PROCEEDINGS OF THE

EWGT2006
INTERNATIONAL JOINT CONFERENCES

11th Meeting of the EURO Working Group on Transportation
Advances in Traffic and Transportation Systems Analysis

and

Extra EURO Conference on
Handling Uncertainty in Transportation:
Analyses, New Paradigms, Applications and Comparisons

September 27 - 29, 2006
TECHNICAL UNIVERSITY OF BARI
BARI - ITALY

Organized by:
Transportation Program – Technical University of Bari
EURO Working Group on Transportation

Edited by The Local Organising Committee
Preface to the Proceedings of the EWGT2006 Joint Conferences

This volume contains the Proceedings of the 11th Meeting of the EURO Working Group on Transportation and the Extra EURO Conference. Both Conferences have taken place at the Faculty of Engineering of the Technical University of Bari, Italy, during September 27-29, 2006.

In the era of rapidly changing technologies and growing globalization of markets, the movements of population and goods have got an unquestionable importance. New methodologies, theories and approaches, interesting case studies addressed to solve complex transportation and logistics problems demonstrate the constant interest of researchers in managing the resources related to transportation, and in improving rationally their effectiveness.

The Conferences have received this interest, covering with innovative and original contributions a wide area of transportation science, from Technological Innovation in transport modes to Transport in Developing Countries, from Freight Transportation and Logistics to Environment and Sustainability.

The response of scientific world demonstrates as well the clear interest that the Conferences aroused: 129 papers of high scientific level were selected, based on opinions of reviewers from Program Committees.

In general, papers present theoretical models; nevertheless, some papers are more oriented to practitioners, giving interesting results obtained by applications of known technologies.

Some of submitted papers, selected by a peer reviewing process performed by the International Committee of the Conferences, will be published on special issues of prominent scientific journals.

We would like to thank all the participants and contributors as well as the Scientific and Organizing Committee for their valuable work in making the Conferences successful.

In addition, we would like to acknowledge the work of the reviewers, whose effort ensured high scientific level of the Conferences.

Finally, we acknowledge the tremendous work of the Organizing Committee members, and especially of the Conference Secretariat in the preparation of the events. Without their peculiar effort, the Conferences could not take place. We are also grateful to the Faculty of Engineering of the Technical University of Bari for the logistic support and to Dr. Maurizio Bielli for his help in organizing the EWGT Meeting.

Bari, September 2006

On behalf of the Organizing Committee

Mauro Dell’Orco and Dino Sassanelli
BACKGROUND OF THE EVENT

Since the first Meeting, held in Landshut (Germany) in 1992, to the last one – the 10th, held in Poznan (Poland) in 2005, the meetings of the EURO Working Group on Transportation represent for scientists all over the world a moment for sharing research results and perspectives on the field of transportation, and encouraging joint researches in a friendly and festive atmosphere.

On the basis of the successful 9th edition, the 11th Meeting of the EURO Working Group on Transportation (EWGT) “Advances in Traffic and Transportation Systems Analysis” was held again in the Politecnico di Bari - Technical University of Bari (Italy), with the main targets of providing a forum to share information and experiences of research activities, and promoting the cooperation among different institutions and organizations in the field of traffic and transportation systems. Moreover, as the Secretariat of pan-European corridor No. 8 is in Bari, the choice of this town was to draw the attention of both transportation researchers and public authorities to the challenge offered by new links across the Mediterranean Sea, and to their related problems and perspectives. The topic of this Meeting is focused on the recent advances in Traffic and Transportation Systems research, covering a wide range of different research themes.

To enhance for all participants the opportunity of introducing new technologies, new ideas, and new paradigms, another scientific event was organised, the Extra EURO Conference "Handling Uncertainty in Transportation: Analyses, New Paradigms, Applications and Comparisons", focusing on new technologies to face Uncertainty in Transportation field.

AIMS OF THE CONFERENCES

The Transportation Systems analysis and modelling have got, especially in recent years, a strong development. The 11th Meeting of the EURO Working Group on Transportation “Advances in Traffic and Transportation Systems Analysis” put together prominent scientists and young researchers to examine new technologies and models to face Transportation problems.

In this meeting, papers have been focused on many topics related to the field of transportation engineering and systems, i. e.:

- Decision Support Systems
- Freight Transportation and Logistics
- Intelligent Transportation Systems
- Safety, Security and Emergency in Transportation
- Technological Innovation in transport modes
- Traffic analysis and control
- Transport Modelling
- Transport Policy
- Transport Systems: Operation, Management, Control and Maintenance
- Transportation Planning and Design
- Transportation, Environment, Development and Sustainability
- Transport in Developing Countries
As for the Extra EURO Conference "Handling Uncertainty in Transportation: Analyses, New Paradigms, Applications and Comparisons", authors were invited to submit original papers that address uncertainty, information, and choices when modelling traffic and transportation systems.

The Program Committee was interested in papers concerning both theory and innovative applications to transportation engineering and planning problems. The considered topics were:

- Understanding Travellers’ Behavioural Processes: New Theories and Paradigms
- Information and Travel Behaviour
- Travel Demand Forecasting
- Users’ choice behaviour
- Intelligent Transportation Systems
- Philosophy of Transportation Science
- Artificial Intelligence and Soft Computing in Traffic and Transportation

ORGANIZATION OF THE EWGT2006 JOINT CONFERENCE

PRESIDING HONORARY CHAIR OF THE CONFERENCES:
Giulio Erberto Cantarella, University of Salerno, Italy

SCIENTIFIC CO-ORDINATOR OF THE CONFERENCES:
Domenico Sassanelli, DVT - Technical University of Bari

in collaboration with:
Mario Binetti, DVT- Technical University of Bari
Mauro Dell’Orco, DVT - Technical University of Bari
Michele Ottomanelli, DIASS -Technical University of Bari

INTERNATIONAL PROGRAM COMMITTEES

The 11th Meeting of EWG on Transportation

CO-CHAIRS:
Giulio Erberto Cantarella, University of Salerno, Italy
Domenico Sassanelli, Technical University of Bari, Italy

MEMBERS:
Jaume Barceló, Technical University of Catalonia, Spain
Francisco García Benitez, Universidad de Sevilla, Spain
Maurizio Bielli, National Research Council, IASI, Italy
Dino Borri, Technical University of Bari, Italy
Roberto Camus, University of Trieste, Italy
Ennio Cascetta, University of Napoli “Federico II”, Italy
Pasquale Colonna, Technical University of Bari, Italy
The Extra EURO Conference

CO-CHAIRS:
Mauro Dell’Orco, Technical University of Bari, Italy
Shinya Kikuchi, Virginia Polytechnic Institute and State University, USA
Dušan Teodorović, University of Belgrade, Serbia

MEMBERS
Andrzej Adamski, Cracow University of Technology, Poland
Erel Avineri, University of the West of England, UK
Michael G. H. Bell, Imperial College London, UK
Michel Bierlaire, EPFL - Swiss Institute of Technology, Switzerland
David Boyce, University of Illinois at Chicago, USA
Mark Dougherty, Dalarna University, Sweden
Jorge Freire de Sousa, University of Porto, Portugal
Liping (Lee) Fu, University of Waterloo, USA
Jonathan Gifford, George Mason University, USA
Chaug-ing Hsu, National Chiao Tung University, Taiwan
Hirotaka Koike, Utsunomiya University, Japan
Hani Mahmassani, University of Maryland, College Park, USA
Angel Marin, Universidad Politécnica de Madrid, Spain
Maria Alice Prudêncio Jacques, Universidade de Brasilia, Brasil
Leena Suhl, University of Paderborn, Germany
Nobuhiro Uno, University of Kyoto, Japan
Henk van Zuylen, Delft University, Netherlands
Jacek Zak, Poznan Technical University, Poland
END-USERS COMMITTEE
Government of Apulia Region
Province of Bari
Municipality of Bari

EWGT2006 LOCAL ORGANIZING COMMITTEE
Mario Binetti, Technical University of Bari
Francesco Civitella, Technical University of Bari
Edoardo De Liddo, Technical University of Bari
Mauro Dell’Orco, Technical University of Bari
Michele Ottomanelli, Technical University of Bari

EURO Working Group on Transportation Coordinator
Maurizio Bielli
Institute of Systems Analysis and Informatics "A. Ruberti".
National Research Council of Italy, Rome
Viale Manzoni 30 - 00185 Rome, Italy
Phone/Fax +39 0677 161
e-mail: bielli@iasi.cnr.it
EWGT web page: http://www.iasi.rm.cnr.it/ewgt/home.html

EWGT2006 Conference Secretariat
Michele Ottomanelli
DVT – Politecnico di Bari
Via Orabona, 4 - 70125 BARI
Phone +39 080596 3380 – Fax. +39 080596 3329
e-mail: m.ottomanelli@poliba.it
web: http://www.poliba.it/ewgt2006
**Proceedings of the EWGT2006 Joint Conferences**

**Articles Index**

<table>
<thead>
<tr>
<th>Id</th>
<th>Title</th>
<th>Authors</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>001</td>
<td>Multiple Objective Signal Control for Urban Environments</td>
<td>Sonal AHUJA, Michael G.H. BELL, Tom VAN VUREN</td>
<td>1</td>
</tr>
<tr>
<td>002</td>
<td>Do pedestrians efficiently use overpass-pedestrian-crossing-facilities in crossing urban streets? Usage assessment and planning parameters</td>
<td>Hassan Abdelzaher Hassan MAHDY</td>
<td>4</td>
</tr>
<tr>
<td>003</td>
<td>Intelligent Transportation Systems: Role of Sensors in Traffic Management</td>
<td>Aman MOUDGIL, Gaurav MOUDGIL, Munish BANSAL</td>
<td>8</td>
</tr>
<tr>
<td>004</td>
<td>Effect of external cost internalisation on transit design problems</td>
<td>Mariano GALLO, Luca D’ACIERNO, Bruno MONTELLA</td>
<td>16</td>
</tr>
<tr>
<td>005</td>
<td>Experimental Study on Data Driven Public Transport Information System in Dhaka City Road Network</td>
<td>Moshiuzzaman MAHMUD, Yeasir RAHUL</td>
<td>22</td>
</tr>
<tr>
<td>006</td>
<td>A multinomial logit for service quality measurement</td>
<td>Laura EBOLI, Gabriella MAZZULLA</td>
<td>30</td>
</tr>
<tr>
<td>007</td>
<td>Modeling the impact of travel information on activity-travel decisions under uncertainty: heuristics and example</td>
<td>Zhongwei SUN, Theo ARENTZE, Harry J.P TIMMERMANS</td>
<td>38</td>
</tr>
<tr>
<td>008</td>
<td>Using neural networks for estimating saturation flow rates at signalized intersections</td>
<td>Isabela QUEIROZ, Maria Alice Prudêncio JACQUES, Mark DOUGHERTY, Li WEIGANG</td>
<td>47</td>
</tr>
<tr>
<td>009</td>
<td>Two-phased parking choice method for pretrip parking guidance system</td>
<td>Ji YANJIE, Deng WEI, Wang WEI</td>
<td>53</td>
</tr>
<tr>
<td>010</td>
<td>A stochastic user equilibrium (SUE) algorithm based on ant colony optimization (ACO)</td>
<td>Luca D’ACIERNO, Bruno MONTELLA, Fortuna DE LUCIA</td>
<td>59</td>
</tr>
</tbody>
</table>
Id: 011
An activity-based system of models using RP and SP methodologies
MAZZULLA Gabriella. ................................................................. 67

Id: 015
A web-based data source platform for modelling and simulation
Marc Philipp MISKA, Hans VAN LINT. ........................................... 74

Id: 016
Dynamic selection of fuzzy signal controllers
Maria Alice Prudêncio JACQUES, Michelle ANDRADE. ..................... 78

Id: 017
Physical and operational design in transit network
Luigi DELL’OLIO, Jose Luis MOURA, Angel IBEAS. ......................... 84

Id: 018
DSS for emergency management on motorway networks
Nenad MLADINEO, Snjezana KNEZIC, Niksa JAJAC. ......................... 90

Id: 019
Using mechanistic design procedures of pavements to optimize transportation planning
Swetha KESIRAJU, Hussain BAHIA. ............................................. 96

Id: 020
Metropolitan areas polarisation and local railway morphology
Marino de LUCA, Francesca PAGLIARA. ......................................... 105

Id: 021
Towards a secure railway transport system
Giuseppe SCIUTTO, Mirco BERRUTTI, Diego MAZZINI, Cristian VEGLIA. 110

Id: 022
Optimal routing during emergency situations by means of prescriptive advices on VMS displays
Frank ZUURBIER, Hans van LINT. .............................................. 118

Id: 023
A hyper-network model for simulating park-and-ride trips
Luca D’ACIERNO, Mariano GALLO, Bruno MONTELLA. ...................... 125

Id: 024
Research on users’ behavior changes during the tram stop-sharing
Shuichi MATSUMOTO, Hironao KAWASHIMA. ................................. 130

Id: 025
Maximization of acoustical airport capacity: parameters and variables
Nicola GUALANDI, Luca MANTECCHINI. ...................................... 137

Id: 026
A new method for offset optimization in urban road networks
Bernhard FRIEDRICH, Essam ALMASRI. ..................................... 144
Validating nested logit models subject to prognosis constraints
Francisco G. BENITEZ, Javier VAZQUEZ. ............................................................. 151

Detection algorithms for intentional car following on smart networks
Banihan GUNAY, Julian HINE. ........................................................................... 159

Path choice modelling for freight road transport: a model for national level
Francesco RUSSO, Antonino VITETTA, Agata QUATTRONE. ............................. 165

Exploring a guidance process for using the public administration genome: an application for transportation in a natural disaster situation
John DICKEY ..................................................................................................... 178

An advanced methodological approach for performance evaluations of alternative functional structures of a ro-ro terminal
Domenico GATTUSO, Antonella POLIMENI, Paolo IARIA. ................................. 183

Experimental research for the evaluation of driver speed behaviour on rural roads
Gianluca DELL’ACQUA, Domenico ABATE. ........................................................ 191

Goods delivery distribution in Rome: an approach to the system reorganization
Antoni MUSSO, CORAZZA Maria Vittoria. .......................................................... 199

Travel time estimation using toll collection data
Francesc SORIGUERA, Leif THORSON, Francesc ROBUSTÈ. .............................. 207

A probit model for departure time choice
Mike MAHER, Andrea ROSA. ........................................................................... 213

Network of magnetic detectors for enhancing airport safety
Kosmas DIMITROPOULOS, Nikos GRAMMALIDIS, Ioannis GRAGOPOULOS, Niovi PAVLIDOU; Thomas HEUER, Haibin GAO, Constanze STOCKHAMMER, Uwe HARTMANN ........................................................................................................ 219

A GIS approach to optimal localization of reserved parking spaces for urban freight distribution
Giuseppe SALVO, Gianfranco AMATO, Pietro ZITO. ......................................... 226

Start-up lost time at urban traffic light junctions
Sergio D’ELIA, Domenico Walter E. MONGelli. ............................................. 232
The influence of advanced traveller information systems (ATIS) on public transport demand
Gianfranco AMATO, Salvatore AMOROSO, Pietro ZITO. ................................. 238

The impact of weather on urban travel demand in the Netherlands
Eric VAN BERKUM, Wendy WEIJERMARS, Anton HAGENS. ......................... 245

Feasibility of CARSHARE in mid-size metropolitan areas in the United States
Ardeshir FAGHRI, Catherine ADAM. ....................................................... 253

Differences among route flow solutions for the user-equilibrium traffic assignment problem
Amos LUZON, Hillel BAR-GERA. ............................................................ 258

Evaluation of on-street parking scheme using virtual reality traffic experiment system
Shinji TANAKA, Masao KUWAHARA. ....................................................... 262

A computer vision based system for off-line quality control of a chemically coated component of a diesel fuel injection pump
Giuseppe MASTRONARDI, Massimiliano RONCON, Vitoantonio BEVILACQUA, Lucia CARIELLO. ................................................................. 269

An ant colony optimization approach for the capacitated vehicle routing problem with simultaneous delivery and pick-up
Bülent ÇATAY. ................................................................. 275

Implementation of an asymmetric network equilibrium problem with detailed representation of unsignalized and signalized urban intersections
Marino LUPI, Federico RUPI, Guido ROSSI. ........................................... 283

Analysis on the mechanism of congestion on the merging section of urban expressway using video image data
Fumitaka KURAUCHI, Nobuhiro UNO, Akito HIGATANI, Yasunori IIDA. ............. 290

On the convergence of weighting methods in multi-attribute assessment of transport projects
Paolo DELLE SITE, Francesco FILIPPI. ............................................... 298

Analysis of the public transport supply in a local level
Guillem ALSINA, Magín CAMPOS, Francese ROBUSTÉ. .......................... 304
Id: 051
A simulation tool for evaluation strategies for productivity improvement of container terminals
DEGANO Cristiana, FICARA Pierluigi, MARANGON Flavio, MODICA Pasquale, and VALFRE Giulio. .......................................................... 309

Id: 052
An ensemble regression approach for bus trip time prediction
João M. MOREIRA, Jorge Freire de SOUSA, Alípio M. JORGE, Carlos SOARES. .. 317

Id: 053
Willingness to pay for access time saving: some evidences
Angel IBEAS, Luigi DELL’OLIO, Jose Luis MOURA. ............................... 322

Id: 054
The uncertainty of delays at optimal pre-timed and vehicle actuated signals
Francesco VITI, Henk J. Van ZUYLEN. ............................................. 328

Id: 055
Two lane highways - a microscopic traffic simulation model
Caroline C. PECKER, Helena B. B. CYBIS. ....................................... 334

Id: 056
Time windows for scheduled trips in multiple depot vehicle scheduling
Natalia KLIEWER, Stefan BUNTE, Leena SUHL. .................................. 340

Id: 058
An algorithm for the pick-up and delivery problem with rear-loading constraints
Giovanni RIGHINI, Federico FICARELLI. ........................................... 347

Id: 059
A dynamic activity based travel demand model to assess Stockholm congestion charging scheme - estimation using Markov chain Monte Carlo and reinforcement learning
Anders KARLSTROM ................................................................. 352

Id: 060
An agent-based approach to combined route and mode choice
Gustavo KUHN ANDRIOTTI, Franziska KLÜGL, Guido RINDSFÜSER. .......... 358

Id: 061
Level crossing risk assessment: analysis and simulation of system
Alessandro BALDASSARRA, Stefano IMPASTATO, Gabriele MALAVASI. .... 365

Id: 062
A time-space network based approach for integrated vehicle and crew scheduling in public transport
Vitali GINTNER, Ingmar STEINZEN, Leena SUHL. .................................. 371

Id: 063
On the potential of social-psychological aspects in modelling travellers’ change of behaviour
Yos SUNITIYOSO, Erel AVINERI, Kiron CHATTERJEE. .......................... 378
Strategic modelling and hierarchical approach to assess the land use and transport effects of transport projects in Madrid
Paul PFAFFENBICHLER, Paula VIEIRA, Andrés MONZÓN, Daniel DE LA HOZ, Ana PARDEIRO.

A contribution to predicting the modal split in urban passenger transport under incomplete data
Jan BRÚHA, Hana FOLTÝNOVÁ.

Cost-benefit analysis of changes in cycling demand and its use in cost-benefit analysis: a case study of Pilsen
Hana FOLTÝNOVÁ and Markéta BRAUN KOHLOVÁ.

Predicting a suitable capacity estimation method for single-lane roundabouts in Izmir, Turkey
Serhan TANYEL, Mustafa ÖZUYSAL, Pelin ÇALIŞKANELLİ, Türkay BARAN.

A methodological approach to optimise temporal integration between regional public transit lines
Domenico GATTUSO, Antonella POLIMENI.

Effect of heavy vehicles at signalized intersections in Turkey
Serhan TANYEL, Gülden (CANSEVEN) KALFAOĞLU, Burak ŞENGÖZ, Pelin ÇALIŞKANELLİ, Mustafa ÖZUYSAL.

A sequential model based on thresholds of log-normal acceptability of the attributes
Massimiliano GASTALDI, Riccardo ROSSI, Romeo VESCOVI.

An online learning solution for data driven freeway travel time prediction models
Hans VAN LINT.

Solving a dynamic traffic assignment problem with spillback congestion by means of the IGSM algorithm
Guido GENTILE, Lorenzo MESCHINI.

Controlling congestion on platforms a case study with London's underground network
Hiroshi SHIMAMOTO, Jan-Dirk SCHMÖCKER, Fumitaka KURAUCHI, Michael G. H. BELL.

Bounded rationality and transportation behavior - Lessons for public policy
Jonathan GIFFORD, Cristina CHECHERITA.
Id: 075  
Analysis of path flow fluctuation on urban expressway using dynamic path flow estimation  
Jin-Young KIM, Fumitaka KURAUCHI, Nobuhiro UNO, Yasunori IIDA. .......................... 458

Id: 076  
Better understanding the potential market of transit ridership and services  
Kevin J. KRIZEK, Ahmed M. EL-GENEIDY. ................................................................. 466

Id: 077  
A GLS estimator for combined path choice model and O-D Matrix aggregate calibration on congested network  
Michele OTTOMANELLI, Leonardo CAGGIANI, Domenico SASSANELLI ....................... 475

Id: 078  
The usage of seat belts in Turkey  
Şafak BİLGİÇ, Polat YALINIZ, Murat KARACASU. ...................................................... 487

Id: 079  
Transport network design under risk-averse traffic assignment: a game theory approach  
Liam O’BRIEN, Wai Yuen SZETO, Margaret O’MAHONY. ........................................ 494

Id: 080  
SITLUM - the Strathclyde integrated transport/land-use model  
Andrea ARAMU, Andrew ASH, Jim DUNLOP, David SIMMONDS. ............................... 503

Id: 081  
Analysis of the factors affecting a regional airport catchment area  
Maria Nadia POSTORINO, Alberto ANDREONI. ......................................................... 511

Id: 082  
A study on the effect of removal center lines of the roads to extend the width of sideways of the roads  
Seiji HASHIMOTO, Toshiomi OGURA, Koji IZUHARA. .............................................. 518

Id: 083  
Carrot and stick tactic: parking taxation to subsidize bus  
Vittorio ASTARITA, Marco SALATINO. ................................................................. 523

Id: 084  
An incident detection method based on the estimation of traffic flow parameters from instrumented vehicles counts  
Vittorio ASTARITA, Giuseppe GUIDO. ................................................................. 530

Id: 085  
Traffic conflict analysis based on moving coordinates system with syncedronized multicameras  
Hiroshi WAKABAYASHI, Shin-ya MURAMATSU. ....................................................... 535

Id: 086  
Transferability of destination choice models with land use variables: a case study on the metropolitan area of Naples  
Luigi BIGGIERO, Cinzia CIRILLO, Gaetano GALANTE. ........................................... 541
Id: 087
A stochastic approach to delays at intersections
Henk J. VAN ZUYLEN. ......................................................... 545

Id: 088
Tracking waves in the LWR model for a progressive dynamic traffic assignment calculation
Vincent HENN, Thomas DURLIN. ........................................... 551

Id: 089
A comparison of macroscopic and microscopic approaches for simulating container terminal operations
Giulio Erberto CANTARELLA, Armando CARTENI, Stefano de LUCA. ............... 556

Id: 090
Object-oriented modeling and simulation of the transportation system
Hanna SAWICKA, Jacek ZAK. .......................................................... 559

Id: 091
Innovative technologies supporting the driving crew in degraded running conditions
Giuseppe SCIUTTO, Marino LUCCHINI, Diego MAZZINI, Cristian VEGLIA. ........... 567

Id: 092
Intersection modelling based on the second order traffic flow model (ARZ)
Salim MAMMAR, Jean-Patrick LEBACQUE, Habib HAJ SALEM. ......................... 570

Id: 093
Risk index model for real time prediction of potential traffic crash on urban motorway
Habib HAJ SALEM, Christophe RAMANANJAONA, Jean-Patrick LEBACQUE, Salim MAMMAR. .............................................................. 581

Id: 094
Existence and stability of equilibriums in regional tax competition
Megan M. KHOSHYARAN, Jean-Patrick LEBACQUE. .......................................... 587

Id: 095
Traffic flow modelling and inter-change vehicle spacing control in the Mont Blanc Tunnel
Vincent HENN, Jean-Baptiste LESORT. .................................................. 595

Id: 096
A shockwave-based methodology for reliable estimation of origin-destination travel-time in urban areas
Manoj K. JHA, Franklin OKONKWO, Wolfgang KUHN. ....................................... 600

Id: 097
PIACON: Robust vehicle trajectory based arterial multi-criteria traffic signal control
Andrzej ADAMSKI. ............................................................................ 608

Id: 099
Modeling Activity choice distribution and road choice behaviour on a network with side constraints
Tai-Yu MA, Jean-Patrick LEBACQUE. ...................................................... 616
On the fractal geometry of the road network and its relevance to transport modelling
Erel AVINERI .......................................................... 624

Stochastic delay at traffic signals: influence of demand patterns
Estelle CHEVALLIER, Ludovic LECLERCQ ......................... 630

Car pooling clubs: viability of the concept applied to Lisbon’s metropolitan area, using GIS and optimization techniques
Gonçalo CORREIA, José Manuel VIEGAS ........................ 639

A dynamic model of point-queue process
Hilmi Berk ÇELIKOĞLU, Mauro DELL'ORCO, Ergün GEDIZLIOĞLU ........ 647

A dynamic node model with a quadratic travel time function based link model
Hilmi Berk ÇELIKOĞLU ............................................. 658

Container stowage optimization with crane utilization balancing: modelling and heuristics
Masoud HAMEDI, Ali HAGHANI, Saini YANG ....................... 667

The 2-period probabilistic TSP: a Markov decision model
Giovanni ANDREATTA, Guglielmo LULLI ........................ 672

Conflict resolution strategies in real time train traffic management
Andrea D’ARIANO, Dario PACCIARELLI, Marco PRANZO .......... 677

Urban transport: European experiences and the Hong Kong case
Elisabetta VENEZIA, Justyna KARAKIEWICZ ......................... 682

Auction based congestion pricing: The basic idea
Dušan TEODOROVIĆ, Konstantinos TRIANTIS, Yueqin ZHAO, Snežana MLADENOVIĆ . 688

Dealing with uncertainties in transport policy making: a new paradigm and approach
J.W.G.M. VAN DER PAS, D.B. AGUSDINATA, W. WALKER, V.A.W.J.MARCHAU .... 694

Providing public participation in transportation decision using span (social participatory allocative networks) method
Murat KARACASU, Şafak BİLGİÇ .................................. 702
Id: 112
Minimum redeployment model for dynamic location of ambulances
Hari K. RAJAGOPALAN, Cem SAYDAM. ............................................................ 708

Id: 113
Dynamic set partitioning applications: a heuristic approach
Luca COSLOVICH, Raffaele PESENTI, Walter UKOVICH. ............................. 714

Id: 114
Handling demand and network capacity uncertainties during major evacuations
Kaan OZBAY, Mustafa Anil YAZICI. ................................................................. 720

Id: 115
Validation and comparison of choice models
Giulio Erbeto CANTARELLA, Stefano DE LUCA, Viviana FEDELE. .............. 728

Id: 116
Decomposition methodology to solve taxi planning
Angel G. MARIN. ............................................................................................. 733

Id: 117
Allocation planning of handling devices for barges unloading
Milorad VIDOVIC, Katarina VUKADINIVIC. .................................................. 740

Id: 118
Univariate and multivariate ARIMA models for air transportation demand estimation: A study about regional air transport
Alberto ANDREONI, Maria Nadia POSTORINO. .............................................. 748

Id: 119
Crossing time micro-analysis: different crossing scenarios for different zones of approach to the crossroads
Edoardo DE LIDDO. .......................................................................................... 756

Id: 120
Simulation and validation of parking and insertion kinematics cycles
Borja BELTRAN, Stefano CARRESE, Emanuele NEGRENTI. ....................... 764

Id: 121
Collaborative demand and supply networks: modelling, design and performances evaluation
Maurizio BIELLI, Mariagrazia MECOLI. ......................................................... 776

Id: 122
A dynamic network loading model for the evaluation and control of traffic pollution
Maurizio BIELLI, Mauro DELL‘ORCO, Ergün GEDIZLIOĞLU, Murat AKAD. ...... 783

Id: 123
Modelling urban traffic networks by a Consumer/Producer architecture
Marco CALABRESE, Vincenzo DI LECCE, Michele OTTOMANELLI, Domenico SASSANELLI. ............................................................... 788
Id: 124
Preserving uncertainty in modeling choice behaviors: from possibility to probability
Mauro DELL'ORCO, Shinya KIKUCHI ................................................................. 793

Id: 125
Petri net model of transhipment activities executed by human operators in a container terminal
Guido MAIONE, Michele OTTOMANELLI .......................................................... 804

Id: 130
Evaluation of O-D multimodal paths in urban networks
Daniela AMBROSINO, Anna SCIOMACHEN ..................................................... 815

Id: 131
Models and methods for route guidance problems on urban and regional networks
Giuseppe BRUNO, Gennaro IMPROTA, Antonio SFORZA .................................. 823

Id: 132
Infomobility and goods distribution in urban and regional areas
Alberto COLORNI, Roberto WOLFLER CALVO, Renato DE LEONE, Paolo DELL’OLMO, Nicoletta RICCIARDI, Daniela AMBROSINO, Anna SCIOMACHEN, Giuseppe BRUNO, Gennaro IMPROTA, Antonio SFORZA ............................................. 830

Id: 133
An integrated information and management system for traditional and demand-responsive mobility services in the city of Milan
Diego CICCARELLI, Alberto COLORNI, Alessandro LUÉ, Valentina MORROCCHI, Damiano ROSSI, Roberto WOLFLER CALVO ......................................................... 842

Id: 134
A web-based software tool for the management of commuters car pooling in the Milan area
Alberto COLORNI, Diego CICCARELLI, Alessandro LUÉ, Valentina MORROCCHI ........ 845

Id: 135
Interaction between signal settings and traffic flow patterns on road networks
Gaetano FUSCO, Guido GENTILE, Lorenzo MESCHINI, Maurizio BIELLI, Giovanni FELICI, Ernesto CIPRIANI, Stefano GORI, Marialisa NIGRO ........................................ 851

Id: 136
Transit-based land use and transport network design
GORI Stefano, NIGRO Marialisa and PETRELLI Marco ...................................... 865

Id: 141
Defining share of Park and Ride trips using Fuzzy Inference System
Andrzej SZARATA ................................................................. 873

Id: 142
ITS framework architecture in a multimodal transport system
Mario BINETTI, Paola AMORUSO, Margherita MASCIA ...................................... 884
Proceedings of the
EWGT2006
Joint Conferences

Authors Index

1. Domenico ABATE, University of Naples “Federico II” - Italy (Paper: 032)
2. Andrzej ADAMSKI, Cracow AGH University of Science and Technology - Poland, (Paper: 097)
4. Sonal AHUJA, Imperial College London - United Kingdom (Paper: 001)
5. Murat AKAD, Istanbul Technical University - Turkey (Paper: 122)
6. Essam ALMASRI, University of Hannover - Germany (Paper: 026)
7. Guillem ALSINA, CENIT - Center for Innovation in Transport - Spain (Paper: 050)
8. Gianfranco AMATO, University of Palermo - Italy (Papers: 037, 039)
9. Daniela AMBROSINO, Università di Genova - Italy (Papers: 130, 132)
10. Salvatore AMOROSO, University of Palermo - Italy (Paper: 039)
11. Paola AMORUSO, DVT - Technical University of Bari - Italy (Papers: 142, 143)
12. Michelle ANDRADE, University of Brasilia - Brazil, (Paper: 016)
13. Giovanni ANDREATTA, University of Padova - Italy (Paper: 106)
14. Alberto ANDREONI, Mediterranea University of Reggio Calabria - Italy (Papers: 081, 118)
15. Andrea ARAMU, David Simmonds Consultancy Ltd - United Kingdom (Paper: 080)
16. Theo ARENTZE, Eindhoven University of Technology - Netherlands (Paper: 007)
17. Andrew ASH, Transport Research Laboratory, Crowthorne House - United Kingdom (Paper: 080)
18. Vittorio ASTARITA, University of Calabria - Italy (Paper: 083, 084)
19. Erel AVINERI, University of the West of England - United Kingdom (Papers: 063, 100)
20. Hussain BAHIA, The University of Wisconsin- Madison - USA (Paper: 019)
21. Alessandro BALDASSARRA, University of Rome "La Sapienza" - Italy (Paper: 061)
23. Türkay BARAN, Dokuz Eylül University - Turkey (Paper: 067)
24. Antonio BARBARA, Salerno Container Terminal S.p.a (SCT) - Italy (Paper: 051)
25. Hillel BAR-GERA, Ben-Gurion University - Israel (Paper: 042)
26. Michael G. H. BELL, Imperial College London - United Kingdom (Papers: 001, 073)
27. Borja BELTRAN, Università degli Studi di Roma TRE - Italy (Paper: 120)
28. Francisco G. BENITEZ, University of Sevilla - Spain (Paper: 027)
29. Mirco BERRUTTI, Consorzio Nazionale Interuniversitario per i Trasporti e la Logistica - Italy (Paper: 021)
30. Vitantonio BEVILACQUA, DEE-Politecnico di Bari - Italy (Paper: 044)
31. Maurizio BIELLI, IASI - Consiglio Nazionale delle Ricerche - Italy (Papers: 121, 125, 135)
32. Luigi BIGGIERO, University of Naples "Federico II" - Italy (Paper: 086)
33. Şafak BİLGİÇ, Eskisehir Osmangazi University - Turkey (Papers: 078, 111)
34. Mario BINETTI, DVT - Technical University of Bari - Italy (Papers: 142, 143)
35. Markéta BRAUN KOHLOVÁ, Charles University - Czech Republic, (Paper: 066)
36. Jan BRŮHA, Charles University Environment Center - Czech Republic (Paper: 065)
37. Giuseppe BRUNO, Università degli Studi di Napoli “Federico II” - Italy (Papers: 131, 132)
38. Stefan BUNTE, University of Paderborn - Germany (Paper: 056)
39. Leonardo CAGGIANI, DVT - Technical University of Bari - Italy (Paper: 077)
40. Marco CALABRESE, DIASS - Technical University of Bari - Italy (Paper: 123)
41. Pelin ÇALIŞKANELLİ, Dokuz Eylül University - Turkey (Papers: 067, 069)
42. José Magín CAMPOS CACHEDA, Center for Innovation in Transport - Spain (Paper: 050)
43. Giulio Erberto CANTARELLA, University of Salerno - Italy (Papers: 089, 115)
44. Lucia CARIELLO, DEE-Politecnico di Bari - Italy (Paper: 044)
45. Stefano CARRESE, Università degli Studi di Roma TRE - Italy (Paper: 120)
46. Armando CARTENİ, University of Salerno - Italy (Paper: 089)
<table>
<thead>
<tr>
<th></th>
<th>Name</th>
<th>Affiliation</th>
<th>Paper(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>47</td>
<td>Bülent ÇATAY</td>
<td>Sabanci University, Turkey</td>
<td>045</td>
</tr>
<tr>
<td>48</td>
<td>Adam CATHERINE</td>
<td>University of Delaware, USA</td>
<td>041</td>
</tr>
<tr>
<td>49</td>
<td>Hilmi Berk ÇELIKOĞLU</td>
<td>Technical University of Istanbul, Turkey</td>
<td>103, 104</td>
</tr>
<tr>
<td>50</td>
<td>Kiron CHATTERJEE</td>
<td>University of the West of England, United Kingdom</td>
<td>063</td>
</tr>
<tr>
<td>51</td>
<td>Cristina CHECHERITA</td>
<td>George Mason University, USA</td>
<td>074</td>
</tr>
<tr>
<td>52</td>
<td>Estelle CHEVALLIER</td>
<td>LICIT (INRETS-ENTPE), France</td>
<td>101</td>
</tr>
<tr>
<td>53</td>
<td>Diego CICCARELLI</td>
<td>Poliedra, Politecnico di Milano, Italy</td>
<td>133, 134</td>
</tr>
<tr>
<td>54</td>
<td>Ernesto CIPRIANI</td>
<td>Università degli Studi di Roma TRE, Italy</td>
<td>135</td>
</tr>
<tr>
<td>55</td>
<td>Cinzia CIRILLO</td>
<td>University of Namur, Belgium</td>
<td>086</td>
</tr>
<tr>
<td>56</td>
<td>Alberto COLORNI</td>
<td>Politecnico di Milano, Italy</td>
<td>132, 133, 134</td>
</tr>
<tr>
<td>57</td>
<td>Maria Vittoria CORAZZA</td>
<td>University of Rome &quot;La Sapienza&quot;, Italy</td>
<td>033</td>
</tr>
<tr>
<td>58</td>
<td>Gonçalo CORREIA</td>
<td>Universidade Técnica de Lisboa, Portugal</td>
<td>102</td>
</tr>
<tr>
<td>59</td>
<td>Luca COSLOVICH</td>
<td>DEEI, University of Trieste, Italy</td>
<td>113</td>
</tr>
<tr>
<td>60</td>
<td>Helena B. B CYBIS</td>
<td>LASTRAN, Federal University of Rio Grande do Sul, Brazil</td>
<td>055</td>
</tr>
<tr>
<td>61</td>
<td>Luca D'ACIERNO</td>
<td>University of Naples &quot;Federico II&quot;, Italy</td>
<td>004, 010, 023</td>
</tr>
<tr>
<td>62</td>
<td>Andrea D'ARIANO</td>
<td>Delft University of Technology, Netherlands</td>
<td>107</td>
</tr>
<tr>
<td>63</td>
<td>Daniel DE LA HOZ</td>
<td>Technique University of Madrid, Spain</td>
<td>064</td>
</tr>
<tr>
<td>64</td>
<td>Renato DE LEONE</td>
<td>University of Camerino, Italy</td>
<td>132</td>
</tr>
<tr>
<td>65</td>
<td>Edoardo DE LIDDO</td>
<td>DVT, Technical University of Bari, Italy</td>
<td>119</td>
</tr>
<tr>
<td>66</td>
<td>Marino de LUCA</td>
<td>University of Naples &quot;Federico II&quot;, Italy</td>
<td>020</td>
</tr>
<tr>
<td>67</td>
<td>Stefano de LUCA</td>
<td>University of Salerno, Italy</td>
<td>089, 115</td>
</tr>
<tr>
<td>68</td>
<td>Fortuna DE LUCIA</td>
<td>University of Naples &quot;Federico II&quot;, Italy</td>
<td>010</td>
</tr>
<tr>
<td>69</td>
<td>Jorge Freire de SOUSA</td>
<td>University of Porto, Portugal</td>
<td>052</td>
</tr>
<tr>
<td>70</td>
<td>Cristiana DEGANO</td>
<td>Università di Genova, Italy</td>
<td>051</td>
</tr>
<tr>
<td>71</td>
<td>Sergio D'ELIA</td>
<td>Universitiy of Calabria, Italy</td>
<td>038</td>
</tr>
<tr>
<td>72</td>
<td>Gianluca DELL'ACQUA</td>
<td>University of Naples “Federico II”, Italy</td>
<td>032</td>
</tr>
</tbody>
</table>

XXIII
73. Paolo DELLE SITE, University of Rome "La Sapienza" - Italy (Paper: 048)
74. Luigi DELL’OLIO, Universidad de Burgos - Spain (Papers: 017, 053)
75. Paolo DELL'OLMO, University of Rome "La Sapienza" - Italy (Paper: 132)
76. Mauro DELL’ORCO, DVT - Technical University of Bari - Italy (Papers: 103, 124, 122)
77. Vincenzo DI LECCE, DIASS - Technical University of Bari - Italy (Paper: 123)
78. Eugenio DI SCIASCIO, DEE - Technical University of Bari - Italy (Paper: 143)
79. John DICKEY, Virginia Tech - USA (Paper: 030)
80. Kosmas DIMITROPOULOS, Centre for Research and Technology Hellas - Greece (Paper: 036)
81. Mark DOUGHERTY, Dalarna University - Sweden (Paper: 008)
82. Jim DUNLOP, Strathclyde Partnership for Transport - Scotland (Paper: 080)
83. Thomas DURLIN, LICIT (INRETS-ENTPE) - France (Paper: 088)
84. Laura EBOLI, University of Calabria - Italy (Paper: 006)
85. Ahmed M. EL-GENEIDY, University of Minnesota - USA (Paper: 076)
86. Ardeshir FAGHRI, University of Delaware - USA (Paper: 041)
87. Viviana FEDELE, University of Reggio Calabria - Italy (Paper: 115)
88. Giovanni FELICI, Istituto di Analisi dei Sistemi ed Informatica, CNR - Italy (Paper: 135)
89. Pierluigi FICARA, Consorzio TRAIN - Italy (Paper: 051)
90. Federico FICARELLI, DTI-Università degli Studi di Milano - Italy (Paper: 058)
91. Francesco FILIPPI, University of Rome "La Sapienza" - Italy (Paper: 048)
92. Hana FOLTÝNOVÁ, Charles University - Czech Republic (Papers: 065, 066)
93. Bernhard FRIEDRICH, University of Hannover - Germany (Paper: 026)
94. Gaetano FUSCO, University of Rome "La Sapienza" - Italy (Paper: 135)
95. Gaetano GALANTE, University of Naplese "Federico II" - Italy (Paper: 086)
96. Mariano GALLO, University of Sannio - Italy (Papers: 004, 023)
97. Haibin GAO, Institute of Experimental Physics, Saarland University - Germany (Paper: 036)
98. Massimilliano GASTALDI, University of Padova - Italy (Paper: 070)
99. Domenico GATTUSO, University Mediterranea of Reggio Calabria - Italy (Paper: 031, 068)

100. Ergün GEDIZLIOĞLU, Technical University of Istanbul - Turkey (Papers: 103, 122)

101. Guido GENTILE, University of Rome "La Sapienza" - Italy (Paper: 072, 135)

102. Jonathan L. GIFFORD, George Mason University - USA (Paper: 074)

103. Vitali GINTNER, University of Paderborn - Germany (Paper: 062)

104. Stefano GORI, Università degli Studi di Roma TRE - Italy (Papers: 135, 136)

105. Ioannis, GRAGOPOULOS, Centre for Research and Technology Hellas - Greece (Paper: 036)

106. Nicola GUALANDI, DISTART - University of Bologna - Italy (Paper: 025)

107. Giuseppe GUIDO, University of Calabria - Italy (Paper: 084)

108. Banihan GUNAY, University of Ulster - Northern Ireland (Paper: 028)

109. Anton HAGENS, University of Twente - Netherlands (Paper: 040)

110. Ali HAGHANI, University of Maryland - USA (Paper: 105)

111. Habib HAJ-SALEM, INRETS/GRETIA - France (Paper: 092, 093)

112. Masoud HAMEDI, University of Maryland - USA (Paper: 105)

113. Uwe HARTMANN, IEP, Saarland Univeristy - Germany (Paper: 036)

114. Seiji HASHIMOTO, Toyota Transportation Research Institute - Japan (Paper: 082)

115. Vincent HENN, LICIT (INRETS-ENTPE) - France (Paper: 088, 095)

116. Thomas HEUER, IEP, Saarland Univeristy - Germany (Paper: 036)

117. Akito HIGATANI, Kyoto University - Japan (Paper: 047)

118. Julian HINE, University of Ulster - Northern Ireland (Paper: 028)

119. Paolo IARIA, University Mediterranea of Reggio Calabria - Italy (Paper: 031)

120. Angel IBEAS, Universidad de Cantabria - Spain (Papers: 053, 017)

121. Yasunori IIDA, Kyoto University - Japan (Papers: 047, 075)

122. Stefano IMPASTATO, University of Rome "La Sapienza" - Italy (Paper: 061)

123. Gennaro IMPROTA, Università degli Studi di Napoli “Federico II” - Italy (Papers: 131, 132)

124. Koji IZUHARA, Toyota Transportation Research Institute - Japan (Paper: 082)
125. Maria Alice Prudêncio JACQUES, University of Brasilia - Brazil, (Papers: 008, 016)
126. Niksa JAJAC, University of Split - Croatia, (Paper: 018)
127. Manoj K. JHA, Morgan State University - USA (Paper: 096)
128. Alípio JORGE, University of Porto - Portugal, (Paper: 052)
129. Gülten KALFAOĞLU, Proset Software Inc., İzmir - Turkey (Paper: 069)
130. Murat KARACASU, Eskisehir Osmangazi University - Turkey (Papers: 078, 111)
131. Justyna KARAKIEWICZ, University of Hong Kong - Hong Kong, (Paper: 108)
132. Anders KARLSTROM, Royal Institute of Technology - Sweden, (Paper: 059)
133. Hironao KAWASHIMA, Keio University - Japan (Paper: 024)
134. Swetha KESIRAJU, The University of Wisconsin- Madison - USA (Paper: 019)
135. Megan M. KHOSHYARAN, ETC Economics Traffic Clinic - France (Paper: 094)
136. Shinya KIKUCHI, Virginia Polytechnic Institute and State University - USA (Paper: 124)
137. Jin Young KIM, Kyoto University - Japan (Paper: 075)
138. Natalia KLIEWER, University of Paderborn - Germany (Paper: 056)
139. Franziska KLÜGL, University of Wuerzburg - Germany (Paper: 060)
140. Snjezana KNEZIC, University of Split - Croatia (Paper: 018)
141. Kevin J. KRIZEK, University of Minnesota - USA (Paper: 076)
142. Wolfgang KUHN, University of Leipzig - Germany (Paper: 096)
143. Gustavo KUHN ANDRIOTTI, University of Wuerzburg - Germany (Paper: 060)
144. Fumitaka KURAUCHI, Kyoto University - Japan (Papers: 047, 073, 075)
145. Masao KUWAHARA, University of Tokyo - Japan (Paper: 043)
146. Jean-Patrick LEBACQUE, INRETS/GRETIA - France (Paper: 092, 093, 094, 099)
147. Ludovic LECLERCQ, LICIT (INRETS-ENTPE) - France (Paper: 101)
148. Jean-Baptiste LESORT, LICIT (INRETS-ENTPE) - France (Paper: 095)
149. Marino LUCCHINI, Centro Italiano di Eccellenza sulla Logistica Integrata - Italy (Paper: 091)
150. Alessandro LUE', Poliedra - Politecnico di Milano - Italy (Papers: 133, 134)
151. Guglielmo LULLI, University of Milano - Italy (Paper: 106)
152. Marino LUPI, DISTART - University of Bologna - Italy (Paper: 046)
153. Amos LUZON, Ben-Gurion University - Israel (Paper: 042)
154. Tai-Yu MA, INRETS/GRETIA - France (Paper: 099)
155. Hassan Abdelzaher Hassan MAHDY, Ain Shams University - Egypt (Paper: 002)
156. Mike MAHER, Napier University - United Kingdom (Paper: 035)
157. Moshiuzzaman MAHMUD, University of Southampton - United Kingdom (Paper: 005)
158. Guido MAIONE, DIASS - Technical University of Bari - Italy (Paper: 125)
159. Gabriele MALAVASI, University of Rome "La Sapienza" - Italy (Paper: 061)
160. Salim MAMMAR, INRETS/GRETIA - France (Paper: 092, 093)
161. Luca MANTECCHINI, DISTART - University of Bologna - Italy (Paper: 025)
162. Flavio MARANGON, D’Appolonia S.p.a - Italy (Paper: 051)
163. V.A.W.J. MARCHAU, Delft University of Technology – Netherlands (Paper: 110)
164. Ángel G. MARÍN, Universidad Politécnica de Madrid - Spain (Paper: 116)
165. Margherita MASCIA, Faber Maunsell - United Kingdom (Paper: 142)
166. Giuseppe MASTRONARDI, DEE-Politecnico di Bari - Italy (Paper: 044)
167. Shuichi MATSUMOTO, Kochi University of Technology - Japan (Paper: 024)
168. Diego MAZZINI, Centro Italiano di Eccellenza sulla Logistica Integrata - Italy (Papers: 021, 091)
169. Gabriella MAZZULLA, University of Calabria - Italy (Paper: 006, 011)
170. Mariagrazia MECOLI, IASI, Consiglio Nazionale delle Ricerche - Italy (Paper: 121)
171. Lorenzo MESCHINI, University of Rome "La Sapienza" - Italy (Paper: 072, 135)
172. Marc Philipp MISKA, Delft University of Technology - Netherlands (Paper: 015)
173. Snežana MLADENOVIC, University of Belgrade - Serbia (Paper: 109)
174. Nenad MLADINEO, University of Split - Croatia (Paper: 018)
175. Pasquale MODICA, University of Salerno - Italy (Paper: 051)
176. Walter E. MONGELLI, University of Calabria - Italy (Paper: 038)
177. Bruno MONTELLA, University of Naples "Federico II" - Italy (Paper: 004, 010, 023)
178. Andres MONZON, Technique University of Madrid - Spain (Paper: 064)
179. João Mendes MOREIRA, University of Porto - Portugal (Paper: 052)
180. Valentina MORROCCHI, Poliedra - Politecnico di Milano - Italy (Papers: 133, 134)
182. Gaurav MOUDGIL, University of California, Berkeley - USA (Paper: 003)
183. Jose Luis MOURA, Universidad de Cantabria - Spain (Papers: 017, 053)
185. Antonio MUSSO, University of Rome "La Sapienza" - Italy (Paper: 033)
186. Emanuele NEGRENTI, ENEA, ENE-TEC - Italy (Paper: 120)
187. Marialisa NIGRO, Università degli Studi di Roma TRE - Italy (Papers: 135, 136)
188. Liam O'BRIEN, University of Dublin - Republic of Ireland, (Paper: 079)
189. Toshiomi OGURA, Toyota Transportation Research Institute - Japan (Paper: 082)
190. Franklin OKONKWO, Morgan State University - USA (Paper: 096)
191. Margaret O'MAHONY, University of Dublin - Republic of Ireland (Paper: 079)
192. Michele OTTOMANELLI, DIASS - Technical University of Bari - Italy (Papers: 077, 123, 125)
193. Kaan OZBAY, Rutgers University - USA (Paper: 114)
194. Mustafa ÖZUYSAŁ, Dokuz Eylül University - Turkey (Paper: 067, 069)
195. Dario PACCIARELLI, Università degli Studi di Roma TRE - Italy (Paper: 107)
196. Francesca PAGLIARA, University of Naples "Federico II" - Italy (Paper: 020)
197. Ana PARDEIRO, Technique University of Naples Madrid - Spain (Paper: 064)
198. Niovi PAVLIDOU, Centre for Research and Technology Hellas - Greece (Paper: 036)
199. Caroline C. PECKER, Federal University of Rio Grande do Sul - Brazil (Paper: 055)
200. Raffaele PESENTI, DINFO - University of Palermo - Italy (Paper: 113)
201. Marco PETRELLI, Università degli Studi di Roma TRE - Italy (Paper: 136)
202. Paul PFAFFENBICHLER, Vienna University of Technology - Austria (Paper: 064)
203. Agnese PINTO, DEE - Technical University of Bari - Italy (Paper: 143)
204. Antonella POLIMENI, University Mediterranea of Reggio Calabria - Italy (Papers: 031, 068)
205. Antonella POLIMENI, Mediterranea University of Reggio Calabria - Italy (Paper: 068)
206. Maria Nadia POSTORINO, Mediterranea University of Reggio Calabria - Italy (Papers: 081, 118)
207. Marco PRANZO, Università degli Studi di Siena - Italy (Paper: 107)
208. Agata QUATTRONE, Mediterranea University of Reggio Calabria - Italy (Paper: 029)
209. Isabela N.F. QUEIROZ, University of Brasilia - Brazil (Paper: 008)
210. Yeasir RAHUL, American International BD. LTD - United Kingdom (Paper: 005)
211. Hari K. RAJAGOPALAN, University of North Carolina at Charlotte - USA (Paper: 112)
212. Christophe RAMANANJAONA, INRETS/GRETIA - France (Paper: 093)
213. Nicoletta RICCIARDI, University of Rome "La Sapienza" - Italy (Paper: 132)
214. Giovanni RIGHINI, DTI-Università degli Studi di Milano - Italy (Paper: 058)
215. Guido RINDSFÜSER, Emch+Berger AG - Switzerland (Paper: 060)
216. Francesc ROBUSTÉ, Technical University of Catalonia - Spain (Paper: 034)
217. Francesc ROBUSTÉ, CENIT - Center for Innovation in Transport - Spain (Paper: 050)
218. Massimilliano RONCONE, CVIT S.p.A. - BOSCH Italia - Italy (Paper: 044)
219. Andrea ROSA, Napier University - United Kingdom (Paper: 035)
220. Guido ROSSI, DISTART - University of Bologna - Italy (Paper: 046)
221. Riccardo ROSSI, University of Padova - Italy (Paper: 070)
222. Damiano ROSSI, Poliedra - Politecnico di Milano - Italy (Paper: 133)
223. Federico RUPI, DISTART - University of Bologna - Italy (Paper: 046)
224. Francesco RUSSO, Mediterranea University of Reggio Calabria - Italy (Paper: 029)
225. Michele RUTA, DEE - Technical University of Bari - Italy (Paper: 143)
226. Marco SALATINO, University of Calabria - Italy (Paper: 083)
227. Giuseppe SALVO, University of Palermo - Italy (Paper: 037)
228. **Domenico SASSANELLI**, DVT - Technical University of Bari - Italy  (Papers: 077, 123)

229. **Hanna SAWICKA**, Poznan University of Technology - Poland (Paper: 090)

230. **Cem SAYDAM**, University of North Carolina at Charlotte - USA  (Paper: 112)

231. **Jan-Dirk SCHMÖCKER**, Imperial College London - United Kingdom  (Paper: 073)

232. **Anna SCIOMACHEN**, Università di Genova - Italy  (Papers: 130, 132)

233. **Giuseppe SCIUTTO**, Consorzio Nazionale Interuniversitario per i Trasporti e la Logistica - Italy  (Papers: 021, 091)

234. **Burak ŞENGÖZ**, Dokuz Eylûl University - Turkey (Paper: 069)

235. **Antonio SFORZA**, Università degli Studi di Napoli “Federico II” - Italy  (Papers: 131, 132)

236. **Hiroshi SHIMAMOTO**, Kyoto University - Japan (Paper: 073)

237. **David SIMMONDS**, David Simmonds Consultancy Ltd - United Kingdom  (Paper: 080)

238. **Carlos SOARES**, University of Porto - Portugal (Paper: 052)

239. **Francesc SORIGUERA**, CENIT - Center for Innovation in Transport - Spain  (Paper: 034)

240. **Ingmar STEINZEN**, University of Paderborn - Germany  (Paper: 062)

241. **Constanze STOCKHAMMER**, HiTec Marketing - Austria (Paper: 036)

242. **Leena SUHL**, University of Paderborn - Germany  (Papers: 056, 062)

243. **Zhongwei SUN**, Eindhoven University of Technology - Netherlands  (Paper: 007)

244. **Yos SUNITIYOSO**, University of the West of England - United Kingdom  (Paper: 063)

245. **Andrzej SZARATA**, Cracow University of Technology - Poland (Paper: 141)

246. **Wai Yuen SZETO**, University of Dublin - Republic of Ireland (Paper: 079)

247. **Shinji TANAKA**, University of Tokyo - Japan  (Paper: 043)

248. **Serhan TANYEL**, Dokuz Eylûl University - Turkey  (Paper: 067, 069)

249. **Dušan TEODOROVIĆ**, University of Belgrade - Serbia  (Paper: 109)

250. **Leif THORSON**, CENIT - Center for Innovation in Transport - Spain  (Paper: 034)

251. **Harry TIMMERMANS**, Eindhoven University of Technology - Netherlands  (Paper: 007)

252. **Konstantinos TRIANTIS**, Virginia Polytechnic Institute and State University - USA  (Paper: 109)

253. **Walter UKOVICH**, DEEI - University of Trieste - Italy  (Paper: 113)
254. Nobuhiro UNO, Kyoto University - Japan  (Papers: 047, 75)
255. Giulio VALFRE, D’Appolonia S.p.a - Italy  (Paper: 051)
256. Eric VAN BERKUM, University of Twente - Netherlands  (Paper: 040)
257. Jan Willem G. M. VAN DER PAS, Delft University of Technology - Netherlands  (Paper: 110)
258. Hans VAN LINT, Delft University of Technology - Netherlands  (Papers: 015, 022, 071)
259. Tom VAN VUREN, Mott MacDonald Ltd - United Kingdom  (Paper: 001)
260. Henk J. VAN ZUYLEN, Delft University of Technology - Netherlands  (Papers: 054, 087)
261. Javier VAZQUEZ, University of Sevilla - Spain  (Paper: 027)
262. Cristian VEGLIA, Consorzio Nazionale Interuniversitario per i Trasporti e la Logistica - Italy  (Paper: 021, 091)
263. Elisabetta VENEZIA, Universita degli Studi di Bari - Italy  (Paper: 108)
264. Romeo VESCOVI, University of Padova - Italy  (Paper: 070)
265. Milorad VIDOVIC, University of Belgrade - Serbia & Montenegro (Paper: 117)
266. José Manuel VIEGAS, Universidade Técnica de Lisboa - Portugal (Paper: 102)
267. Paula VIEIRA, Technique University of Madrid - Spain  (Paper: 064)
268. Antonio VITETTA, Mediterranea University of Reggio Calabria - Italy  (Paper: 029)
269. Francesco VITI, Delft University of Technology - Netherlands  (Paper: 054)
270. Katarina VUKADINIVIC, University of Belgrade - Serbia & Montenegro (Paper: 117)
271. Hiroshi WAKABAYASHI, Meijo University - Japan (Paper: 085)
272. W. WALKER, Delft University of Technology - Netherlands  (Paper: 110)
273. Deng WEI, Southeast University - People's Republic of China (Paper: 009)
274. Wang WEI, Southeast University - People's Republic of China (Paper: 009)
275. Li WEIGANG, University of Brasilia - Brazil (Paper: 008)
276. Wendy WEIJEMARS, University of Twente - Netherlands  (Paper: 040)
277. Roberto WOLFER CALVO, Universite de Technologie de Troyes - France (Papers: 132, 133)
278. Polat YALINIZ, Eskisehir Osmangazi University - Turkey  (Paper: 078)
279. Saini YANG, University of Maryland - USA  (Paper: 105)
280. Ji YANJIE, Southeast University - People's Republic of China (Paper: 009)
281. Mustafa Anil YAZICI, Rutgers University - USA  (Paper: 114)
282. Jacek ZAK, Poznan University of Technology - Poland (Paper: 090)
283. Yueqin ZHAO, Virginia Polytechnic Institute and State University - USA  (Paper: 109)
284. Pietro ZITO, University of Palermo - Italy  (Paper: 037, 039)
285. Frank Stendert ZUURBIER, Delft University of Technology - Netherlands  (Paper: 022)
MULTIPLE OBJECTIVE SIGNAL CONTROL FOR URBAN ENVIRONMENTS

Sonal AHUJA¹, MGH BELL¹, Tom van VUREN²

Abstract

This paper focuses on the development of a multi objective traffic signal controller using artificial intelligent optimisation techniques. Using fuzzy logic to define linguistically the various conflicting objectives of a traffic signal controller and a genetic algorithm to refine the definitions of the fuzzy membership functions, a hybrid signal control mechanism is proposed to improve traffic flow conditions at two case study sites in the UK. An adaptive fuzzy logic genetic algorithm (FLGA) signal control mechanism is proposed which can respond to fluctuating vehicular and pedestrian demands and priorities. The results of the simulation of the proposed signal control algorithm have been compared with fixed time, vehicle actuated and 'scramble' signal controllers.

Traffic signal control is increasingly seen as a way to implement social policy and not just a way to improve traffic efficiency. Local authorities expect it to deliver priority for public transport, easier and safer walking and cycling facilities, secure access for the mobility impaired, clean air and revitalised town centres. The objectives have therefore become multi-dimensional and the scope of control multi-modal.

The objectives of signal control may be conflicting and inconsistent. For example one may like to provide greater pedestrian priority at a signalised intersection and reduce bus delays while not increasing the delays to general traffic inordinately. The above can perhaps be a local objective at a particular intersection, and other intersections may have different local objectives all trying to aim for the network-wide objective of providing smooth flow and lowering delays.

¹ Centre for Transport Studies – Dept. of Civil and Environmental Engineering, Imperial College London, London, SW7 2AZ, United Kingdom. Email: s.ahuja@imperial.ac.uk
² Mott MacDonald Ltd, Birmingham, United Kingdom
To help fulfil the expectations of local authorities, a software platform for multi-objective traffic signal optimisation has been developed, using fuzzy logic for the control and a genetic algorithm for the optimisation. Innovative aspects of the work relate to the use of fuzzy logic for multi-modal control, multi-objective optimisation taking user reactions into account, the consideration of time dependency in modelling user reactions, and the inclusion of the mobility impaired.

This paper covers the first attempt at producing a Fuzzy Multiple Objective traffic Signal Control (FMOSC) algorithm for a real world application. The developed FLGA signal controller has been simulated in the VISSIM microsimulation package for two proposed traffic signals in the highly congested Stechford area in Birmingham UK and the traffic signals in West Croydon where there is conflicting demand between pedestrians, trams, buses and other vehicles.

For these study intersections the traffic engineer is posed with a notoriously difficult problem to satisfy several conflicting objectives that must be achieved by the signal control system. These objectives include:

- Satisfy traffic demand from all competing arms and reduce queues on approaches so that they do not block back into other intersections;
- Reduce pedestrian delays and conflicts with vehicles while not increasing vehicular delays;
- Reduce not only queues but delays for vehicular users and optimise the performance of the traffic signal controller as a whole;
- Address the fluctuating and changing nature of vehicular and pedestrian demand between various arms of the intersection.

Compared to a conventional complex optimisation situation which may be represented through complex sets of equations and the manipulation of variables, fuzzy logic uses linguistic variables combined with human reasoning. The use of linguistic variables (e.g. “long delay”) characterises the properties of the approximate in a semantic and computationally useful way. At the core of the fuzzy logic system is a set of rules which describe the relationship between the inputs and the outputs in a qualitative linguistic way.

In a real life situation the traffic engineer may respond to logical conditions such as – “IF LONG queue on a link, THEN give HIGH green time”. The conditions that define LONG and HIGH are subjective conditions based on an expert’s experience. In the proposed signal controller, fuzzy logic works in a similar way, attempting to translate the crisp knowledge in a process, such as green time allocation, to linguistic or fuzzy knowledge.

As in a knowledge-based expert system, these rules provide an easily understood scheme for explaining the input/output mapping. In contrast to expert systems, however, a fuzzy logic rule base can be quite simple and concise, due to the mapping of the individual discrete input and output values onto user-defined fuzzy set.
Crisp inputs in the proposed fuzzy controller include, queue lengths, delays, demand and flow for different arms and different users (pedestrians and vehicles) and the crisp output is allocation of green time for the demanded stages.

The key to the success of the fuzzy logic system is the definition of the membership functions which define the fuzzy sets. These definitions often come from ‘expert knowledge’ or engineers who are well conversant with the operation of the system. In this paper we propose an adaptive genetic algorithm that mimics the role of the ‘expert’ in the system.

The average network delay per vehicle and average travel time during the peak hour are used to determine the performance of the fuzzy logic controller. The crisp output from the simulation runs is formulated as a fitness function in a genetic optimiser. The Genetic Algorithm (GA) is used to evaluate the performance of the network and to suggest other definitions (fine-tuning) of membership functions (input and output functions) of the fuzzy controller with the aim of minimising network delays and travel time.

The results of the study (Birmingham site) indicate that compared to a vehicle actuated controller, the manually tuned fuzzy logic controller can give a 4% reduction in network travel time and a 10% reduction in delay.

The developed platform aims to help practitioners and decision-makers achieve a better integration of modes (walking, cycling, bus and car) with each other and with the urban environment. Through the explicit consideration of pedestrians, cyclists and those with mobility difficulties, the platform can be used to promote social inclusion.
DO PEDESTRIANS EFFICIENTLY USE OVERPASS-PEDESTRIAN-CROSSING-FACILITIES IN CROSSING URBAN STREETS? USAGE ASSESSMENT AND PLANNING PARAMETERS

Dr. Hassan Abdelzaher Hassan Mahdy¹

Abstract: In urban areas, where intensive mixed land-uses "traffic generators" exist, overpass-pedestrian-crossing-facilities OPCF are normally used to safeguard pedestrians while crossing expressways, primary streets, railway, and metro lines. Also, they contribute in reducing casualties and preventing traffic interruption. Unfortunately, pedestrians don't use OPCF efficiently; where portion of pedestrians crossing urban roads without using them. This paper focuses on assessing the OPCF usage efficiency based on questionnaires performed on a sample of pedestrians crossing urban roads and streets without using OPCF. Also, planning parameters for OPCF will be suggested based on the questionnaires analysis.

1. Methodology

The used research includes the following activities:

- Selecting survey location;
- Identifying of the physical characteristics of survey location,
- Surveying of road furniture and traffic control devices TCD of survey location
- Collecting data; the following data are simultaneously collected,
- Counting of road-vehicular-traffic,
- Counting of pedestrians (moving on sidewalk and crossing road or street),
- Interviewing sample of crossing pedestrians using OPCF
- Analyzing collected data;
- Assessing OPCF usage
- Recommending OPCF parameters

¹ Lecturer of Highway and Airport Engineering, Faculty of Engineering, Ain Shams University, Cairo, Egypt
Email: h9840417@hotmail.com   Cellular: +2 0127483563
2. Study location and its physical characteristics

To achieve research objectives, OPCF located on Autostrad road (at Mansha'a Naser, Cairo, Egypt) is selected. This site has the following characteristics: a) Heavy traffic volumes and high speed road, b) Wide road cross section with a narrow central island, c) Relatively heavy pedestrian crossings, and d) No barriers or fences on sidewalks or central island.

3. Analysis

3.1. Analysis of traffic and pedestrian data

This section presents the analysis of collected data to reflect the current performance indicators of pedestrians and the usage efficiency of OPCF. Figure (1) shows the fluctuations in road traffic volumes and crossing pedestrians over the survey time which extended from 7:00 to 19:00 on a normal working day. The left axe represents the measured directional traffic volumes of the road in vehicles per hour, while the right axe represents the volumes of crossing pedestrians in pedestrians per hour. The observed directional traffic volumes vary from 3948 to 6030 vph.

There are two approximately equal and extended traffic peaks; morning peak from 8:00 to 11:00 and afternoon peak from 14:00 to 16:00 with about 6000 vph per direction. Also, two pedestrian-crossing flow peaks occur; the first is about 670 ped/h, from 8:00 to 10:00 and the second is about 790 ped/h from 14:00 to 15:00.

The peak periods of crossing pedestrians who don't use OPCF are similar to traffic flow peak periods (morning peak from 8:00 to 11:00 and afternoon peak from 14:00 to 16:00) because during these periods the road is approaching congestion situation (low vehicles' speed approaching stop and go condition) which encourage pedestrian to cross the road without using OPCF. But during the off-peak traffic flow (from 11:00 to 14:00) vehicles' speed are much higher. That explains why the peak of crossing-pedestrian using OPCF occurred in this time period (see the gray part of Figure 1).

In general and over the total survey time, crossing pedestrians using OPCF are lower than those who don't. The overall percentages are 28 and 72% respectively e.g. the usage efficiency of OPCF is about 28% (see Figure 2).

![Figure (1) Fluctuations in road traffic and crossing pedestrian volumes over the survey time](image1)

![Figure (2) Fluctuations in volumes & percentages of crossing pedestrians over survey time](image2)
The gray part in Figures 1 & 2 shows the peak of crossing-pedestrian using OPCF, at the same time the off-peak of traffic volume and crossing pedestrian don't use OPCF. Figure (3) shows the fluctuations in the total crossing-pedestrians (using or not OPCF) classified according to pedestrian gender. The overall average percentages of females crossing-pedestrian using or not OPCF are 17 and 18% respectively, compared to 11 and 54% respectively, for males.

The ratio between percentages of males and females crossing roads without using OPCS is 3.18 which means that males don't seriously feel the risk of crossing roads without using OPCF as females do, and they expose themselves to accident risks 3 times the females. (Figure 3).

3.2. Questionnaire results

Figure (4) summarizes the results of question number 2 in the questionnaire form. The horizontal axe displays item numbers, while vertical axe displays the weight of each item. These weights are calculated based on the questionnaire data.

The most influenced items, ranked according to their weights, are: 2, 5, 1, 7, and 6 respectively, as their cumulative weight exceeds 85%. Items 3, 4, 8, and 9 (the gray parts in Figure 4) are considered ineffective as their total weight is less than 15%.

The most influence items of why pedestrians don't use OPCF, as extracted from the interview results, are:
- OPCF is high and stairs are not comfortable (item # 2, weighting 46.1%).
- Always occupied by vendors (item # 5, weighting 13.2%)
- Not in the proper locations (item # 1, weighting 13.2%)
- No lights (item # 7, weighting 6.6%)
- Slippery surface (item No. 6, weighting 6.6%)

4. Suggesting planning criteria of OPCF

Figure (5) summarizes the results of the questionnaire form. The horizontal axe displays item numbers, while the vertical axe displays the weight of each item. These weights are calculated based on the questionnaire data.

The most influenced items, ranked according to their weights, are: 3, 4, 2, 5, and 1 respectively, as their cumulative weight exceeds 85%. Items 6, 7, and 8 (the gray parts in Figure 5) are considered ineffective as their total weight is less than 15%.
5. The most influence parameters in OPCF planning

The most influence items that encourage pedestrian to use OPCF, as extracted from the interview results, are:

- Ramps should be used for elders and handicaps (item # 3, weighting 24.4%).
- Elevators should be used (item # 4, weighting 19.5%)
- The reasonable heights of OPCF (item # 2, weighting 19.5%)
- Providing transit nodes at or nearby OPCF locations (item # 5, weighting 17.1%)
- The proper location of the OPCF (item No. 1, weighting 6.6%)

Therefore, in planning OPCF, the above planning parameters should be carefully addressed to encourage pedestrians using OPCF in crossing urban roads and street and consequently reducing potential for vehicle-pedestrian accidents and enhance safety.

6. Conclusions and recommendations

The following conclusions and recommendations can be drawn from this research:

a) A questionnaire form is designed and used in pedestrian interview, b) Pedestrians and vehicles survey forms are designed and used, c) OPCF usage efficiency is assessed, d) Males don't seriously feel the risk of crossing roads without using OPCF, e) as they expose themselves to 3 times the risk that females do, f) The most influence factors that make pedestrians don't use OPCF in crossing roads are determined, g) The most influence OPCF planning parameters are suggested. h) Traffic education, particularly traffic safety, should be a subject of primary and high schools courses.

7. Bibliography


INTELLIGENT TRANSPORT SYSTEMS: ROLE OF SENSORS IN TRAFFIC MANAGEMENT

Aman MOUDGIL¹ Gaurav MOUDGIL² Munish BANSAL³

Abstract. Transportation Infrastructure is the backbone of every country. This is the gateway to economic development and subsequently, a healthy economy. Urbanization process in most of cities has reached an alarming stage, flooding the urban regions. Various attempts have been made to provide the necessary transport services and related infrastructure to meet the requirement. The rapid increase in transportation facilities has led to large amount of traffic problems including violating traffic rules leading to accidents (sometimes mat be fatal). This paper outlines the application of Information Technology in controlling transportation systems to solve many of the problems.

1. Introduction

With the rising automobile industry, and the banks helping mankind in fulfilling their dreams, the number of automobiles has grown exponentially in the past few years. This has led to the fact that once considered a luxury, vehicles have now turned out to be a necessity and almost each and every person owns some vehicle or the other (be it a car, scooter, motorcycle, etc). This sharp increase in vehicles has led to large amount of traffic problems and the ones in a hurry prefer to skip traffic lights even when signaling Red.

Now in order to help out the unlucky ones or rather to help the traffic authorities in managing this tumultuous traffic, the use of sensors at traffic junctions are proposed that should ease the working of the traffic authorities. The use of this technology should help the traffic management in the way that they will not have to strain their lives (as they have to nowadays) in keeping a continuous check on the disobedient freaks.

¹ Department of Computer Science and Engg, Sector-26 Chandigarh-Pin:160019, aman_moudgil3000@yahoo.com
² MS-SEMM University of California, Berkeley, gaurav_moudgil2000@yahoo.com
³ B.E-Civil Engg Punjab Engg. College, Sector-12 Chandigarh-Pin:160012, munishbansal29@yahoo.com
2. Sensors and Transducers

Sensors and Transducers are devices used to measure mechanical, thermal, electrical and chemical changes occurring around. The sensor is responsive to changes in the quantity to be measured, for example, temperature, position, or chemical concentration. The transducer converts such measurements into electrical signals, which can be fed to instruments for the readout, recording, or control of the measured quantities. Sensors and transducers can operate at locations remote from the observer and in environments unsuitable or impractical for human beings.

Some devices act as both sensors[1] and transducer[1]. A thermocouple has two junctions of wires of different metals; these generate a small electric voltage that depends on the temperature difference between the two junctions. A Thermistor is a special resistor, the resistance of which varies with temperature. A variable resistor can convert mechanical movement into an electrical signal. Specially designed capacitors are used to measure distance, and photocells are used to detect light. Other devices are used to measure velocity, acceleration, or fluid flow. In most instances, the electric signal is weak and must be amplified by an electronic circuit.

For such a situation, use of SPZ & SPT Fiber optic traffic sensors[5] interfaced via MD-200[5] can be made. These sensors are sensitive to vertical pressure only but are very flexible and elastic. These sensors do not contain any metallic parts. Traffic sensors are not sensitive to environmental conditions and also can function smoothly under a wide range of temperature. Moreover unlike other types of sensors, these are quite simple to install.

These sensors and many other Traffic control equipments are being manufactured by SIEMENS, which are the global leaders in Traffic Control and Regulation systems. SIEMENS have also developed Traffic control computers for controlling traffic signals under the names SITRAFFIC CENTRAL, ACTRA, and ICONS. These computer systems were developed in Germany, England and the U.S and provide specific solutions for all the requirements encountered in international transportation business. In fact SIEMENS manufactures total traffic control equipments right from Traffic Sensors to Traffic control computers to the maintenance and regulation of these equipments.

Figure 1. Traffic Sensor
3. Proposal

3.1. Installing a Transmitter in the automobiles while issuing licenses for the vehicles

It is proposed that while issuing licenses, the licensing authorities should direct the car dealers to install a Transmitter in the vehicles and once done, they (authorities) should key-in the owner’s name, car number, model number, owner’s address, etc in the chip. The Transmitter should be the first thing to start in the car and it should be connected to the engine of the car in such a way that the engine of the car cannot be switched on if it (The Transmitter) is missing.

3.2. Laying down Sensors at traffic junctions[2]

It is proposed that if we lay down sensors at the traffic junctions in the manner as shown in the figure, then it would help to punish the defaulters and make respective authorities more efficient. Now referring from fig.1 let us suppose that as TRAFFIC LIGHT 1 switches to green then the sensor 1 should switch off (indicated by blue) and with TRAFFIC LIGHT 2 switching to red the sensor 2 should switch ON automatically. If VEHICLE 2 tries to skip TRAFFIC LIGHT 2 then the sensor should direct the receiver to collect the information from the Transmitter of the VEHICLE and redirect this information to the server (maintained by the TRAFFIC REGULATORY AUTHORITIES).
Figure 3. Depicting the concept of Traffic Management through the use of Sensors

Figure 4. Bird’s Eye-view of a modernized Traffic Junction

This Traffic Police Headquarters will maintain a database of this information. The respective authorities can now collect this information. For converting these signals into some other readable form of energy, Transducers may be employed.

3.3. Installing Traffic control computers for controlling Traffic

Responsible for controlling traffic signals, they hold the reins on the entire flow of traffic in the city. Intelligent software combined with powerful hardware enables the traffic
data to be visualized in a clear and understandable manner. Siemens has developed such efficient and user-friendly systems under the names SITRAFFIC, ACTRA, and ICONS. Such traffic control systems should be installed with utmost care and expertise at the traffic junctions.

3.4. Installing the new generation of pay machines

The Database maintained by the Traffic authorities would list the defaulters on their website along with the due date for the payment of the challan. The defaulters could pay their challans either online with a Credit card/cheque or offline by the current procedures of payment. This would benefit the people, as it would become easier for them to repay the challans.

In case the person does not pay the due amount, his/her license shall be cancelled and the person will be subjected to the necessary legalities that are entitled when a person does not pay in for the challan.

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>ID</td>
<td>Name of Person</td>
<td>Address</td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Aman Moudgil</td>
<td>232, akk road, chandigarh</td>
</tr>
<tr>
<td>3</td>
<td>Gaurav Moudgil</td>
<td>232, sdfdsfsd, chandigarh</td>
</tr>
<tr>
<td>4</td>
<td>Henry gourd</td>
<td>345, edgedgdg, chandigarh</td>
</tr>
<tr>
<td>5</td>
<td>amarbir singh</td>
<td>234, sdfdsdfdf, chandigarh</td>
</tr>
</tbody>
</table>

Figure 5. Database maintaining the particulars of the defaulters

4. A More Detailed Description

This process can broadly be subdivided into four parts. These are explained as below:

Figure 6. Representing the grid of sensors laid down at Traffic Light Junction
When the SENSORS are not ON, i.e. when the traffic light is signaling GO, then the network will remain switched OFF and the circuit will not respond to the Transmitter (which will be transmitting data forever till the engine of the car is switched OFF). But as soon as the Traffic light signals RED, the SENSOR will automatically be activated. Now if a car tries to cross the Traffic light then the following sequence of steps will take place:

1. When activated, the sensor sends a signal to the Receiver.
2. After receiving the signal, the Receiver switches itself on and searches for the possible signal.
3. When the receiver establishes connection with the transmitter, it receives all the information transmitted by the transmitter and then passes it on to the server, which maintains a database of this information.

5. Benefits of Installing such a Network

1. This would develop the technical skills of the present working staff.
2. Would lessen the strain on the Traffic Police.
3. Would prevent the defaulters from getting rid of fines.
5. Should lead to a decrease in the number of road accidents occurring at Traffic junctions.
6. Would create more decent drivers.
7. Will make the system more transparent.
8. Could also be used for Flow monitoring and Traffic data acquisition like Volume, speed, gap time, headway, occupancy, concentration, etc.
10. Could also be used for controlling the Urban and Interurban Traffic.
11. Besides at Traffic junctions, these can also be utilized as parking guidance systems.
12. Could also be utilized for vehicle classification

6. **Method of Laying Down the Sensors[5]**

The method of installing such a grid of Sensors can be categorized as:

1. Initially cut a slot of suitable dimension into the road.
2. Clean and dry the road.

![Cross-Sectional view of fiber optic traffic sensors](image)

3. Pour small amount of an Embedding material in the bottom of the slot.
4. Place the sensor into the Embedding Material.
5. Now fill the cavity with any Embedding material like SL Cast.
6. When dried, the road is ready for use.

![Final view of an installed Grid](image)
7. Conclusion

With the rising Traffic congestion, modernization of current traffic-control system will have to be necessary at some time or the other. Its e-modernization is essential as it will infer better output in terms of traffic regulation and control and will enable efficiency and transparency in the current system. This will also lessen the rising burden on the current transport regulatory bodies. This article basically directs the use of a technology, hoping that it will be implemented. This technology can be fruitful if this methodology is implemented for an entire city. Currently, similar Traffic systems are be developed by SIEMENS and such systems are being used by countries like Germany, England and the United States of America for managing their ever increasing traffic problems.

References


EFFECT OF EXTERNAL COST INTERNALISATION ON TRANSIT DESIGN PROBLEMS

Mariano GALLO†, Luca D’ACIERNO*, Bruno MONTELLA*

Abstract. This paper proposes a multimodal transit supply design model with elastic demand, which takes external costs into account. The proposed model is applied in the case of transit fare optimisation in the region of Campania (southern Italy). Initial results show that internalisation of external costs yields a decrease in system costs and an increase in transit system use. Finally, the proposed models are also applied in the case of reciprocal influence between road and transit systems, yielding rather similar solutions.

1. Introduction

In the literature, much attention has been focused on the transit supply design problem, analysing different aspects and proposing numerous solution approaches. However, it is important to note that most models are based on the following simplifying assumptions:

1. analysis only of the transit system;
2. hypothesis of rigid travel demand;
3. lack of reciprocal influence among transportation systems;
4. lack of external cost terms in the objective function.

The first two hypotheses considerably simplify the problem even if they are hardly ever acceptable in real cases. Indeed, supply changes in the transit system can yield non-negligible impacts on user travel choices and therefore on Origin-Destination matrices as well as road system performances. In fact, a transit design model should normally have elastic demand and be able to evaluate proposed solutions simulating effects both on transit and road systems. In the literature there are some examples of design with elastic demand, whether in the case of problem general formulation ([5], [19]) or transit fare design ([17], [20]).

† Department of Engineering, University of Sannio, piazza Roma 21, 82100 Benevento, Italy, gallo@unisannio.it
* Department of Transportation Engineering, University of Naples “Federico II”, via Claudio 21, 80125 Naples, Italy, {dacierno|montella}@unina.it
Several design models with elastic demand neglect the influence that variations in level-of-service in the road transportation system could have on transit system performances (assumption 3). This side of the problem is related to choices of the simulation model, which can be classified as “with mutual congestion” if it does not neglect influences between the two transportation systems. Some assignment models with mutual congestion are proposed by [2] as well as [6], and utilised to formulate design models ([3], [7], [16]).

An often-neglected aspect in the formulation of transit design models is the evaluation of external costs (assumption 4). An effective transit system, reducing the use of private cars, yields a decrease in system external costs (air and noise pollution, accidents, etc); the evaluation of these costs could be fundamental for an optimal choice of planning configuration. The importance of external cost evaluation is highlighted by extensive research developed in international contexts ([12], [18], [24]). In the literature external costs have been considered chiefly in road system design models ([1], [3], [7], [8], [9], [13], [14], [16], [23], [25]).

2. Our proposal

A general multimodal transit design problem with elastic demand, which overruns assumption 1 and 2, can be formulated as a constrained multidimensional optimisation model, that is:

$$\hat{y} = \arg\min_{y \in S_y} \left( w(y, f_{\text{transit}}, f_{\text{road}}) \right)$$

subject to:

$$\left[ f_{\text{transit}}, f_{\text{road}} \right] = A(y, f_{\text{transit}}, f_{\text{road}})$$

$$\text{NFC}(y, f_{\text{transit}}) \leq \text{Budget}$$

$$f_{\text{road}}(y, f_{\text{transit}}) \leq \phi_l(y, \text{Cap}_l) \quad \forall l$$

where $y$ is the design variable vector, $S_y$ is the feasibility set of design variable vectors; $w(\cdot)$ is the objective function to be minimised, $f_{\text{transit}}$ ($f_{\text{road}}$) is the equilibrium flow vector of transit [road] system, $A(\cdot)$ is the assignment function, $\text{NFC}(\cdot)$ is the net operational cost of the transit firm, $f_{\text{road}}(\cdot)$ is the maximum number of on-board users of line $l$, $\phi_l$ is the service frequency of line $l$, $\text{Cap}_l$ is the vehicle capacity of line $l$.

Equation (2) expresses the assignment constraint, equation (3) indicates the budget constraint and equations described by relation (4) are the constraints of satisfaction of travel demand (as shown by [4]).

In general, objective function $w$ is equal to the sum of transit user costs ($\text{UC}_{\text{transit}}$), road user costs ($\text{UC}_{\text{road}}$) and net cost of transit firm ($\text{NFC}$), that is:

$$w(y, f_{\text{transit}}, f_{\text{road}}) = \text{UC}_{\text{transit}}(y, f_{\text{transit}}) + \text{UC}_{\text{road}}(y, f_{\text{road}}) + \text{NFC}(y, f_{\text{transit}})$$

with:

$$\text{UC}_{\text{transit}}(y, f_{\text{transit}}) = \sum_i \sum_a f_{\text{transit}}^{a,i}(f_{\text{transit}}) \left( \beta_{\text{transit}}^{a,i} \cdot \text{UT}_{\text{transit}}^{a,i}(y, f_{\text{transit}}) + \text{UC}_{\text{transit}}^{a,i}(y) \right)$$

$$\text{UC}_{\text{road}}(y, f_{\text{road}}) = \sum_i \sum_a \delta^i \cdot f_{\text{road}}^{a,i}(f_{\text{road}}) \left( \beta_{\text{road}}^{a,i} \cdot \text{UT}_{\text{road}}^{a,i}(y, f_{\text{road}}) + \text{UC}_{\text{road}}^{a,i}(y) \right)$$
where \( f_{a,i}^{\text{transit}} \) is the transit user flow on link \( a \) for \( i \)-th user category, \( \beta_{i}^{\text{transit}} \) is the value-of-time on link \( a \) for \( i \)-th user category, \( UT_{a,i}^{\text{transit}} \) \( [UT_{a,i}^{\text{road}}] \) is the transit [road] travel time on link \( a \) for \( i \)-th user category, \( UC_{a,i}^{\text{transit}} \) \( [UC_{a,i}^{\text{road}}] \) is the transit [road] monetary cost on link \( a \) for \( i \)-th user category, \( \delta^{i} \) is the occupancy index (the average number of road users for each vehicle) for \( i \)-th user category, \( f_{a,i}^{\text{road}} \) is the road vehicle flows on link \( a \) for \( i \)-th user category.

Moreover, the transit firm cost is equal to the difference between operational costs of the transit firm (OFC) and ticket revenues (TR), that is:
\[
NFC(y, f_{a,i}^{\text{road}}) = OFC(y) - TR(y, f_{a,i}^{\text{road}})
\]
with:
\[
OFC(y) = \sum_{l} c_{i}^{l} \cdot L_{l}(y) \cdot \varphi_{l}(y)
\]
\[
TR(y, f_{a,i}^{\text{road}}) = \sum_{i} \sum_{a} f_{a,i}^{\text{road}} \cdot (f_{a,i}^{\text{road}}) \cdot UC_{a,i}^{\text{road}}(y)
\]
where \( c_{i}^{l} \) is the standard operational cost for each \( km \) of line \( l \), \( L_{l} \) is the length of line \( l \) and \( \varphi_{l} \) is the service frequency of line \( l \).

In particular, we show that optimisations without external costs yield solutions that are mathematically correct but fail to comprehend the analysed problem in real terms. Indeed, neglecting external costs not only underestimates real costs but also yields different solutions. Therefore we propose an objective function that also includes an external cost term \( (EC) \) that can be expressed as \((15)\):
\[
EC(f_{a,i}^{\text{road}}) = ec_{\text{road}} \cdot \sum_{a} \sum_{i} f_{a,i}^{\text{road}} \cdot L_{a}
\]
where \( ec_{\text{road}} \) is the external cost for each \( \text{vehicle-km} \) produced by the road system, \( L_{a} \) is the length, expressed in \( \text{km} \), of link \( a \). Moreover, \([15]\) estimated that \( ec_{\text{road}} \) value is equal to \(0.1460 \text{ €/(vehicle-km)}\) in the region of Campania (southern Italy).

In this paper we apply the proposed multimodal design problem, which takes external costs into account (overrunning of assumption 4), in the case of transit fare optimisation in the region of Campania, comparing results among present fares, optimised fares without external costs and optimised fares with external costs. This approach represents the case of a decision-maker that optimises the transit system by allowing for the presence of external costs.

Instead, the internalisation of external costs entails affecting road user choices as a function of external costs by means of suitable road pricing (which obviously has to be determined). Therefore, equation \((7)\) can be replaced by the following relation:
\[
UC_{a,i}^{\text{road}}(y, f_{a,i}^{\text{road}}) = \sum_{a} \sum_{i} \delta^{i} \cdot f_{a,i}^{\text{road}}((f_{a,i}^{\text{road}}) \cdot \beta_{i}^{\text{transit}} \cdot UT_{a,i}^{\text{road}}(y, f_{a,i}^{\text{road}}) + UC_{a,i}^{\text{road}}(y) + \Delta p_{\text{road}}^{i})
\]
where \( \Delta p_{\text{road}}^{i} \) is the suitable road pricing on link \( a \) for \( i \)-th user category for internalising external costs.

In particular, in this paper we show that with the assumption that revenues from suitable road pricing have to be equal to external costs (calculated by means of equation \(7)\), the increase in pricing for road users has to be equal to:
\[
\Delta p_{\text{road}}^{i} = ec_{\text{road}} \cdot L_{a}
\]
Hence, in this paper we propose another optimisation model that allows external costs to be internalised, and we apply it in the same context of Campania. Our results show that internalisation of external costs yields a decrease in system costs and an increase in transit system use.

Finally, the proposed models are also applied in the case of reciprocal influence between road and transit systems (overrunning of assumption 3), yielding rather similar solutions.

3. Conclusions and research prospects

In this paper we show that overrunning of assumptions in general utilised in transit supply design models (such as rigid demand, lack of external costs and lack of any internalisation) could yield results that are substantially different and model reality better. It is considered necessary to internalise external costs from a system point of view, since society will allocate resources anyway to reduce impacts produced by the road system by using, for example, transit firm subsidies.

In terms of future research, we propose model development as well as further real network applications. Finally, from a management point of view we suggest analysis of the effects produced by the proposed approach in the case of “first best” versus “second best” approaches (as proposed by [10], [22]), “implementation paths” (as proposed by [10], [11], [22]) and taking into account effects of different government levels (as proposed by [21]).

References


EXPERIMENTAL STUDY ON DATA DRIVEN PUBLIC TRANSPORT INFORMATION SYSTEM IN DHAKA CITY ROAD NETWORK

Moshiuzzaman MAHMUD¹, Yeasir RAHUL²

Abstract. Intelligent transportation system is providing a strong base for better and efficient transport facilities in the developed countries. Economy, socio-political instability and lack of resources hinder this sort of system to work properly in the less developed countries. However Transport system of such area can be benefited from that service if infrastructure and technology can be catered intelligently with the available resources. Among many other key issues, proper bus management is vital for the expanding city like Dhaka, Capital of Bangladesh. The availability and the dissemination of bus arrival and departure information are still very poor, although the need of it is extremely important. This experimental study was undertaken with a view to blend the latest smart technology with local resources that can meet the economic constraints and at the same time can give a strong base to future transport system of Dhaka. Today’s state of the art technology can be used to provide the commuters with the information of their next available bus, time required. This paper will give a feasible solution to such issue with respect to a developing country. The design and application of this system have been done considering the transport budget for Dhaka city. This bus information system requires web technology and that need to have a strong communication backbone between the Bus, Bus station and the central data storage system. The facilities such as World Wide Web (WWW), Wireless Application Protocol (WAP) or small scale Variable Message Sign (VMS) act as the information gateway for this purpose. This paper mainly focuses on the system architecture of this data driven Bus information system and take a brief look at its feasibility. This paper also addresses the implementation issues of the system.

¹ M.Sc Transportation Planning and Engineering, University of Southampton, UK
² American International BD. LTD
1. Introduction

Public transport system is broadly recognized as the solution to the Urban Congestion of today’s developing world. All the public transports are needed to be made more attractive to the commuters to get a congestion free road network. This is the reason why this sector of the transportation requires proper management. Though the rail transit systems provide high capacity, but are well sited for high-density corridors. They are expensive both in initial investment and maintenance, and take a long time to build. Hence, many countries opt for buses as low cost and quicker alternative to increase the capacity of public transport system. In most of the south Asian region, public are considered as the main means of transport for the city dwellers. Therefore it demands an efficient management to yield the maximum benefit out of the present infrastructure.

It is perceived that highest level of congestion, created in the urban centers, is caused by those commuting to work. The level of congestion depends upon the type of vehicle they are using. More these commuters are encouraged to use the public transport less the amount of congestion will occur. A proper information system of the availability of public transport can reduce the unnecessary congestion and which in turn reduce the loss of time to the commuters too. Providing the bus user with lucid and easy to follow information such as bus time tables, time required for next bus to come etc, are essential if the public transport is to become more useful and attractive. Today’s state of the art technology is helping this process at large in most of the developed country. Problem lays with the developing countries where economy and the resources are vital the issue. This case study mainly investigates such intelligent system with respect to a developing country. Here the capital city of Bangladesh, Dhaka, has been chosen as the site of the study. Hence an intelligent approach of Bus information system has been described which can be implemented keeping the economy and the resource constraints upfront.

2. Site Description

Dhaka, Capital city of Bangladesh, is one of the most densely populated cities in Asia. Being the capital city of the country, it attracts a huge amount of people from the rural area. The population of the city is continuously growing due to the natural growth and the rural-urban migration. Currently it holds a population of 12.62 million, 15.5% peripheral and 60.5% established urban area (DTCB 2005). It is causing enormous pressure on the existing transportation system and the infrastructure. With the limited resources Dhaka has a complex road network with three main arterials with several link roads (DTCB 2005).

Due to the presence of low income group of people Public transport is fairly a popular mode among the city dwellers. One of the major reason for the less person trips in pubic transport compared to non-motorized three wheelers, commonly known as rickshaw, is it’s journey time unreliability and its accessibility. Due to high level of congestion to the arterial and the link road, it is a common phenomenon that public transport fails to maintain its schedule. As a result people have to stand in the queue for long time in the bus stoppage.
Due to this uncertainty of bus arrivals, most often people prefer to use different other alternatives.

Understanding the need of proper transport system, Government of Bangladesh sought assistance of UNDP. Based on UNDP report a Five year plan had been taken under implementation from 1997-2002 (DTCB 2005). This is a part of the transport policy that Dhaka undertook for its development. The Land Transport Policy of Dhaka city has been formulated in the light of the Government’s pledge to establish a transport system which is a safe, cheap, modern, technologically dependable, environment friendly and acceptable in the light of globalization. The study, that has been elaborated here, can prove worthy in the context of Dhaka city.

3. Public Transport and the Passengers

A clear conception of the requirements road users should be present in the mind of the system operators, since this is directly related to the popularity of the public transport system. The identified key requirements are Proximity and Frequency of Service (Vijaybalaji 2000). A strong information system can reduce the gap between the passenger’s expectations and the existing system. This system will also increase the level of confidence of the passenger on the public transport. Most developed countries, is stressing hard on the reliability and the punctuality issue of the public transport operation these developed countries have already achieved success in that regard. However the issue is how conveniently these technologies can be implemented in the developing country?

4. System Framework

The whole proposed system comprises of three different units. These units are named as,

1. DAT- Data Acquisition and Transfer
2. DSP- Data Storage and Processing
3. DAD- Data Access and Display

All the units are connected one with another by a strong communication network (Figure.1). These networks are established by different means of communication such as telephone, radio link etc.

4.1. Data Acquisition and Transfer

In this segment of the system, the bus driver is required to provide their information at every bus stoppage. Data is collected by the Point of service (POS). Data such as the bus number, driver’s specifications, bus stoppage no, number of seats available in the bus etc can be entered through this process. It does not take not more than 10 seconds to enter these
data in the POS. Other information such as bus number, driver specification, rout card etc is preset information for the POS in a particular bus. After getting all the information in a stoppage, the POS will dial to the central data storage system in DSP unit and transfer the information. In common practice it is always desirable to have in-vehicle system, which involves less amount road side installations, at the same time it is more flexible in operation. In this In-Vehicle system, POS establishes communication with the DSP unit using Global System for Mobile (GSM) telecommunication system or radio-link.

4.2. Data Storage and Processing

This unit of the system receives data from the DAT unit through the communication channel established by the DAT. This server-end, the data bank for the whole system, stores all the data acquired from the previous unit and makes available for the next DAD unit. For the proper operation of the DAD unit, this data storage system should have a robust database, which is capable of high volume of data handling capability. Whenever a bus reaches a stoppage, the DAT unit will connect with the DSP unit and transfer data. In this circumstance, the DSP unit will have to handle a number of calls or data receiving request at a time. This request handling is maintained by a multi request handling hardware i.e. SmartNet. This technology is capable of shaking hand with a number of calls at the same time, and sends the actual data to the appropriate database using TCP/IP protocol.

As mentioned above, a continuous communication can be established with the GSM telecommunication system. With this established data bridge and the Global Positioning System (GPS), installed in the bus, can continuously update about the position of the bus. This position information is an important input to the bus arrival time prediction procedure in the digitized route maps with link-node representation (Cox and Love 1999). The digitized map can be used to determine the actual location of a bus with respect to its schedule. At the same time GPS data can be used for actual journey time prediction. Lin and Zeng (1999) have shown how the journey time prediction can be possible with the GPS data.

As mentioned earlier, all the data are stored in a central database. These data will be accessed by the users from the DAD unit via some web application, which will reside in the DSP unit. The front end of this web application is described in the next section of this paper. The backend should mainly consist of some PHP and WAP applications along with the main database. All the HTTP requests sent by the DAD unit will be handled here. For the monitoring purpose of the whole bus rout, the DSP can also have an administrative application which enables the administrator to keep an eye on the whole system. Any irregularity of data acquisition or collapsing of bus network can be identified immediately by that application.
4.3. Data Access and Display

This is the last unit of the system where the user will interact with the data by the mean of accessing data through the web services. The web interface will show the map of the bus rout along with the latest bus information in the stoppage locations. A Green Dot shall indicate the position of the stoppage. The corresponding information will be shown by text just beside the Green Dots (Figure 2). The system user will be able to browse through the whole road network in the web interface. This self refreshed interface will refresh itself in a certain interval so that the user can be benefited with the actual real time data, and can make decision about their travel schedule accordingly. The same information will also be shown in Variable Message Sign (VMS) in the stoppage.
5. Requirements of Integration of the System

This application in the developing countries can bring manifold benefits to the commuters, transport planners, system operators and the country as a whole. However implementation should follow some key issues. These key issues are identified as (Yokota,T 2004),

- Institutional Requirements
  - awareness of the scheme to be taken
  - Sufficient budget and capital investment
  - A good practice of traffic laws and regulations
  - Developed/development of Human Resources and

- Infrastructure Requirements
  - Communication standard for data exchange
  - A common data model for efficient information coordination and exchange
  - Wire or wireless communication backbone with appropriate bandwidth
5.1. Implementation Issue: Institutional

The main Institutional issue is the Budget. In the fiscal year 2004-05 the total allocated budget for Transport and Railway sector was 4480 crore taka which was equivalent to $746 million (SDNBD ORG 2004). A major portion of this budget is allocated for Dhaka city. A comparison (Table 1) with almost similar but conventional technology installed in Taipei can portray how conveniently this technology can be incorporated in Dhaka.

<table>
<thead>
<tr>
<th>No of AVL Equipped Vehicle</th>
<th>Type of AVL</th>
<th>Total Capital Cost of AVL (Reported)</th>
</tr>
</thead>
<tbody>
<tr>
<td>135</td>
<td>GPS</td>
<td>$270,000</td>
</tr>
</tbody>
</table>

**Table 1: Cost involved in the conventional system in Taipei** (Source: TCRP 2003)

5.2. Implementation Issue: Infrastructural

Infrastructural issue can be addressed according to the three stages of the system as described before,

a) At the current state of infrastructure, Dhaka does not have any roadside detectors or any GPS equipped vehicle. Volvo double-deckers are mainly used for the public transport in Dhaka city. GPS equipments can easily be installed in these buses or direct GSM technique can be used for the communication. It is more feasible than to install detectors on the road. It would be much easier for maintenance too.

b) A strong client-server application has to be developed for a central control unit and well trained human resource is prerequisite for the maintenance of this central control unit.

c) Information gateway installations would be much easier than the first two. Mobile phone or transistor radios are very common for the city dwellers. The service providers are already providing different sort of information such as latest weather report, news etc though mobile phone. As a result this bus information can easily
be incorporated. Internet is also within the reach of most of the people. The main installation involves the VMS at the bus stoppage and its maintenance. Small size VMS is suitable for this purpose with proper data receiver from the central control unit.

6. Conclusion

The Urban public transportation system continues to loose market, especially in developed countries, since they fail to meet certain specification regarding their time, frequency of availability etc (Lin and Zeng 1999). Information system is providing them much better solution to that problem. The technology discussed here can have a perfect solution for the developing country like Bangladesh. The describe data driven system can ensure a perfect operation of bus network keeping a good coordination between all the relevant issues such as economy, resource, time and safety. The infrastructural and the implementation issues have to be taken care for the successful operation of the system. It surely can promise to increase the number of public transport users to a great level.

References

A MULTINOMIAL LOGIT MODEL FOR SERVICE QUALITY MEASUREMENT

Laura EBOLI and Gabriella MAZZULLA

Abstract. In this paper a Multinomial Logit model (MNL), calibrated by using mixed RP and SP data, is proposed. The model provides a way for measuring service quality in public transport. By means of this model the importance of service quality attributes on global customer satisfaction can be identified; in addition, a Service Quality Index (SQI) can be calculated. This index provides an operationally appealing measure of current or potential service effectiveness.

1. Introduction

Over the last few decades, Logit models have been widely used for the calibration of the mode choice models in which the alternatives are different transport modes. However, more recently “within mode” models have been proposed, in which the alternatives relate to an only transport mode, usually public transport mode. In this case, each alternative of choice is a bus service package characterized by some service quality attributes. For this type of models the use of Revealed and Stated Preferences data (RP, SP) is more convenient because SP data permit the introduction of some attributes and attribute levels not available in the actual context [16, 18].

In this paper a Multinomial Logit model (MNL), which allows the measurement of public transport service quality, is proposed. The model was calibrated by using mixed RP and SP data collected by a sample survey.

The paper is organized as follows. In the next section, the techniques usually adopted for measuring public transport service quality are introduced. The third section is about the description of the sample survey. The forth section deals with the results of the model calibration. In the last section, a brief conclusion is reported.

1 Department of Territorial Planning, University of Calabria, via P. Bucci, cubo 46/B, 87036 Rende (CS), laura.eboli@unical.it, g.mazzulla@unical.it
2. The Service Quality Measurement

Generally, transit agencies have given too much importance to saving money at the expense of service quality levels; therefore they have essentially focused on cost efficiency and cost effectiveness. A measure of cost efficiency is typically defined as produced services (e.g. vehicle kilometres), while a measure of service effectiveness is defined as consumed services (e.g. passenger kilometres).

However, transit agencies actually have an interest in obtaining a high service quality level taking into account passengers priorities and requirements.

By consulting the literature, a large set of service quality attributes can be identified, see for example [19, 22]. All the attributes can be grouped in macro-factors defined by one or more attributes. Some examples of these are transport network design (e.g. number and regularity of bus stops, having stops near destination), service supply and reliability (e.g. frequency, regularity and punctuality of rides), comfort (e.g. availability of seats on bus, bus overcrowding), fare (e.g. fairness/consistency of fare structure, ease of paying fare), information (e.g. availability of information on schedules/maps, explanation and announcement of delays), safety (e.g. safe and competent drivers, security for crimes), relationship with personnel (e.g. friendly, courteous personnel), customer preservation (e.g. repayment, complaint number), environmental protection (e.g. use of vehicles with low environmental impact), quality of system (quality of stops furniture, cleanliness of bus exterior).

All the attributes have an influence on service quality level, each one in a different measure. Therefore there is the necessity to quantify the importance of each one. In the literature there are many techniques for measuring importance of service quality attributes. These techniques can be identified in two different categories.

The first one includes methods of statistical analysis for determining the impact of the attributes on global service quality and customer satisfaction, based on the statements of a sample of passengers. They express an opinion by ranking or rating some service quality attributes. For a more detailed discussion of these techniques one should refer to Akan [1], Berger [3], Bhave [4], Cuomo [7], Hartikainen et alii [9], Hill [13, 14], Kano [15], and Parasuraman et alii [17].

The second category consists in the estimation of the coefficients by modelling. The models relate global service quality (dependent variable) to service quality attributes (independent variables). There are linear models, like multiple regression models, and non-linear models, like Structural Equation Models (SEM), which relate latent, observed and error variables [5], and Logit models, i.e. discrete choice models based on random utility theory [2, 6].

An example of modelling was proposed by Jones [21], in which a linear regression model for each interviewed user was calibrated; for each coefficient, a mean of estimated coefficients from individual models was calculated.

Examples of SEM are reported in Vilares and Coelho [23] and in Gronholdt and Martensen [8].
Some Logit models were proposed by Hensher [10, 11, 12, 19]. By the estimation of coefficients of the models, like Multinomial or Mixed Logit, the importance of service quality attributes on global customer satisfaction is evaluated. A Service Quality Index (SQI) is calculated as a linear combination of attributes by using estimated coefficients. These models were calibrated by using the combination of RP and SP data.

3. Experimental Context

The sample survey was addressed to the University of Calabria students. The campus is sited in the urban area of Cosenza, in the South of Italy. It is attended by approximately 32,000 students and 2,000 members of staff (March 2006).

At the present time, the University is served by bus services which connect the urban area with the campus; extra-urban bus services connect the campus with the other towns of Calabria. In a working day, about 8,000 students travel by urban bus. Interviews were directed to a sample of 320 students who live in the urban area and habitually use the bus to reach the campus. Therefore, the sampling rate is approximately equal to 4%. Respondents were asked to provide information about their trip habits regarding getting to the university and, in addition, about public transport service quality.

The interview is divided into four sections. In the first and second section some information about socioeconomic characteristics (gender, age, income, car availability, number of cars owned in family, number of family members, number of family members with driving licence) and travel habits was elicited.

In the third section the user gave some information about 16 service quality attributes. The users expressed a rate of importance and a rate of satisfaction on each attribute, on a scale from 1 to 10.

In the last section an RP-SP experiment was proposed to the sample. In the experiment each user described his RP context and compared it with some SP contexts in order to make a choice. In this way, it has been possible to consider some hypothetical contexts not available for the users at the time of the survey. In order to define the SP treatments, 9 service quality attributes were selected. Each attribute varies on two levels producing different alternatives, i.e. bus packages (table 1).

<table>
<thead>
<tr>
<th>Service quality attributes</th>
<th>Levels</th>
</tr>
</thead>
<tbody>
<tr>
<td>Walking distance to the bus stop</td>
<td>same as now (1); 10 minutes more (0)</td>
</tr>
<tr>
<td>Frequency</td>
<td>every 15 minutes (1); same as now (0)</td>
</tr>
<tr>
<td>Reliability</td>
<td>on time (1); late (0)</td>
</tr>
<tr>
<td>Bus stop facilities</td>
<td>bus shelter, seats and lighting (1) no shelter, no seats, no lighting (0)</td>
</tr>
<tr>
<td>Bus crowding</td>
<td>no overcrowded (1); overcrowded (0)</td>
</tr>
<tr>
<td>Cleanliness</td>
<td>clean enough (1); not clean enough (0)</td>
</tr>
<tr>
<td>Fare</td>
<td>25% more than the current fare (1); same as now (0)</td>
</tr>
<tr>
<td>Information</td>
<td>timetable, map, announcement of delays (1) no timetable, no map, no announcement of delays (0)</td>
</tr>
<tr>
<td>Transit personnel attitude</td>
<td>very friendly (1); very unfriendly (0)</td>
</tr>
</tbody>
</table>

Table 1. Service quality attributes and levels.
The full factorial design includes all the possible combinations among the attribute levels. In this case, it consists of 29 combinations producing 512 treatments. We restricted the number of treatments to 50 representing the SP choice alternatives. These alternatives were coupled in order to produce 32 different couplings. Each RP-SP experiment is composed by the RP alternative and one SP couple. Two experiments were proposed to each user and then, 640 observations were generated. The attribute weights were estimated according to the user choices; in this way, the current and potential service quality levels were calculated. An example of RP-SP experiment is shown in table 2.

The sample is evenly spread between male and female. The age range is between 18 and 31, but 73.1% of the sample is between 19 and 22. The sample was divided, also, in “in course”, and “out course” students; in Italy, the “out course” condition relates to a university student who has not finished his studies in the prescribed time. The “in course” students represent a percentage of 77.5% of the total.

<table>
<thead>
<tr>
<th>Attributes</th>
<th>Actual service</th>
<th>Service bus A</th>
<th>Service bus B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Walking distance to the bus stop</td>
<td>same as now</td>
<td>10 minutes more</td>
<td>same as now</td>
</tr>
<tr>
<td>Frequency</td>
<td>same as now</td>
<td>same as now</td>
<td>every 15 minutes</td>
</tr>
<tr>
<td>Reliability</td>
<td>no shelter, no seats, no lighting</td>
<td>Bus shelter, seats and lighting</td>
<td>no shelter, no seats, no lighting</td>
</tr>
<tr>
<td>Bus stop facilities</td>
<td>no shelter, no seats, no lighting</td>
<td>Bus shelter, seats and lighting</td>
<td>no shelter, no seats, no lighting</td>
</tr>
<tr>
<td>Bus crowding</td>
<td>overcrowded</td>
<td>overcrowded</td>
<td>no overcrowded</td>
</tr>
<tr>
<td>Cleanliness</td>
<td>clean enough</td>
<td>clean enough</td>
<td>not clean enough</td>
</tr>
<tr>
<td>Fare</td>
<td>same as now</td>
<td>same as now</td>
<td>25% more than the current fare</td>
</tr>
<tr>
<td>Information</td>
<td>no timetable, no map, no announcement of delays</td>
<td>timetable, map, announcement of delays</td>
<td>no timetable, no map, no announcement of delays</td>
</tr>
<tr>
<td>Transit personnel attitude</td>
<td>very friendly</td>
<td>very friendly</td>
<td>very unfriendly</td>
</tr>
</tbody>
</table>

Table 2. Example of an RP-SP experiment proposed to the interviewed.

About 45% of interviewed students belong to a middle class of family income and about 40% to a lower-middle class. Almost all the students do not have the possibility of using a car to reach the campus (91.8%).

4. Results

An MNL choice model, in which all random components are independently and identically distributed, is proposed. The model was specified and calibrated by using mixed RP and SP data. Each alternative is a bus service package, real or hypothetical. The utility function of the SP alternatives is a linear combination of the service quality attributes. In the utility function of the RP alternative some socioeconomic variable are included in addition to the
service quality attributes. The utility of each alternative is an index of service quality (SQI) of each bus package and the parameter values are the attribute weights.

All the service quality attributes are defined as dichotomous variables, except WALK and FARE that are continuous, measured in minutes and in Euros respectively. The values of the dichotomous variables are defined like the attribute levels reported in table 1.

Two socioeconomic characteristics are included in the RP utility function: gender (SEX, a dummy variable equal to 1 if the student is female and 0 otherwise) and car availability (CAR, a dummy variable equal to 1 if the student does not have the possibility of using a car to reach the campus and 0 otherwise).

Considering the joint use of RP and SP data, the scaling estimation methodology was adopted. The scale factor of the conjoint RP-SP model was estimated by means of the procedure proposed by Swait and Louviere [20]. However, in this case it assumes a value about equal to 0.9, because the variability between the RP and SP alternatives is not considerable.

The results of calibration are shown in table 3. All parameters have a correct sign and assume a value statistically different from zero, at a 95% level of significance.

<table>
<thead>
<tr>
<th>variable</th>
<th>acronym</th>
<th>parameter</th>
<th>estimation</th>
<th>t-student</th>
</tr>
</thead>
<tbody>
<tr>
<td>Walking distance to the bus stop</td>
<td>WALK</td>
<td>$\beta_{WALK}$</td>
<td>-0.115</td>
<td>-4.1</td>
</tr>
<tr>
<td>Frequency</td>
<td>FRE</td>
<td>$\beta_{FRE}$</td>
<td>2.923</td>
<td>11.3</td>
</tr>
<tr>
<td>Reliability</td>
<td>REL</td>
<td>$\beta_{REL}$</td>
<td>1.229</td>
<td>7.4</td>
</tr>
<tr>
<td>Bus stop facilities</td>
<td>STOP</td>
<td>$\beta_{STOP}$</td>
<td>0.608</td>
<td>3.8</td>
</tr>
<tr>
<td>Bus crowding</td>
<td>CROW</td>
<td>$\beta_{CROW}$</td>
<td>0.648</td>
<td>3.2</td>
</tr>
<tr>
<td>Cleanliness</td>
<td>CLEAN</td>
<td>$\beta_{CLEAN}$</td>
<td>0.815</td>
<td>5.2</td>
</tr>
<tr>
<td>Fare</td>
<td>FARE</td>
<td>$\beta_{FARE}$</td>
<td>-5.962</td>
<td>-4.0</td>
</tr>
<tr>
<td>Information</td>
<td>INF</td>
<td>$\beta_{INF}$</td>
<td>0.580</td>
<td>3.6</td>
</tr>
<tr>
<td>Transit personnel attitude</td>
<td>PER</td>
<td>$\beta_{PER}$</td>
<td>0.482</td>
<td>3.2</td>
</tr>
<tr>
<td>Gender</td>
<td>SEX</td>
<td>$\beta_{SEX}$</td>
<td>0.406</td>
<td>2.2</td>
</tr>
<tr>
<td>Car availability</td>
<td>CAR</td>
<td>$\beta_{CAR}$</td>
<td>0.659</td>
<td>3.7</td>
</tr>
</tbody>
</table>

Table 3. Results of the model estimation.

As expected, WALK and FARE assume a negative sign because an increase of fare and distance from the bus stop involves a decrease of utility. SEX and CAR assume a positive sign; this result indicates that the RP utility has a higher value, ceteris paribus, for the students of female gender and for the students who do not have car availability. These categories of students are more satisfied with the actual bus service.
The model verifies the statistical tests on the goodness-of-fit. Rho-squared is equal to 0.343; the LR statistics is widely higher than critical value with 11 d.o.f.; the % Right has a good value (about 67%).

The parameter of the variable FARE has a high absolute value because the sample is mainly composed by students with a middle or lower-middle family income.

Short of the attributes with a negative parameter, the frequency is the attribute with the higher weight on global service quality. Indeed, an increase of the frequency from a bus every an hour to a bus every 15 minutes produces, ceteris paribus, an increase of about 2.9 on SQI.

Other important attributes are the reliability, the cleanliness and the bus stop facilities. A simultaneous improvement of the three attributes is comparable to an improvement in terms of frequency.

5. Conclusions

In this research a tool of measuring public transport service quality is provided. An MNL model has been introduced. This model was calibrated by using a mixed RP and SP data collected by a survey addressed to the University students.

The model calibrated verifies the statistical tests on the goodness-of-fit. All parameters have a correct sign and assume a value statistically different from zero, at a 95% level of significance. The service frequency is a statistically strong attribute. The parameter of the variable FARE has a high value because the sample is composed by students with a middle or lower-middle family income.

The utility of each alternative is an index of service quality (SQI) of each bus package and the values of the parameters are the attribute weights.

This index is useful to planners and transit operators for measuring the importance of service quality attributes and for investing on some attributes in order to improve the service.

Similar models can be proposed by introducing other service quality attributes or socioeconomic characteristics.

References


MODELING THE IMPACT OF TRAVEL INFORMATION ON ACTIVITY-TRAVEL DECISIONS UNDER UNCERTAINTY: HEURISTICS AND EXAMPLE

Zhongwei SUN¹, Theo ARENTZE², H.J.P TIMMERMANS³

Abstract. Travel information is widely available in different forms through various sources. This wide availability of travel information is believed to influence travellers’ daily activity-travel patterns. To understand and model the impact of travel information on traveller’s decision processes, the authors proposed a Bayesian framework to represent the decision and learning process under conditions of uncertainty, given travel information and non-stationary transportation environments. The value of travel information is conceptualized as the extent to which the information allows an individual to make better decisions by reducing the uncertainty during both the scheduling and rescheduling stages.

1. Introduction

Traffic information is assumed to have a positive impact on the transportation network by affecting individuals travel behaviour. Researchers believe precise and real-time traffic information leads to effective usage of current transportation networks (e.g. [11], [4],[7]). Personal Intelligent Travel Assistant (PITA) or Advanced Travel Information System (ATIS) are viewed as valuable means of optimising the use of the existing transport capacity and thus has been seen as an effective mitigation to congestions.

Findings indicate that the adoption of ATIS information is linked to both the traveller’s past experience and the credibility of the information source. Studies(e.g.[6], [12]) found in their laboratory experiments about how individuals adjust their departure time choice and route choice behaviour in response to previous experience, that the most recent information, essentially the previous trip’s travel time, is the most important factor in current decisions.

¹ Eindhoven university of technology, PO box 513, the Netherlands, z.sun@bwk.tue.nl
² Eindhoven university of technology, PO box 513, the Netherlands, t.a.arentze@bwk.tue.nl
³ Eindhoven university of technology, PO box 513, the Netherlands, h.j.p.timmermans@bwk.tue.nl
Many similar studies address one aspect of travel behaviour, such as departure time or route choice and produced plenty of results regarding different facets of travel decisions. Nevertheless, focusing on only one aspect of travel decisions is limited in the sense that travel decisions are interrelated to each other in a complete daily activity-travel pattern.

The provision of travel information, pre-trip and en-route, implies a potential reduction of uncertainty about the state of the transportation network. Based on the information and his/her own knowledge of the travel environment, a traveller can evaluate alternatives and make decisions under uncertainty. By executing these decisions the traveller compares his/her beliefs against reality and gradually he/she will learn the real traffic situation and the credibility of the information source.

To understand and model the impact of travel information on dynamics of traveller’s decision processes, the authors have proposed a Bayesian framework ([5], [2], [15], [1]). The value of information is conceptualized as the extent to which information about travel time (or any other relevant state of the system) improves activity-travel scheduling decisions at the beginning of the day and during execution of a schedule. The system takes the complete activity travel pattern of one day into account and explicitly represents the uncertainty and the perception of credibility of information source. This framework is based on expected utility theory and Bayesian principles of probabilistic inference, belief updating and (parameter and structure) learning, which automatically incorporates the decision process and learning process of travellers under information provision. The purpose of this paper is to describe the implementation and an illustration of the framework. First, we will briefly describe the modelling framework in more detail.

2. MODEL FRAMEWORK

Activity-based modelling looks at travel as a derivative of the activities an individual wishes to conduct in space and time [10], [17], [3]. It is considered an important approach for travel demand management because it produces more insight into the traveller’s decision-making process and behaviour. It is also considered a better method of analysis for policy evaluation and impact assessment because direct and indirect effects of travel demand measures are captured.

An important concept in the present model is the possible re-scheduling of activities in response to unforeseen events during the execution of the schedule in a particular day. Re-scheduling is assumed to be considered only if due to an unforeseen event or unanticipated travel time, the end time of the last (activity or travel) episode differs from the scheduled end time or when the individual receives new information regarding future events (conditions, circumstances).

Figure 1 shows the conceptual composition of this framework. A concise description of the framework is as follows. A traveller has his own perception or belief about the traffic situation and credibility of information service. Based on his belief together with information received, the traveller makes travel decisions and formulates his schedule from scratch or changes an existing schedule. After execution of his decisions, i.e. a trip or activity, the traveller updates his belief about the traffic and the credibility of information service using his new experience. Following sections describe the basic components of this framework.
2.1. Decision under uncertainty and learning

The decision problem is modelled as a decision tree representing all paths of possible decisions and outcomes of uncertain events. To avoid the exponential grow of the branches, we further develop the model in terms of heuristics individuals may use to reduce the number of combinations, in intelligent ways. For details we refer readers to previous work [2], [16].

We assume that a traveller only has partial knowledge of the travel environment. And we further assume that to deal with uncertainty, the traveller uses a scenario-based approach. Each scenario represents a possible outcome of the combinations of possible decisions and outcomes of uncertain events. The traveller assigns probabilities to each outcome indicating how likely this outcome will be in his perception. Under the assumption of rational behaviour, possible schedule decisions are evaluated under each possible outcome and the decision that maximizes the expected utility is chosen. The value of information is conceptualised as the expected difference in the utility of the best choice between the situation with and without information.

A probability distribution across the outcomes of an event to represents the uncertainty. It is further conditioned upon context variables (which are the outcome of structure learning). Assume $Y$ is the information received (before the event) and $X$ is the outcome experienced (after the event) towards the same uncertain event, $P(Y | X)$ then represents the individuals beliefs regarding the probability of a message $Y$ (i.e., ‘the travel time will be $Y$ minutes’) given the true state $X$ (i.e., the true travel time is $X$ minutes). Hence, it reflects the degree of credibility the individual attaches to the information source.
The probabilities representing an individual’s beliefs are updated each time an event is experienced during the implementation of the schedule. Spiegelhalter [14] method derived from Bayesian principles is used to update the beliefs:

\[
P^{t+1}(y_j | x_k) = \frac{P^t(y_j | x_k)W^t(x_k) + w^t}{W^t(x_k) + w^t}
\] (1)

\[
P^{t+1}(y_j | x_k) = \frac{P^t(y_j | x_k)W^t(x_k)}{W^t(x_k) + w^t}
\] \( \forall j \neq i \) (2)

\[W^{t+1}(x_k) = W^t(x_k) + w^t
\] (3)

\[
P^{t+1}(y_j | x_i) = P^t(y_j | x_i)
\] \( \forall j,i \neq k \) (4)

\[W^{t+1}(x_i) = W^t(x_i)
\] \( \forall j,i \neq k \) (5)

where \( y_i \) denotes the observed state of \( Y \), \( x_k \) is the observed state of \( X \), \( w_i \) is the weight assigned to the \( t \)-th case and \( P^t(y | x) \) is the probability representing the individual’s updated belief in \( Y = y \) given that \( X = x \). In applications, \( Y \) represents some uncertain event and \( X \) represents the contextual variables on which the beliefs are dependent. \( W^t \) denotes a running total which can be seen as a measure of the amount of experience accumulated until time \( t \). It implies the impact of the past experience on belief updating. In non-stationary systems it is rational to assign more weight to more recent cases, namely, \( w^t \) should decay as a function of time. The perceived credibility of an information source is updated based on the same mechanism.

2.2. Scheduling Engine

An evolutionary algorithm (in short EA) scheduling engine was developed to generate near optimal schedules for the traveller, similar to the work of Meister[13]. This schedule engine use the same form of utility as Aurora ([9], [8]) which assumes the utility of an activity is a continuous function of the activity’s duration, an asymmetric S shape, for each activity.

\[U_a = \frac{U_a^{\text{max}}}{(1 + \gamma_a \exp[\beta_a(\alpha_a - v_a)])^{1/\gamma_a}}
\] (6)

where \( U_a \) is the utility of activity \( a \), \( U_a^{\text{max}} \) is the asymptotic maximum utility, \( v_a \) is the activity duration and \( \alpha_a, \beta_a \) and \( \gamma_a \) are parameters. \( \alpha_a \) denotes the duration where the marginal utility changes from increasing to decreasing; \( \beta_a \) refers to the slope of the curve and \( \gamma_a \) defines the degree of a-symmetry of the S-shape. For the detailed definition of \( U_a^{\text{max}} \), readers are referred to [9].

In Aurora, the utility of a travel is defined as a linear, negative function of travel time with mode specific intercept and slope as parameters. Utility effects of location, time of day and possibly other attributes of activities are represented by increases or decreases of \( U^{\text{max}} \).
Our schedule engine follows this. The total utility of a given schedule is then defined as the sum of the utilities across all travel and activity episodes in the schedule.

The EA schedule engine is trying to maximize the sum of utilities across all activities and their connected travel episodes. We encoded the whole schedule as a chromosome and each single activity as a gene section. For simplicity, the engine uses only mutation actions and no crossovers. Although this may somehow slowdown the searching process and deviates from the classic genetic algorithm, it reduces the complexity of the program. The fitness function is defined as the sum of utilities across activities and travel in the solution. The space-time and schedule resource constraints are captured by means of penalties in the fitness function. If a shopping activity is scheduled at 5AM, which is out of the time window for the shopping activity, the chromosome fitness receives a relatively large negative penalty. Based on the fitness, a tournament selection applied to all the populations with a tournament size 2 and 0.8 probability of copying winner to new population.

The input of this schedule engine is a predefined schedule (during execution) or an empty schedule (start of the day) and the current time. The output is a schedule that is near optimal regarding the section starting from the current time to the end of the day. The part that has already been executed, i.e. the section before the current time, is taken into account in the fitness function but, obviously, is not subject to change.

2.2.1. Myopic heuristic in scheduling

We propose that individuals adopt a myopic view when they generate or revise a schedule under conditions of uncertainty for a given day [16]. This heuristic is mainly meant to avoid the combinatorial problem arising when facing multiple uncertain events. This heuristic assumes the traveler considering one uncertain event, with all its possible outcomes, at a time rather than considering the combinatorial space of all possible outcomes of all uncertain events in a day. It also assumes that uncertain events relate exclusively to travel times and, therefore, only trips and trip facet choices are considered uncertain. The heuristic is formulated as:

1. Select one activity from the activity list at time T, (Initial T = 0)
2. Schedule this activity to be the next activity after T and generate a schedule for the remaining day based on best guesses of travel times for each link in the network
3. Determine the decision tree representing all possible Decision-Outcome (in short, DO) scenarios
4. Evaluate the decision tree and make a decision regarding choice facets of the trip
5. Determine the utility of the resulting schedule
6. Select next activity in the list and repeat from step 2 until the last
7. Choose the activity that yields the highest utility and choose its decision sequence that gives this utility and remove that activity from the list
8. Execute the trip
9. Set T to the current time and repeat from 1

This heuristic is an extended version of the one described in earlier work ([16]).
3. Numerical illustration

In this section, we apply the decision heuristic to a hypothetical example to illustrate how the framework works. The framework is implemented in Java.

The hypothetical transportation network used in this simulation is shown in Figure 2. A traveller has to make his schedule and executes it for one day. There are activity types, a mandatory work activity with fixed time, location and duration, a flexible (may be excluded) shopping activity with two possible locations and a mandatory in-home activity. A time window constraint exists saying that no shopping can happen before 8am or after 8pm. The traveller has initial beliefs on the traffic situation, in terms of a probability distribution across traffic speed categories (High, middle, low) for each link. For each route, we always categorize the travel time into 3 categories. This is based on modal values for each of the 3 possible speed categories, namely 100km/h (high), 70km/h (middle) and 40km/h (low).

The credibility of PITA was set to 80 percent by means of the conditional probability table shown in Table 1.

<table>
<thead>
<tr>
<th>PITA Travel time</th>
<th>High</th>
<th>mid</th>
<th>Low</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>0.8</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>mid</td>
<td>0.1</td>
<td>0.8</td>
<td>0.1</td>
</tr>
<tr>
<td>Low</td>
<td>0.1</td>
<td>0.1</td>
<td>0.8</td>
</tr>
</tbody>
</table>

Table 1. Conditional probabilities of travel time information given true travel times (represented in rows)

In the schedule generated under best guesses of travel times, the shopping activity is positioned directly after the work activity (i.e., without returning home in-between) at location 4 (shop2). In fully detail, this initial schedule of the day can be represented as:
Starting from the first home activity, the traveler has two options, to go shopping first or to go to work. Following the Steps 2 to 5 of the heuristic, the system evaluates each DO scenario using two decision trees, DT1 for work and DT2 for shopping. DT1 and DT2 have their highest utilities 151.769 and 152.011 respectively. Thus the decision is the branch in DT2 which gives the utility of 152.011, the corresponding decision attributes are location 4, mode 1 route 5 and travel start time 538(8:58am).

To execute this decision, the system generates a random real travel time of 19 minutes based on the true speed probability distribution and the traveler checked whether this is consistent with his believed travel time which was used to make the decision. If not, a rescheduling process has to be activated to adapt to the new travel time. This procedure continues until the end of the day is reached and results in a final schedule:

**Activity** | StartTime | Duration | Location | Route | TravelTime
--- | --- | --- | --- | --- | ---
HOME | 0 | 571 | 1 | 2 | 14
Shop | 948 | 32 | 3 | 3 | 14
HOME | 994 | 446 | 1 | 1 | Utility=152.476

| **Activity** | StartTime | Duration | Location | Route | TravelTime |
--- | --- | --- | --- | --- | ---
HOME | 0 | 538 | 1 | 5 | 19
Shop | 557 | 25 | 4 | 8 | 15
HOME | 972 | 468 | 1 | 2 | Utility: 149.671

### 4. CONCLUSION AND DISCUSSION

In this paper we implemented and further developed a previously suggested conceptual framework for travel (re)scheduling behaviour under information provision. We have assumed that travellers hold beliefs with respect to links in the transportation network. Whereas most previous studies have considered a single source of uncertainty, making inferences based on links may accommodate the case that travellers may be faced with multiple delays occurring at different links. To mitigate the problem of combinatorics, we assume that travellers adopt a simplified decision heuristics in a myopic way. That is, they only consider the uncertainty related to a next trip in a preliminary schedule (based on best guesses). An evolutionary algorithm scheduling engine was set up to generate, revise and evaluate a full-day schedule at every decision moment at the start of the day and during execution. With this framework, we can test and observe the single travel facet decision as well as the full-day schedule pattern. By running across many days, learning effects on the level of the transport system as well as information service can also be taken into account.

Several aspects of this framework can be further improved. The heuristic does not take into account uncertainty further away from the current decision. This makes it difficult to represent travellers’ attitude towards later uncertainties. The decision tree construction and evaluation may be further refined. Branches with minor probabilities could be eliminated according to some other, complementary heuristics such as leave out extremely unlikely happened outcomes or use some satisfying decision rule. Though the EA schedule engine
works fine, it is still quite primitive and lacks some importance characteristics such as time preferences and mode change features.

Reference:


USING NEURAL NETWORKS FOR ESTIMATING SATURATION FLOW RATES AT SIGNALIZED INTERSECTIONS

Isabela QUEIROZ\textsuperscript{1}
Maria Alice Prudêncio JACQUES\textsuperscript{1}
Mark DOUGHERTY\textsuperscript{2}
Li WEIGANG\textsuperscript{1}

Abstract. Field measurements of saturation flow rates are not possible in many cases. Therefore, this factor is usually estimated by models. Usually these models have been developed based on regression analysis. However, new modeling techniques are now available and have been applied in the transportation field, as is the case with artificial neural networks (ANNs). The objective of this work is to investigate the advantages of the use of ANNs in the development of saturation flow models vis-à-vis the use of regression analysis. In light of this, models using the two modeling techniques were developed for the same database comprising data from forty traffic lanes located in Brasília, Brazil. Afterwards, the models were tested on ten traffic lanes in the same city, and the results showed that the use of ANNs produced better estimates than those generated by the corresponding regression model.

1. Introduction

The control of the traffic at road junctions may be carried out through different means, traffic lights being one of the most used in central urban areas. When signal use is technically justified and its timings are properly calculated to handle the traffic needs at each intersection, it constitutes a very efficient way of promoting the safe movement of pedestrians and vehicles at the intersection level. However, the definition of signal timings is highly dependent on knowledge of traffic behaviour at intersection approaches. This is particularly pertinent in relation to saturation flow rates. Saturation flow rate is defined as the maximum hourly traffic flow rate that can cross a lane or intersection approach under

\textsuperscript{1} University of Brasilia, Brazil, mapj@unb.br
\textsuperscript{2} Dalarna University, Sweden, mdo@du.se
prevailing traffic and road conditions. In this case, it is assumed that the green signal indication is available during the whole hour and that there are no lost times.

This study aims to investigate the use of ANNs as an alternative technique in the development of a model for saturation flow rate estimation vis-à-vis a regression analysis model. This investigation follows two procedures. Firstly, the results generated by the model developed with the aid of neural networks are analysed to evaluate the generalisation capacity of the trained network. Secondly, the results produced by the trained neural network are compared to those generated by a regression model. This regression model was developed based on the same database used for the neural network's training. The development of the regression model is detailed in Queiroz (2001).

2. Background

Many studies related to the analysis of the main factors affecting saturation flow rates at signalized intersections are presented in the literature on the state of the art. Among them, Kimber and Semmens (1982), Kimber et al. (1986), Niittymaki and Pursula (1997) investigate different factors related to the physical and operational characteristics of the approaches (or traffic lanes) and to drivers’ behavior. Despite the differences in the existing models’ expressions, some variables are presented in almost all of them, as is the case with traffic composition, turning movements, lane width and grade. This means that although there is a reasonable agreement among the variables identified as explanatory for saturation flow estimation purpose, their relationships differ from model to model.

The major problem in regression modeling is to define the model structure that should be calibrated (the type of relationships among the independent variables). It is specially a hard task to identify non-linear terms (quadratic and higher orders, cross terms, etc.) to be incorporated into the model.

Neural networks, as well as other modeling techniques, present advantages and disadvantages. According to Dougherty (1995) it is important to investigate the extent to which the use of neural network for a given problem is worthwhile. In this sense, one important aspect is to test the neural network's performance in relation to traditional modelling techniques, as is the case with regression analysis. Artificial Neural Networks are parallel and distributed systems. They can be classified into different models based on the network's topology and on the learning algorithm used. The multilayer Perceptron (MLP) used with the Back Propagation learning algorithm is the model chosen for the present application.

3. Models Development

The development of the models for estimating saturation flow rates was elaborated according to the procedure presented below.
3.1. Data collection

Twenty distinct intersections in the central area of Brasília were selected for the study. This produced a sample of fifty (50) traffic lanes to be observed. The necessary data at the traffic lanes were gathered through different means, such as recording of the traffic discharge process, direct field observations and topographic surveys. By the recording of intersections, information about the saturation flow rate, traffic composition and types of movements were obtained. The saturation flow rate was calculated according to the method described in the Road Note 34 (RRL, 1963). In the field, one trained person took notes about situations related to the traffic operational conditions. The data on road geometric characteristics such as grade, lane width and turning radius were collected from topographic surveys.

3.2. The generation of the Model Using ANN

3.2.1. Preliminary data treatment and network architecture definition

The data treatment involves preparing the available data for their use in the network modelling process. This comprises two stages. The first is data normalisation, which aims to produce a certain degree of homogenisation in the scales of the different variables by compressing their values into a common interval. The second stage consists of splitting the available data into two sets called training sequence and testing sequence. Forty of the fifty traffic lanes were selected for the network training sequence, and ten lanes for the network testing sequence.

Two normalisation intervals were tested in this work: the interval between zero and one; and the interval between 0.05 and 0.95. Within these intervals three strategies of normalisation of the database were tested. Strategy Type 1, which normalises the entire database globally; Strategy Type 2 that is the normalisation of each variable separately; and strategy Type 3 that is the normalisation of the input data and the output data separately. The results of the network training for each of the six possibilities are presented in Section 3.2.2.

Many tests were performed in order to define the architecture of the neural network which is as follows:

- Neural network model: multilayer perceptron (MLP);
- Learning algorithm: backpropagation;
- Number of hidden layers: one;
- Number of neurons in the hidden layer: four
- Activation function: linear function and logistic function;
- Rate of learning: 0.02.
3.2.2. Training of the network and corresponding results

In order to train the network, a sequence of training data is presented to it. The training sequence data is a set of vectors presenting values for the input variables and the corresponding value for the output variable. The output variable is the saturation flow rate, and the input variables are: grade; width; location; fluidity condition; proportion of right turning vehicles; proportion of left turning vehicles; and turning radius. For this training, two different data sets were developed. The first set was formed by 40 (forty) training vectors. The second set was formed by 1035 (one thousand thirty five) vectors. The two sets were then used for training the network according to the 6 (six) different training conditions specified in Sub-section 3.2.1.

The results of these different training conditions applied to the lanes belonging to the testing sequence data are presented in Table 1. These results were obtained by calculating the estimate relative error between the values of the saturation flow measured in loco ($FS_{\text{loc}}$) and the value estimated by the neural network ($FS_{\text{ob}}$). The analysis of Table 1 shows that the minimum mean relative error was produced when the sequence of training with forty vectors was used and normalised globally (strategy Type 1) over the interval of 0 to 1. The results also show that the performance of the trained networks is very sensitive to differences in the training specified conditions and to the number of vectors used for this training.

<table>
<thead>
<tr>
<th>Normalisation Interval</th>
<th>Number of vectors in the training set</th>
<th>Mean relative error (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Type 1</td>
</tr>
<tr>
<td>0 to 1</td>
<td>40</td>
<td>4.10</td>
</tr>
<tr>
<td></td>
<td>1035</td>
<td>5.98</td>
</tr>
<tr>
<td>0.05 to 0.95</td>
<td>40</td>
<td>4.73</td>
</tr>
<tr>
<td></td>
<td>1035</td>
<td>4.86</td>
</tr>
</tbody>
</table>

Table 1: Mean error for the normalisation intervals studied per training condition.

3.3. Development of the Basic Regression Model

Initially the original model developed by Kimber et al. (1986) was calibrated to the complete database related to the fifty surveyed traffic lanes in Brasília. Based upon the results of this calibration, many analyses were performed in order to define the convenience of adding new variables to the initial model, as well as to remove from the model the variables that proved to be not statistically significant. The final expression for the calibrated analytical model is presented in Equation (1). The coefficient for determination of the calibrated model ($R^2$) is equal to 0.525, which is comparable to the coefficient found by Kimber et al. In addition, all the model’s parameters proved to be statistically significant at the level of 5%.

$$FS = \frac{(1705.65 - 26.48 \cdot G + 182.89(w - 3.25) - 140.58L + 256.51F)}{(1 + 1.5(VD + VE)/R)}$$
where:

$FS$: saturation flow rate per lane (FS), expressed in pcu/hg;

$G$: grade, expressed in % (positive values for uphill grades and negative values for downhill);

$w$: lane width, expressed in metres;

$L$: dummy variable related to the intersection location ($L=1$ in CDB areas; $L = 0$, otherwise);

$F$: dummy variable related to the fluidity condition downstream the approach; $F = 1$ when the condition is good/regular; and $F = 0$, otherwise;

$VD$: rate of right turning vehicles;

$VE$: rate of left turning vehicles;

$R$: turning radius, expressed in metres.

### 4. Models Estimating Performance

A comparison between the two modelling techniques is presented in Table 2. The results show that the model generated with the neural networks performed better than the analytical model. The benefit of using neural networks for saturation flow modelling was bigger especially when the 40-vector data base was used.

### 5. Conclusions

The model for estimating saturation flow rates developed in this work produced estimated errors that are in general 50% less than those produced by the corresponding regression models. Therefore, the advantage of using neural networks to model complex and random processes such as vehicle departures from signalized approaches was demonstrated.

However, it was also verified that modeling using MLP neural networks requires long and systematic investigation of all the important factors affecting these networks’ performance for each different application.
<table>
<thead>
<tr>
<th>Approach</th>
<th>Lane</th>
<th>Error (%)</th>
<th>Error (%)</th>
<th>Error (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Analytical models</td>
<td></td>
<td>Neural network</td>
</tr>
<tr>
<td></td>
<td></td>
<td>40 vectors</td>
<td>1035 vectors</td>
<td>40 vectors</td>
</tr>
<tr>
<td>Cruzeiro Novo</td>
<td>Left</td>
<td>12.878</td>
<td>16.406</td>
<td>1.72</td>
</tr>
<tr>
<td>Exit from the SIG/Eixo M.</td>
<td>Right</td>
<td>14.195</td>
<td>19.002</td>
<td>3.75</td>
</tr>
<tr>
<td>SCLN 105/106</td>
<td>Single</td>
<td>-10.032</td>
<td>-9.760</td>
<td>3.87</td>
</tr>
<tr>
<td>SCLN 107/108</td>
<td>Single</td>
<td>-11.065</td>
<td>-10.641</td>
<td>2.71</td>
</tr>
<tr>
<td>L2 Sul (Col. Santa Rosa)</td>
<td>Central</td>
<td>-3.003</td>
<td>1.027</td>
<td>-4.50</td>
</tr>
<tr>
<td>SCN (BSB Shopping)</td>
<td>Left</td>
<td>8.408</td>
<td>12.957</td>
<td>-1.24</td>
</tr>
<tr>
<td>Memorial JK</td>
<td>Central</td>
<td>-2.392</td>
<td>2.044</td>
<td>-1.64</td>
</tr>
<tr>
<td>Memorial JK</td>
<td>Right</td>
<td>-1.572</td>
<td>2.738</td>
<td>-1.13</td>
</tr>
<tr>
<td>W3 Norte-SHN (S-N dir.)</td>
<td>Left</td>
<td>-8.092</td>
<td>-2.070</td>
<td>-11.45</td>
</tr>
<tr>
<td>Mean relative error</td>
<td></td>
<td>9.16</td>
<td>9.909</td>
<td>4.10</td>
</tr>
</tbody>
</table>

Table 2: Estimate relative errors associated with the modelling techniques considered.

References


TWO-PHASED PARKING CHOICE METHOD FOR PRETRIP
PARKING GUIDANCE SYSTEM

Ji Yanjie¹, Deng Wei², Wang Wei³

Abstract. Aim at solving vehicles parking and traveling in a journey at the same
time, a two-phased parking choice method for pretrip parking guidance system
was proposed. In the first phase, four aspects including walking distance from
parking lots to destination, parking classification, real-time parking space and
parking fee were considered during the selection of candidate parking lots around
destination. In the second phase, five indexes consisting of walking distance from
parking lots to destination, parking feasibility, parking cost, parking safety and
journey time from origination to parking lots were selected to be optimization
parameter. Then multi-object decision making model based on fuzzy preference
was put forward to rank candidate parking lots.

1. Introduction

A problem faced in major metropolitan area, is the search for parking space that results in
invalid traffic flow on the road, excess pollution, and driver frustration. Parking guidance
system is an efficient way to solve the problem. Providing drivers with information on
parking availability would reduce search time and impact on congestion. Searching for
parking could be simplified to two kinds of choice: one is parking lots choice, and the other
is traveling routes choice. Without authentic parking and travel information before trips or
en-route, drivers usually search for parking lots with experience. And the searching results
could be divided into three kinds: one is arriving at parking lots successfully, one is having
no parking spaces availability, and the last one is that traffic congestion happens in the
route. The latter two results in traffic congestion and drivers can not arrive at destination on
time. Therefore it is an efficient way to develop parking choice advice before trips using
web site, telephone kiosk, etc. To realize the function, corresponding parking choice

¹ Transportation College, Southeast University, Nanjing, People's Republic of China;
email: yimuzi309@163.com
² ditto; email: dengwei@seu.edu.cn
³ ditto; email: wangwei@seu.edu.cn
method should be developed to help operation control center to process data efficiently and pass to terminal users through partition centers.

Although parking facilities’ locating models have been developed much in many countries in the past, and choice models which can simulate travelers’ parking choice behavior were developed to predict and manage [1]. The study on parking choice models for pretrip parking guidance system is sparse, however. Considering time-variant characteristic of traffic network and parking berth, two-phased parking choice method for pretrip parking guidance system is proposed in this paper. This paper begins by the discussion of parking decision evaluation index. It is followed by two-phased parking choice method: the first phase analyzed qualitatively the parking lots and yielded the candidate parking lots; and in the second phase, a multi-object decision making model is used to rank quantitatively those selected in the first phase. Finally, advantages of the two-phased method and issues need to further research are presented.

2. Parking decision evaluation index

Figure 1 presents the results of questionnaire survey on drivers’ parking behavior and it indicates that the most important factors influencing drivers’ selection in turn are walking distance from parking lots to destination, parking safety, parking fee and drivers’ familiarity with parking [2]. In addition, considering parking feasibility and drive route choice, real-time parking space availability and drive time should be included in the parking decision evaluation indexes. The real-time parking space can be predicted through different methods. Choice of travel route have different optimization criterion. Because travel time can reflect drivers’ familiarity and parking lots’ accessibility, shortest travel time is chosen as the optimization criterion. In sum, parking decision evaluation index system is composed of five indices as shown in table 1.

![Figure.1 The distribution proportion of main forces of parking choice](image)

<table>
<thead>
<tr>
<th>Evaluation aspect</th>
<th>Evaluation index</th>
<th>unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parking accessibility</td>
<td>walking distance from parking lots to destination</td>
<td>meter</td>
</tr>
<tr>
<td>Parking feasibility</td>
<td>parking space</td>
<td>entries</td>
</tr>
<tr>
<td>Parking safety</td>
<td>type of parking lots</td>
<td></td>
</tr>
<tr>
<td>Parking cost</td>
<td>parking fee</td>
<td>Yuan(RMB)/hour</td>
</tr>
<tr>
<td>Travel cost</td>
<td>travel time</td>
<td>minute</td>
</tr>
</tbody>
</table>

Table.1 Evaluation indices of parking decision
3. Two-phased parking choice method

Some assumptions of the proposed method are that (1) walking distance from parking lots to destination is calculated by the real line distance between parking lots and destination; (2) user optimal is the objective of the choice of parking lots and travel route; and (3) drivers will not wait when parking lots have no empty space.

3.1. First phase

To reduce combination of different parking lots and travel route, two-phased method is used to choice parking lots. The first phase is divided into four steps and four factors are included: walking distance from parking lots to destination, parking type, parking space availability and parking fee.

According to the investigation of driver behavior, it is inadvisable if the ideal radii are longer than 500 meters. So it is advised to choose parking lots within a 500-meter-radii circle whose center is the destination as candidates. Distance from parking lots to destination could be calculated the coordinates of the parking lots ($P_i$, $i=1, 2, \ldots, n$) and the destination ($P_v$). If drivers have special requirements on parking type, parking lots inconsistent with drivers’ requirement can be eliminated. Based on the two steps mentioned above, travel routes from start point to initializing candidate parking lots and the corresponding travel time could be decided. Then combined with parking space prediction, whether parking lots have empty space when drivers arrived at initializing candidate parking lots can be ascertained and parking lots without empty space should be eliminated.

Parking space is changed with time and the prediction method can reference [3]. The travel time from origination to parking lots ($T_{pi}$) is correlated with the travel route. Actually the optimal route is a one-to-one shortest path problem and the optimal path problem has many optimization criterions. According to the parking decision evaluation indexes mentioned above, travel time is chose as the optimization criterion. Travel time is time-dependent and stochastic because traffic volume is transformed dynamic and fluctuation randomly. The optimal path problem of time-dependent and stochastic traffic network can be found by SPTDN algorithm [4], or by a dynamic routing algorithm based on chaotic neural network [5], and so on. Combined with the time arrived at initializing candidate parking lot and intending parking time, parking fee is calculated. Parking cost is usually decided by parking time and there are two conditions. One is uniform price and the other is discriminatory price based on different parking time. The parking lot whose cost is less than the upper limit of expected parking cost can be chosen to candidate parking lots. Through the choice mentioned above, candidate parking lots ($P_1, P_2, \ldots, P_m$) could be confirmed.

3.2. Second phase

In the second phase, a simple and practical method for multi-object decision-making under fuzzy-preference is adopted to rank candidate parking lots selected in the first phase. In the paper, five indexes consisting of walking distance from parking lots to destination ($L$), parking feasibility ($B$), parking cost ($F$), parking safety ($S$) and travel time from origination...
to parking lots (T) are selected as optimization parameter and composed the object vector set \( G = \{L, B, F, S, T\} \). Candidate parking lots selected in the first phase composed decision-making alternatives set \( P = \{P_1, P_2, \ldots, P_m\} \). The attribute of \( P_i \) can be defined as \( \{y_{i1}, y_{i2}, y_{i3}, y_{i4}, y_{i5}\} \), where \( y_{i1}, y_{i2}, y_{i3}, y_{i4}, y_{i5} \) represent walking distance from parking lots \( P_i \) to destination, parking feasibility of \( P_i \), parking cost of \( P_i \), parking safety of \( P_i \), and journey time from origination to parking lots \( P_i \) respectively. The steps as followed:

Step 0: set up decision-making alternatives set \( P = \{P_1, P_2, \ldots, P_m\} \), objective vector set \( G = \{L, B, F, S, T\} \) and decision matrix \( A_{m \times 5} = (y_{ij})_{m \times 5} \);

Step 1: transform decision matrix \( A_{m \times 5} \) to standardization matrix \( Z_{m \times 5} = (z_{ij})_{m \times 5} \);

Step 2: compute the optimal weight set of multi-object \( W \);

Step 3: compute synthetic attribute of every candidate parking lots: \( D_i(w), i = 1, 2, \ldots, m \);

3.2.1. Quantification of indices

Walking distance, parking cost and journal time are all quantitative benefit indices and can be computed in the first phase. Real-time parking spaces reflected parking feasibility is a quantitative cost index and can be predicted. Quantitative indices of parking safety could be got as follows: parking lots within road is 3-5, parking lots off road is 5-7; parking lots underground is 7-9; stereoscopic garage is 7-9 and mechanical parking lots is 5-7 [2].

3.2.2. Data standardization

Decision matrix \( A_{m \times 5} \) shows the basic information of the decision-making problem and could be standardization. To three cost indices \( L, F, T \), standardization may be represented by equation (1):

\[
Z_{ij} = \frac{y_{ij}^{\text{max}} - y_{ij}^{\text{min}}}{y_{j}^{\text{max}} - y_{j}^{\text{min}}} (i = 1, 2, \ldots, m; j = 1, 3, 5)
\]

To two benefit indices \( B, S \), standardization may be represented by equation (2):

\[
Z_{ij} = \frac{y_{ij}^{\text{max}} - y_{ij}^{\text{min}}}{y_{j}^{\text{max}} - y_{j}^{\text{min}}} (i = 1, 2, \ldots, m; j = 2, 4)
\]

where, \( y_{j}^{\text{max}} = \max\{y_{1j}, y_{2j}, y_{3j}, y_{5j}\}, y_{j}^{\text{min}} = \min\{y_{1j}, y_{2j}, y_{3j}, y_{5j}\} \).

3.2.3. Weight of sub-objective

The weight of every sub-objective is composed of weight vector \( W = (w_1, w_2, w_3, w_4, w_5)^T \) and satisfied that the summation is equal to one, and \( w_j \geq 0, j = 1, 2, 3, 4, 5 \). In standardization matrix \( Z = (z_{ij})_{m \times 5} \), to a sub-objective \( G_j \), the difference between the candidate parking lot \( P_i \) and other candidate parking lots may be defined as:

\[
v_j(w) = \sum_{i=1}^{m} |z_{ij} - z_{ij}^*| w_j
\]
It is supposed that \( v_j(w) \) represents the sum varying dispersions among the candidate parking lots:

\[
v_j(w) = \sum_{i=1}^{m} v_i(w) = \sum_{i=1}^{m} \sum_{k=1}^{n} (z_{ij} - z_{kj}) w_j, \tag{4}
\]

where, \( i = 1, 2, \ldots, m; j = 1, 2, 3, 4, 5 \). The choice of weighting vector \( w_j \) should make the sum varying dispersions among the candidate parking lots maximization. So the objective function may be represented by the following function:

\[
\text{max } F(w) = \sum_{j=1}^{5} v_j(w) = \sum_{j=1}^{5} \sum_{i=1}^{m} \sum_{k=1}^{n} (z_{ij} - z_{kj}) w_j \tag{5}
\]

The solution of the function (5) could be represented by the following:

\[
w_j^* = \frac{\sum_{i=1}^{m} \sum_{k=1}^{n} (z_{ij} - z_{kj})}{\sqrt{\sum_{j=1}^{5} \left( \sum_{i=1}^{m} \sum_{k=1}^{n} (z_{ij} - z_{kj}) \right)^2}}, j = 1, 2, 3, 4, 5 \tag{6}
\]

Traditional weight vector is satisfied with normalization constraints and not satisfied with units constraints. So unit weight vector \( w_j^* \) should be dealt with normalization:

\[
\bar{w}_j^* = w_j^* \left/ \sum_{j=1}^{5} w_j^* , j = 1, 2, 3, 4, 5 \right.
\]

### 3.2.4. Decision ordering

Through the steps mentioned above, objective weight vector under fuzzy preference is got: \( \mathbf{W} = (w_1^*, w_2^*, w_3^*, w_4^*, w_5^*)^T \). And then, weighted normalization decision matrix may be represented by the following function:

\[
\mathbf{C} = \begin{bmatrix}
w_1^* z_{11} & w_2^* z_{12} & w_3^* z_{13} & w_4^* z_{14} & w_5^* z_{15} \\
w_1^* z_{21} & w_2^* z_{22} & w_3^* z_{23} & w_4^* z_{24} & w_5^* z_{25} \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
w_1^* z_{m1} & w_2^* z_{m2} & w_3^* z_{m3} & w_4^* z_{m4} & w_5^* z_{m5} 
\end{bmatrix}
\]

According simple additive weighting method, the following function represents the multi-objective comprehensive evaluation value: \( D(w) = \sum_{j=1}^{5} z_{ij} w_j^*, i = 1, 2, \ldots, m \). The bigger the multi-objective comprehensive evaluation value, the better the decision-making project.
4. Evaluation of two-phased parking choice method

4.1. Advantages

The proposed method considered personal attributive characteristic is an efficient way to serve for pretrip parking guidance system. In the first phase, considering factors such as walking distance, parking feasibility, parking cost and so on, candidate parking lots can be selected. Through the method, searching area is shrunken, solving efficiency is improved and drivers’ preferences are taken into account. In the second phase, multi-object decision making model based on fuzzy preference is put forward to rank candidate parking lots selected in the first phase.

Two-phased parking choice method resolves vehicles parking and traveling in a journey at the same time. Not only parking choice is accomplished, but also the drive route is considered. Through the method, drivers’ time for searching parking spaces is reduced, blindness manufacture in a travel is avoided effectively and trip efficiency is increased.

4.2. Issues need to further research

In the paper, shortest travel time is chosen as the optimization criterion. But different users have different preferences, so the parking choice system should develop different criteria, such as shortest distance, lowest jam degree, least travel cost and so on. Not only parking choice but also drive route choice choose user optimal as object. It is feasible in the case of numbered users. When the amount of user is large, system optimal is also needed to be considered. The paper is mainly to build two-phased parking choice method for before-trip parking guidance system. Further research about prediction method of real-time parking space, optimal path problem of time-dependent and stochastic traffic network is necessary.

References


A STOCHASTIC USER EQUILIBRIUM (SUE) ALGORITHM BASED ON ANT COLONY OPTIMISATION (ACO)

Luca D’ACIERNO*, Bruno MONTELLA*, Fortuna DE LUCIA*

Abstract. In this paper we propose a Stochastic User Equilibrium (SUE) algorithm that can be adopted as a simulation model when cost functions depend on the number of vehicles using network elements. Indeed, analyses of real dimension networks need simulation algorithms that allow network conditions and performances to be rapidly determined. Hence, we developed an MSA (Method of Successive Averages) algorithm based on the Ant Colony Optimisation paradigm that allows transportation systems to be simulated in less time but with the same accuracy compared with traditional MSA algorithms.

1. Introduction

In design problems or in real-time management of transportation systems, it is necessary to have a simulation model that allows network performances and features to be defined for each alternative project or each management strategy. In analyses of real dimension networks, both in urban and rural contexts, it is important that simulation models allow solutions to be obtained swiftly such that it is possible to explore a large number of alternative projects or simulate beforehand consequences of a strategy in terms of future (minutes or hours) network conditions. Most simulation algorithms used in the case of steady-state conditions (assumption of inter-period and intra-period stationarity) are based on the calculation of a sequence of network loading (assignment with a fixed-point approach, as proposed by [5]).

In this paper, we verify the possibility of developing a meta-heuristic algorithm that allows network flows to be calculated more quickly than by using traditional algorithms proposed in the literature. In particular, we steered our research into ant-based algorithms. Indeed, these algorithms, developed about a decade ago (the first papers were by [8], [9], [14]) and based on the food source search of ant colonies, have in many cases shown their efficiency in terms of calculation time, such as in travelling salesman problems ([4], [8], [16]), in quadratic assignment problems ([20], [24], [29]), in job/shop scheduling ([2], [10], [5]).
In the case of transportation systems, an algorithm based on the Ant Colony approach for solving a network design problem, consisting of choosing among a set of alternative projects, was proposed by [27]. Moreover, [25] proposed an assignment equilibrium algorithm based on ACO that tends to load mainly minimum cost paths (deterministic approach). The same authors also compared its performance with those of Frank and Wolfe’s algorithm ([18]), showing the efficiency of the ACO-based approach in simulation problems.

An important aspect concerning the development of solution algorithms is the theoretical proof of convergence. Nevertheless, initially most of the literature analysed convergence properties only from a numerical point of view. However, in recent papers (such as [7], [15], [21], [22]) convergence is stated for some classes of ACO-based algorithms.

In this context, the aim of the paper is to develop an ACO-based algorithm to solve the Stochastic User Equilibrium (SUE) problem and prove its convergence and efficiency. The paper is organised as follows: section 2 shows the proposed assignment algorithm, and the first results are reported in section 3; section 4 concludes and comments on prospects for future research.

2. The proposed assignment algorithm

In this paper, we proposed an ACO solution algorithm for the well-known fixed-point problem proposed by [5]:

\[ f^* = AP\left(-A^T c(f^*) - C^{N^A}\right) \]  

(1)

where \( f^* \) is the link flow vector that yields costs that generate a network loading vector equal to itself, \( A \) is the link-path incidence matrix, \( P \) is the path choice probability matrix (known as the path choice map), \( c \) is the vector of link cost functions, \( C^{N^A} \) is the vector of non-additive path costs, and \( d \) is the vector of demand flows.

In the literature, two papers ([5] and [6]) state that the fixed-point problem, expressed by (1), has at least one solution if: choice probability functions, \( P(-C^t) \), are continuous, link cost functions, \( c(f) \), are continuous and each OD pair is connected (i.e. \( I_{od} \neq \emptyset \) \( \forall od \)).

Moreover [5] and [6] state that the above fixed-point problem has at most one solution if: route choice models are expressed by strictly increasing functions with respect to systematic utilities and cost functions are expressed by monotone non-decreasing functions with respect to link flows.

It may be stated that these conditions are always satisfied by almost all functions proposed in the literature and, therefore, the fixed-point solution exists and is unique.

An extension of Blum’s theorem ([1]) was proposed by [5] where it was shown that we may develop two solution algorithms based respectively on the following sequences:

\[ f^{n+1} = f^n + \frac{1}{t_n} \left( f^n c(f^n) - f^n \right) \in S_r \quad \text{with} \quad f^n \in S_r \]  

(2)

\[ c^{n+1} = c^n + \frac{1}{t_n} \left( c^n f(c^n) - c^n \right) \in S_r \quad \text{with} \quad c^n = c(f^n) \in S_r \quad \text{and} \quad f^n \in S_r \]  

(3)
where \( f(\cdot) \) is the network loading function that provides link flows as a function of path costs, and \( \mathcal{F}_r[\mathcal{S}_r] \) is the feasibility set of vectors \( f(e) \).

In order to prove the convergence of algorithms, assuming that solution existence and uniqueness conditions hold, it is necessary to verify that link cost functions have a symmetric continuous Jacobian \( \text{Jac}[c(f)] \) over set \( \mathcal{F}_r \), for algorithm expressed by (2), and choice map functions \( P(\cdot) \) are additive and continuous with continuous first derivative for algorithm expressed by (3).

In this paper, we propose an ACO solution algorithm for the fixed-point problem (1) based on the following assumptions:

– for each \( od \) pair there is an ant colony with its nest (centroid \( o \)) and its food source (centroid node \( d \)). Since every colony has a distinctive kind of pheromone, ants can recognise only paths utilised by the same colony. This hypothesis was formerly introduced by [25];

– the initial intensity of the pheromone trail on each link \( l \), associated to ant colony \( od \), indicated as \( \tau_{od,l}^0 \), is a function of path costs, that is:

\[
\tau_{od,l}^0 = \sum_{k \in I_{od}} T_{od,k}^0
\]

with:

\[
T_{od,k}^0 = \begin{cases} 
\exp(-C_k^0/\theta) & \text{if } k \in I_{od} \\
0 & \text{if } k \notin I_{od} 
\end{cases}
\]

\[
C_k^0 = c_i(f^*)
\]

where \( f^* \) is the initial vector of link flow that belongs to feasibility set \( \mathcal{F}_r \) and \( C_k \) is the cost of \( k \)-th path.

– the probability of choosing link \( l \), with \( l = (i, j) \), at diversion node \( i \) (known in the ACO literature as transition probability), at iteration \( t \), can be expressed as:

\[
p'[l | i] = \frac{\tau_{od,l}^t}{\sum_{l \in FS(i)} \tau_{od,l}^t}
\]

where \( FS(i) \) is the set of links belonging to the forward star of node \( i \). In this case, the visibility term (generally indicated as \( \eta_{od} \)) is equal to 1;

– the “ant” (or “vehicle”) flow on the path \( k \) is equal to:

\[
F_{od,k} = d_{od} \prod_{l \in k} p'[l | i]
\]

which is equal to the flow expression proposed by [25];

– the updating of the pheromone trail can be expressed as:

\[
\tau_{od,l}^t = (1-\rho)\tau_{od,l}^{t-1} + \rho \Delta \tau_{od,l}^t
\]

where evaporation coefficient \( \rho \) is variable and equal to \( 1/t \). The variability of this coefficient was introduced by [23], even though the expression was different;

– each ant is provided with a memory that stores the sequences of used links. This property allows ants to update the pheromone trail at each iteration only if they are the first to arrive, after reaching the food source, at their nest;

– the increase in the pheromone trail is based on a global approach, that is all links are simultaneously updated. This assumption, in combination with the ant memory property, can be expressed by a function of path costs, that is:
\[ \Delta T'_{ad} = \sum_{k,h} \Delta T'_{ad, k} \]  

with:

\[ \Delta T'_{ad, k} = \begin{cases} 
\exp\left(-\frac{C_i}{\theta}\right) & \text{if } k \in I_{ad} \\
0 & \text{if } k \notin I_{ad} 
\end{cases} \]  

The usefulness of a global approach in transportation problems was highlighted by [27].

With the above hypotheses, it may be stated that the application of the proposed ACO algorithm in the case of a transportation network is equivalent to the application of an MSA algorithm where the successive averages are applied to weights of Dial’s algorithm ([11]), that is:

\[ \Delta \sum \in = (10) \]

Moreover, with the use of the extension of Blum’s theorem ([1]) proposed by [5], convergence of the proposed ACO-based MSA algorithm may be stated theoretically, assuming that existence and uniqueness conditions hold, if \( \text{Jac}[e(\tau)] \) is symmetric and continuous. Indeed, this condition satisfies all hypotheses of Blum’s theorem in the case of the fixed-point problem \( \tau = \tau(f(\tau)) \) where \( \tau \) is a vector whose generic element is the pheromone trail \( \tau_{ad} \). The proof of convergence in this case is similar to that proposed by [5], hence for brevity it is not reported in this paper. However, sufficient conditions to verify the \( \text{Jac}[e(\tau)] \) hypothesis are that link cost functions are separable (i.e. \( c_i = c(f_i) \)) and have a symmetric and continuous Jacobian \( \text{Jac}[e(f)] \) over set \( S_P \), and choice map functions \( P(\cdot) \), are additive and continuous with the continuous first derivative. Indeed, since \( \text{Jac}[e(\tau)] = \text{Jac}[e(f)\text{Jac}[f(\tau)] \) the condition on \( \text{Jac}[e(\tau)] \) is verified. Moreover, the above conditions are generally satisfied by almost all functions proposed in the literature and therefore convergence of the proposed algorithm may be postulated.

3. First results

In order to verify the efficiency of the proposed MSA algorithm based on Ant Colony Optimisation, it was applied to simulate traffic conditions in the case of two Italian real dimension networks: the network of Salerno (a city of about 140,000 inhabitants) and the network of Naples (a city of about 1,000,000 inhabitants). Table (1) shows features of the analysed networks, whereas tables 2 and 3 indicate algorithm performances. In both networks the proposed algorithm (indicated as MSA-ANT) is shown to provide the same solution as the traditional algorithm (algorithm thresholds being equal to 0.1% and 1.0% respectively in the case of Salerno and Naples) in lower calculation times. Instead, these values become 20.0% and 32.1% with respect to the case of algorithm MSA-CA.
In terms of number of iteration, the proposed algorithm requires 90.7% and 84.7% less with respect the algorithm MSA-FA; these values become 28.6% and 35.7% in the case of algorithm MSA-CA. Finally, if we accept that the convergence of algorithm is achieved when only 90% of links satisfy termination test than the reductions in number of iterations would be 75.0% and 63.6% in the case of algorithm MSA-FA, and 40.0% and 0.0% in the case of algorithm MSA-CA.

All numerical tests were performed with a PC Pentium IV 3.0 GHz.

<table>
<thead>
<tr>
<th>City name</th>
<th>Number of links</th>
<th>Number of nodes</th>
<th>Number of centroid nodes</th>
<th>Number of OD pairs</th>
<th>Number of peak-hour trips</th>
</tr>
</thead>
<tbody>
<tr>
<td>Salerno</td>
<td>1,133</td>
<td>529</td>
<td>62</td>
<td>3,844</td>
<td>21,176</td>
</tr>
<tr>
<td>Naples</td>
<td>5,750</td>
<td>3,075</td>
<td>167</td>
<td>26,916</td>
<td>118,764</td>
</tr>
</tbody>
</table>

Table 1. Network features.

<table>
<thead>
<tr>
<th>Algorithm name</th>
<th>Number of interactions</th>
<th>Convergence &gt;90.0 %</th>
<th>Calculation time [s]</th>
<th>Solution error</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSA-FA</td>
<td>54</td>
<td>12</td>
<td>29</td>
<td>Reference</td>
</tr>
<tr>
<td>MSA-CA</td>
<td>7</td>
<td>5</td>
<td>5</td>
<td>0.004 %</td>
</tr>
<tr>
<td>MSA-ANT</td>
<td>5</td>
<td>3</td>
<td>4</td>
<td>0.004 %</td>
</tr>
</tbody>
</table>

Table 2. Algorithm performance on the Salerno network.

<table>
<thead>
<tr>
<th>Algorithm name</th>
<th>Number of interactions</th>
<th>Convergence &gt;90.0 %</th>
<th>Calculation time [min]</th>
<th>Solution error</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSA-FA</td>
<td>59</td>
<td>11</td>
<td>42.15</td>
<td>Reference</td>
</tr>
<tr>
<td>MSA-CA</td>
<td>14</td>
<td>4</td>
<td>11.26</td>
<td>0.013 %</td>
</tr>
<tr>
<td>MSA-ANT</td>
<td>9</td>
<td>4</td>
<td>7.65</td>
<td>0.006 %</td>
</tr>
</tbody>
</table>

Table 3. Algorithm performance on the Naples network.

4. Conclusions and research prospects

In this paper we proposed an MSA assignment algorithm based on Ant Colony Optimisation and stated the perfect equivalence in terms of path choice behaviours between artificial ants (with the proposed approach) and road users (simulated with the traditional algorithms of traffic assignment). Moreover we stated theoretically the convergence of the proposed MSA-ANT algorithm by means of the extension of Blum’s theorem ([1]) proposed by [5], and numerically its efficiency in terms of calculation time with respect to traditional MSA algorithms.

Hence, the proposed algorithm can be considered an extension of ACO-based assignment algorithms in the case of stochastic assumption on path choice model (the previous paper [25] was based on a deterministic approach). Moreover, theoretical proof on convergence overcomes limitations resulting from the use of numerical proof in [25]. Finally, the new MSA-ANT could be utilised for developing an ACO-based algorithm also in the lower level of paper [27] in order to speed up solution search.
In terms of future research, we propose an extension of the proposed algorithm in the case of more complex path choice models (such as C-Logit and Probit) and in the case of preventive-adaptive choice behaviour that is typical of mass transit system users (i.e. the hyper-path choice approach proposed by [26]).

Finally, we advocate using the proposed algorithm as a simulation model in network design problems or in real-time management of transportation systems in order to highlight the advantages of the ACO approach.

References


AN ACTIVITY-BASED SYSTEM OF MODELS USING RP AND SP METHODOLOGIES

Gabriella MAZZULLA

Abstract. In this paper an original formulation of a system of random utility behavioural models is introduced. User decisional procedure is simulated according to a sequential approach, or rather through a set of linked sub-models that reproduce the different choice dimensions for consecutive stages. An activity-based approach has been adopted. In the proposed system, some user choice dimensions and, specifically, the transport mode choice, is modelled using both RP and SP data.

1. Theoretical framework

Over the last 30 years, considerable advances in modelling travel demand have been made and, particularly, in discrete choice analysis. The trip-based approach has been widely studied and applied. Afterwards, the tour-based approach has allowed some complexities to be addressed, such as trip-chaining and the interrelation between travel from home to one or more activity locations and back home again. The systems of more advanced models simulate the travel-pattern and analyze every tour in one or more day, taking into account the existing conditionings among the tours that people make in a day or a week; in some cases, also the tours realized by other family members are taken into account.

As is well-known, the mobility demand is not an end in itself, but is derived from the need to conduct several activities in different places. Activities and trips have a cause and effect relationship. Therefore, since the early seventies and more markedly about ten years later, many authors, instead of the traditional trip-based approach, have preferred an innovative approach, in the literature named activity-based, because it is based on activities rather than on trips analysis. For this reason, activity-based models can be defined as participation activity models. According to the activity-based approach, the mobility demand is simulated taking account of the relationship between activities and trips and the spatio-temporal constraints in which people make activity and travel decisions.

1 Department of Territorial Planning, University of Calabria, via P. Bucci, cubo 46/B, 87036 Rende (CS), g.mazzulla@unical.it
Adler and Ben-Akiva are the first authors who introduced the activity duration and activity scheduling concepts [1]. Travel demand analysis according to the activity-based approach was analyzed by Pas [7]. For a more detailed discussion of activity-based travel theory one should refer to Axhausen and Garling [2], Damm [3], Ettema and Timmermans [4], Golob and Golob [5], Kitamura [6].

Even if the activity-based approach has widely been discussed in the literature, there are still few proposed applications; besides, only recently the activity-based travel theory has been combined with the conjoint use of the Revealed and Stated Preferences methodologies [8]. The conjoint use of such data allows the analysis and the simulation of both actual and future consumer behaviour in real scenarios, but also in hypothetical scenarios, with the aim of considering, for example, non-existing transport mode among the choice alternatives.

2. Experimental context

An experimental survey was realized on the campus of the University of Calabria, situated in the urban area of Cosenza (in the South of Italy). The campus is attended by 32,000 students and 2,000 members of staff approximately (November 2004). The survey, realized in the autumn of 2004, involved a sample of 1,477 students, with a sampling rate equal to 4.5% approximately.

In the survey the students described activity and trip sequences in the tours made in a day. The tours are the trip sequences with first origin and last destination at home (home-based tours), and with one or more sojourns on the campus. Each activity is characterized by typology and duration, and each trip is defined by origin, destination, duration and transport mode used. In total, the students made 1,631 tours with at least one stop in the university area; 1,323 students made only one tour in a day; 154 students made two tours. As expected, the tours were realized primarily for carrying out didactic activity; in 66% of the cases attending a lesson, in 14% of the cases conducting other study activities and in 13% of the cases effecting a different didactic activity from the first ones. However, other activities, like the handling of personal business or recreation/relational activities, were undertaken in the tours but were not considered as primary activities. About 64% of the tours were trip-tour (i.e. with a single outward and one return trip); the remaining 36% were tours with three or more trips (trip-chain), in which two or more activities were combined.

Additionally, the survey was planned to collect RP and SP data on the same sample. The RP survey was done in order to collect some students’ socio-economic characteristics and information regarding transport mode actually used to reach the university campus.

The SP experiment was made in order to estimate a strategy for increasing the use of collective transport that connects the urban area with the campus; the strategy consists in conjugating policies for the improvement of the collective transport services with demand management policies, i.e. car-park pricing. A new transport system, which connects the urban area with the campus, was proposed; this system is characterized by a high frequency and a low travel time. At the same time, park pricing policies in the university area were adopted. In the SP experiment, the users expressed their degree of preference (according to a semantic scale from 1 to 5) on 7 hypothetical choice scenarios; each scenario includes both car and bus alternatives. The car alternative is characterized by the parking cost
attribute, varying from the actual level (free) to the intermediate and high level; the bus alternative is characterized by frequency (low and high) and travel time attributes (equal or lower than the actual time).

3. The proposed system of models

In this paper, a system of models has been proposed. These models have the purpose to simulate the sequence of activities and trips made by the university students, according to the activity-based approach. Specifically, the home-based tours, realized by the students for making one or more activities on the campus, were analyzed. Some typical activities are: attending a lessons or other didactic activities, handling personal business, meal, recreation and/or relational activities. The system of models simulates the sequence of activities in a trip-tour or trip-chain; in the first case the students conduct a single activity in the university area; in the second one the students undertake several activities in the same tour, some inside the campus and other outside the campus, in the urban area of Cosenza. No hypotheses were made on the connection among the tours realized by the same student in a day, but only the “first tour” was analyzed and simulated.

Each simulated tour is characterized by a predominant activity, which produces the need to make the tour, named primary activity PA. For defining the primary activity of each tour, a hierarchy of activities was formalized, as a function of the spatial and temporal constraints that each activity has on the organization of the same tour. Specifically, the place in which the activity is carried out is taken into account, the activity type and the chronological sequence. It is supposed that the primary activity in the tour is made inside the campus, considering that each activity effected outside the university area is secondary. Among the activities carried out inside the campus, a hierarchy of the activity was defined, according to the following order: attending a lesson; other didactic or study activities; meal; handling personal business; recreation and/or relational activity. If two activities of the same hierarchical level are made in a tour, the primary activity is the first in chronological order.

The proposed system of models simulates the user decisional procedure in six choice dimensions, each included in a different sub-model, whose sequence is introduced in the flow chart shown in figure 1. The higher hierarchical level is represented by the activity program choice model AP. The next hierarchical level is represented by the models related to the primary activity, and specifically: choice of the primary activity PA, choice of the primary destination PD, in which carrying out this activity, and choice of the transport mode MPD used to reach this destination from home. Equally, a set of models relates to the possible secondary activities SA in the tour, and specifically: choice of the secondary activity SA, choice of the secondary destination SD, in which effecting this activity, and choice of the transport mode MSD used for reaching this destination. In some cases, the choice dimensions are not only conditioned according to the order of the introduced sequence, but the choice dimensions at a lower level condition those at a higher level. This is shown, in figure 1, by a discontinuous line of connecting arrows among the different models.
Figure 1. General structure of system of models.
In analytical terms, this reciprocal conditioning is represented by a specific variable inserted in the utility function of the alternatives of the model at a higher level, which considers the choices also made by users at a lower level. Specifically, this link is inserted between the models for choosing the type of activity and where the activity takes place, and the destination choice model, in reference to both the primary and secondary activities.

The structure of the system of models indicates the sequence with which the choices on each dimension are made by users and their reciprocal influence. The order of the sequence with which the sub-models are introduced could be different from the one described. For instance, the choice related to the activity program could be made after the choice of the primary activity in the tour, because the user chooses to make a possible secondary activity as a function of the spatio-temporal constraints related to the primary activity. The order of the sequence is usually verified in the model calibration phase, verifying the different hypotheses by statistical tests performed on the single parameters and sub-models. Nevertheless, in this case, the verification was not effected.

The sequence of the sub-models introduced in figure 1 can be expressed, in analytical terms, by the following relation:

\[
p(\text{AP, PA, PD, MPD, SA, SD, MSD/T}) = p(\text{AP/T}) \times p(\text{PA/T, AP}) \times p(\text{PD/T, AP, PA}) \times p(\text{MPD/T, AP, PA, PD}) \times p(\text{SA/T, AP, PA, PD, MPD}) \times p(\text{SD/T, AP, PA, PD, MPD, SA}) \times p(\text{MSD/T, AP, PA, PD, MPD, SA, SD})
\]

The symbols in formula (1) have the following meaning:
- \( p(\text{AP, PA, PD, MPD, SA, SD, MSD/T}) \) represents the probability that the student, considered that he has decided to go from home to the campus and therefore to make the tour T, after he has chosen the activities program AP, undertakes the primary activity PA, in the primary destination PD, reached with the transport mode MPD, with the secondary activity SA, in the secondary destination SD, reached with the transport mode MSD;
- \( p(\text{AP/T}) \) represents the probability, calculated through the activity program choice model, that the student chooses to combine or not the primary activity in the tour with one or more secondary activities, given the tour T;
- \( p(\text{PA/T, AP}) \) represents the probability that the student, after he has chosen the activity program AP, undertakes the primary activity PA in the campus, given the tour T;
- \( p(\text{PD/T, AP, PA}) \) represents the probability that the student, after he has chosen the activity program AP, undertakes the primary activity PA in the primary destination PD inside to the campus, given the tour T;
- \( p(\text{MPD/T, AP, PA, PD}) \) represents the probability that the student, after he has chosen the activity program AP, uses the transport mode MPD to reach the primary destination PD, in which makes the primary activity PA, given the tour T;
- \( p(\text{SA/T, AP, PA, PD, MPD}) \) represents the probability that the student, after he has chosen the activity program AP, undertakes the secondary activity SA, combined with the
primary activity PA in the primary destination PD, reached with the transport mode MPD, given the tour T;

- $p(SD/T, AP, PA, PD, MPD, SA)$ represents the probability that the student, after he has chosen the activity program AP, undertakes the secondary activity SA, in the secondary destination SD, inside or outside to the campus, combined with the primary activity PA in the primary destination PD, reached with the transport mode MPD, given the tour T;

- $p(MSD/T, AP, PA, PD, MPD, SA, SD)$ represents the probability that the student, after he has chosen the activity program AP, undertakes the secondary activity SA, in the secondary destination SD, reached with the transport mode MSD, combined with the primary activity PA in the primary destination PD, reached with the transport mode MPD, given the tour T.

In the proposed system of models the generative phase of the tour was not simulated, because the survey was conducted "at the destination", i.e. the student was interviewed in the place in which he had undertaken his activity.

The proposed system is composed of random utility behavioural models, with a multinomial or hierarchical Logit structure, according to the degree of similarity perceived by the user among the different alternatives. It was calibrated using the data of a sample of University of Calabria students. The results obtained are satisfactory and the statistical tests show that the experimental data were well replied. Additionally, the attribute coefficients of the considered systematic utilities show a good statistical significance.

4. Conclusions

In this paper an original formulation of a system of random utility behavioural models has been introduced. The system of models allows student mobility simulation in a university campus. User decisional procedure has been simulated according to a sequential approach, or rather through a set of linked sub-models that reproduce the different choice dimensions for consecutive stages. An activity-based approach has been adopted. The proposed models are multinomial or hierarchical Logit models.

Additionally, the conjoint RP and SP techniques have been adopted for investigating on the user choice behaviour in the hypothesis that an innovative transport system is realized to access the university campus. Indeed, in the proposed system of models some user choice dimensions, and specifically the transport mode choice, is modelled using both RP and SP data. The conjoint use of this data allows the analysis and the simulation of both actual and future consumer behaviour in real scenarios, but also in hypothetical scenarios, in order to consider non-existing transport mode among the choice alternatives.

The system of models was calibrated by a survey made for an Italian university campus, attended by around 30,000 students. Model calibration has allowed investigation of the general structure of the proposed system and of the systematic utility functions of the choice alternatives and, specifically, it has allowed the choice of the best variables to be taken into account for the analysis of student travel behaviour.

Although the proposed system of models is not relevant to general planning, the model structure is very realistic for student mobility simulation.
In some cases, the individual models proposed in the system structure have a simple formulation, like the destination choice models; however, more complex individual models can replace existing simpler models if the need or desire arises.

More significant results can be obtained by using a wider statistical sample. These results can be also used in similar territorial contexts, with the aim of supplying better, more efficient transportation services to university students.

Acknowledgements

The author is grateful to Rosanna Guzzo for her work in the preliminary phase of the research as PhD candidate in the University of Calabria (Italy).

References


Abstract. For testing and analyzing models and application, researches spend an enormous amount of time to gather suitable data and to convert this data to a format that can be processed by their developed models. This work slows down the research and leads to tons of unused data in archives that is inaccessible for other researchers in the same field. This paper presents a web based approach for a data source research platform to overcome this drawback and to open the possibilities for further collaborations in the European or even world-wide collaboration of transportation research.

1. Introduction

Gathering real life data is a time consuming job in research. A lot of data is measured and stored in several places and different formats. While a lot of it is not used, other institutions gather similar data on different locations. In this way a lot of money and time is spend unnecessary. During the developing of the microscopic online simulator MiOS (Miska, Muller, van Zuylen, 2004) at the Delft University of Technology, the issue of gathering real world input (online and offline) was rising throughout new challenges.

Online measurements had to be gathered from different servers and different structure for the input to the simulation and for the driving behavior model (Miska, Muller, van Zuylen, 2006) offline trajectory data from different sources was used. This led to a framework to gather information of different structure and sources for the processing.

As a logical extension of this framework a web based data source platform is under development to provide researchers with an easy access to the data they need in a format they can use. This paper describes the structure of this platform and shows the possibilities for research and practice to develop new and assess existing models with various sets of networks and measurement data.
2. Structure of the Platform

2.1. Overall structure

The basic structure of the platform is defined in three different layers. The bottom layer is a set of databases connected to the platform and an extension of it with bots (software modules, searching the web) that includes the gathered information. On top of that, a second layer is used for data fusion and converting the data in a requested format. The third layer is the user front end, which is a web based application to inquire needed information. Figure 1 illustrates the different layers.

![Figure 1. Structure of the data platform with three layers](image)

2.2. Data Layer

The data layer consists of a set of hard connected databases and bots to access remote databases known to the platform. While the connected databases have a known structure, database queries can be performed directly. For remote databases or web pages, from which data can be obtained, bots are taking over this part. A bot is a software program, aware of the network and searching predefined web locations for certain data. Better known than bots are agents, which are bots with additional artificial intelligence to deal with different situations, inputs and are able to make decisions. These kinds of abilities are at this stage not needed and so the simple form of a bot has been chosen. The data stored in this layer is consistent of network descriptions, measurements information and can be extended in various ways.
2.3. Process Layer

The process layer has two major tasks: first the fusion of data from the databases and second, the reformatting of the data to a required form. Data fusion is getting more and more important recently. Traffic control centers receive measurements from various sources and have to process them to a common picture of the network state. Online models have to assemble data from different sources as an input for their simulations and so there is a growing need for fused data in research.

If volume and speed data for a certain network stretch is required, it should be possible to obtain this data from one single source, even though the measurements are coming from inductive loop detectors, cameras, or other devices. The data fusion module at this early stage simply assembles the measurement for the defined network part and results in a dataset of all measurements with location information. It is planed to add to this layer various tools for data mining to allow researchers to clean the data in advance. Since this is a manipulation of data, it will only be used at user demand.

The reformatting of the data is a simpler task. When the data is found, the existing variables are known and the user can input a required format for the output. If the information is inconsistent between the data sources it is possible to request the common information or to fill missing data with void values.

This allows users to check in an easy and simple way if the required data for their research is available already or if further gathering is required.

2.4. Front end

The front end is a web based application, well known from other internet portal in the web. It gives the possibility to receive common information and after registering to access the available data by requests. Registration will be free of charge for research purposes and the login data is only used for statistical purposes. If the user gives the permission, it will be published what institutions work with which kind of data, which might lead to further collaboration opportunities.

Since various data is not freely accessible by restrictions of companies, governments or research sponsors, the possibility is given to add charges for certain information. For commercial use their might be a sign in fee in the future. The user will get after the request a complete list of data, including the information for costs of the download. Even though the data might not be available for free, the costs are expected to be much lower than to perform measurements from scratch. Uploaded data will always be accompanied by a detailed description of the methods used to obtain the measurement, so that the confidence in the quality of the data can be given.

Next to these mostly for researchers interesting options, the platform also includes the possibilities for practitioners to test different simulation models before making the decision to buy them. It is well known that different simulation models have their strength in certain applications. Without looking into detail, nobody can tell about the quality without testing. Therefore, practitioners can download complete network structures and input data for a simulation as well as the real world results. This allows an easier assessment of different models for real applications and purposes they are going to be used for.
3. Realization

In contrary to the Next Generation Simulation NGSIM project (http://ngsim.camsys.com), the proposed platform is not a simple download page for data sets. A Java based application allows the user to search through the data and its description and to define what to download. This results into a query to the processing layer. Underlying software tools generate the necessary queries to the databases and triggers, if applicable, the bots to collect data. The data fusion in this stage assembles the data according to the wishes of the user and creates a standardized data file. This file is used by the conversion module and returns it to the front end Java application to be send then to the user.

The procedure is tested with data from the Regiolab Delft. The Regiolab Delft is a combined research effort of the Delft University of Technology, the research school TRAIL, the Dutch Ministry of Transport, the province of South Holland, the municipality of Delft and the traffic industry. A variety of existing and newly developed traffic detection systems enable Regiolab Delft to monitor traffic (flow rates, speed, density, travel time) and to estimate the actual origins and destinations of flows. In addition, in the near future Regiolab Delft will also receive information about the actual status of the freeway control systems (ramp metering, speed control, lane use), weather conditions, and so on. For research goals, a data storage and analysis system is developed supplying different users with dedicated information.

4. Conclusion and Further Research

Even though the presented data source platform is not a new scientific contribution to the transportation world, it clearly closes a gap between data measurement and data usage. There is an added value for researchers as well as practitioners which will save a lot of time and therewith money in project budgets. The platform is operational from September 2006 and can be reached under the web address: http://www.trafficdata.info.

References


DYNAMIC SELECTION OF FUZZY SIGNAL CONTROLLERS

Maria Alice Prudêncio JACQUES¹
Michelle ANDRADE²

Abstract. Previous studies have proven that modifications to fuzzy set definitions exert a significant impact on fuzzy signal controller response. Further, the literature shows some procedures aimed at adjusting fuzzy set parameters in response to prevailing traffic conditions. Among other techniques, neural networks have been used to provide this adjustment online. This approach, however, presents some limitations and so indicates the need for further investigations. In this sense, this paper presents a procedure, called dynamic selection of fuzzy signal controllers. The procedure seeks to allow for the use of different fuzzy control strategies in response to prevailing traffic conditions. It is based upon off-line studies regarding combinations of fuzzy sets so as to better traffic control under different traffic volume levels and operation characteristics. The evaluation of the procedure proposed has shown promising results in terms of reducing queue and cycle length at an isolated intersection.

1. Introduction

Fuzzy control techniques have been adopted to improve signal controllers’ responsiveness to traffic volume variations at signalized intersection approaches. That is, they aim to make these controllers more flexible and, therefore, more efficient. However, recent studies have drawn attention to the fact that in many cases, one of the more important elements in fuzzy signal controllers, their related input and output fuzzy sets, should be adjusted to improve controller response to certain modifications in traffic situations. In other words, it has been recognized that the conceptual meaning of one or more particular fuzzy sets can be modified depending on general traffic conditions. For instance, the concept of “small queue” changes from peak to non-peak periods. In addition, other studies have indicated that small modifications to the partition of a given variable’s universe of discourse may have a significant impact on controller response and, therefore, on the performance of the

¹ University of Brasília, Brazil, mapj@unb.br
² University of Brasília, Brazil, maccivil@gmail.com
traffic controlled [3]. These findings justify investigations geared towards conducting online adjustments to fuzzy signal controller parameters.

2. Adaptive Fuzzy Signal Controllers

The literature shows some procedures aimed at adjusting fuzzy set parameters in response to prevailing traffic conditions. Among them, the statistical-adaptive and fuzzy-adaptive methods by [4] can be pointed out. These methods apply adaptive strategies to fuzzy traffic controllers by combining different defuzzification and decision-making criteria. The statistical-adaptive method works on the fuzzy sets belonging to the fuzzy rules’ antecedent variables (arrival and queue), and the fuzzy-adaptive method deals with the fuzzy rules’ consequent variable (extension). In both methods, the value for the adjustment defined is similar for all fuzzy sets related to the same fuzzy variable to be modified. Another procedure developed comprises the use of a neural network structure for providing an adjustable fuzzy control system for traffic control named neurofuzzy system [2]. The modifications provided by this method are more generic than their counterparts in the previously mentioned systems, as they allow for different adjustments for distinct fuzzy sets related to the same fuzzy variable. However, this total freedom for fuzzy set adjustments has proven in some cases to lead to unsuitable results. For example, there are some elements in the discourse universe with no pertinence to any fuzzy sets in the related fuzzy variable. Therefore, because its result has not been very encouraging, it is more sensible to first verify the real potential benefits of a more flexible fuzzy control strategy on traffic performance before going further in the development of such a traffic signal controller type. This is the main objective of this present work.

3. Dynamic Selection of Fuzzy Signal Controllers

With a view to evaluating the advantages of on-line adjustments to fuzzy signal controller parameters, a procedure called dynamic selection of fuzzy signal controllers, developed off-line, is proposed. In this present study the controllers developed differ from each other only in relation to the partition of the universes of discourse for their input variables (queue and arrival). Given that the expected impact of each table on the traffic performance had to be previously evaluated in order to build in the rules for the tables’ choice criteria, other possibilities for alteration to the fuzzy controller’s decision-making unit were not included in this exploratory investigation. However, because the generation of each table is off-line, and independent of each other, there are no limitations as to the possible modifications to be made.

The basic fuzzy signal controller from which the alternative ones were generated is that considered by [3], and is named in the following text as Case 1. Two other controllers were selected from [3] to be a part of the group of fuzzy signal controllers to be considered in the present work: Cases 4 and 7. The three controllers present a general distinguishable pattern of green signal extensions: smaller extensions (Case 7), original extensions (Case 1) and bigger extensions (Case 4).
Each of the fuzzy control strategies, defined by its fuzzy sets, fuzzy rules and defuzzification method, is represented by means of a fuzzy control table, generated for MATLAB software. These fuzzy tables provide extension values, being “extension to the current green time” the output variable for the fuzzy signal controlled considered, related to field measurements of the controller’s input variables, which are “queue” and “arrivals”. Then, based on a predefined traffic performance control variable, the system is able to select the best control table to be used under a given traffic condition. To test this procedure, the queue at the end of the current green at the approach receiving green signal indication (Qia), and the queue at the end of the red for the halted traffic (Qib) were defined as the performance control variables [1]. Both queues are measured in terms of number of stopped vehicles. Furthermore, the test has considered two different approaches. The first took into account the performance control variables values for table selection at the end of each extension of the green time; the second only considers them at the very end of the green time (the new table selected would be effective only from the beginning of the green for the halted traffic). By considering the number of vehicles departing from the stop line during the green signal indication (Da), the following basic criteria were defined.

\[
\begin{align*}
\text{a)} & \quad \text{When the fuzzy controller selection is made at the end of each signal phase} \\
& \quad \text{D}_{a} \geq \text{Q}_{i a}, \quad \text{if} \quad \begin{cases} 
\text{Q}_{f b} \leq 4 & \rightarrow \text{Reduce extensions in the next phase} \\
4 < \text{Q}_{f b} \leq 8 & \rightarrow \text{Keep extension in the next phase at an average level} \\
\text{Q}_{f b} > 8 & \rightarrow \text{Increase extensions in the next phase}
\end{cases} \\
& \quad \text{D}_{a} < \text{Q}_{i a} \rightarrow \text{Reduce extensions in the next phase}
\end{align*}
\]

The above relations imply that if the initial queue (Qia) has been discharged, the controller is selected based on the number of vehicles queued at the conflicting approach (Qib). The selection considers that the greater this number the bigger the extensions to be provided during the next phase. In case the initial queue has not been completely discharged, the controller is selected so as to reduce the green time for the next phase to the greatest extent, bearing in mind the prevailing traffic conditions at the approach to be served.

\[
\begin{align*}
\text{b)} & \quad \text{When the fuzzy controller selection is made at the end of each extension given} \\
& \quad \text{if} \quad \begin{cases} 
\text{Q}_{f b} \leq 4 & \rightarrow \text{Increase extensions in the next phase} \\
4 < \text{Q}_{f b} \leq 8 & \rightarrow \text{Keep extension in the next phase at an average level} \\
\text{Q}_{f b} > 8 & \rightarrow \text{Reduce extensions in the next phase}
\end{cases}
\end{align*}
\]

For this situation, the controller selection criterion takes into account the queue build up pattern at the approach with red signal indication. This criterion results in a more immediate action than that proposed by the first selection strategy. The controller is
selected in order to adjust the extensions given in the current green indication. This specially benefits the discharging queue in cases where the queue build up with opposite traffic is slow.

The values 4 and 8 considered in the previously presented decision-making criteria were defined based on analyses of the limit values for fuzzy set membership functions present in the controllers studied. In other words, for each application of the procedure, there must be a direct relation to the fuzzy sets reflecting control and intersection characteristics.

3.1. Simulations

A simple simulation program was developed to test the two approaches for the procedure proposed. This allows for registering the tables selected, the number of vehicles in queue at the end of each green, and the cycle length at the end of each cycle. A related program was developed to produce the same results for the case in which each one of the tables considered is used continuously (no table selection was allowed).

Initially, simulations were conducted for the three controllers operating isolatedly (only one fuzzy controller table was used at a time). Subsequently, the two approaches for the dynamic selection of fuzzy signal controllers were simulated, where the three fuzzy tables were used interchangeably, according to the respective criteria. The simulations took into account three different initial situations at the controlled approaches regarding the initial queue at the approach receiving the green signal indication (Qia). They are: Qia=0; Qia=10; e Qia=15. Each initial condition was tested for three predefined volume levels [1] considering both fixed and random arrival headways. For all tests 15 (fifteen) cycles were simulated. The simulated intersection is an isolated junction of two one-way streets, one main and one minor street, operating under different traffic volume levels.

3.2. Results and Analyses

Part of the simulation results are presented in Tables 1 and 2. They are related to the following traffic volume under fixed arrival headways: 400 vph at the minor street and 1200 vph at the main street. The analyses performed considered the number of queued vehicles and the green times at the end of the 15th cycle, where these times have been stabilized. For the case of random arrival headways, whose results are not presented here due to space constraints, the analyses took into account the average value over the 15 cycles simulated.

The results showed that the dynamic selection of fuzzy controllers can present better results than the use of a single plan (represented by a single control table), even when a small number of possibilities has been explored. In addition, the table selection at the end of each extension (Fextension) proved to be, in general, more efficient than selection only at the end of the green time (Fphase) for a given intersection approach. For the case of Qia = 0 (empty system at the beginning of the simulation), there is no difference in the control results for all tests made. However, for the other initial situations, the use of dynamic
selection at the end of each extension provided a 25% reduction in the final queue at the main street approach in relation to the better result from the isolated controller operation; the associated green times and cycle length were almost the same.

| Case 1 | Qia = 00 | 7 | 4 | 33.5 | 12 | 52.5 |
| Case 1 | Qia = 10 | 13 | 6 | 50.2 | 19.3 | 76.5 |
| Case 1 | Qia = 15 | 13 | 6 | 50.2 | 19.3 | 76.5 |
| Case 4 | Qia = 00 | 6 | 4 | 33.4 | 12 | 52.4 |
| Case 4 | Qia = 10 | 13 | 6 | 56.7 | 19.9 | 83.6 |
| Case 4 | Qia = 15 | 13 | 6 | 56.7 | 19.9 | 83.6 |
| Case 7 | Qia = 00 | 7 | 4 | 33.3 | 12 | 52.3 |
| Case 7 | Qia = 10 | 12 | 5 | 43.5 | 15 | 65.5 |
| Case 7 | Qia = 15 | 12 | 5 | 43.5 | 15 | 65.5 |

Table 1. Results for the three controllers operating isolatedly.

<table>
<thead>
<tr>
<th>Qia</th>
<th>Stabilized Queue (veh)</th>
<th>Green Time (s)</th>
<th>Cycle (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Main Street Approach</td>
<td>Minor Street Approach</td>
<td>Main Street Approach</td>
</tr>
<tr>
<td>Fphase</td>
<td>Qia = 00</td>
<td>7</td>
<td>4</td>
</tr>
<tr>
<td>Fphase</td>
<td>Qia = 10</td>
<td>12</td>
<td>6</td>
</tr>
<tr>
<td>Fphase</td>
<td>Qia = 15</td>
<td>12</td>
<td>6</td>
</tr>
<tr>
<td>Fextension</td>
<td>Qia = 00</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>Fextension</td>
<td>Qia = 10</td>
<td>9</td>
<td>5</td>
</tr>
<tr>
<td>Fextension</td>
<td>Qia = 15</td>
<td>9</td>
<td>5</td>
</tr>
</tbody>
</table>

Table 2. Results for the dynamic selection of fuzzy signal controllers.

For the minor street approach the dynamic selection results reproduce those produced by the best isolated controller.

4. Conclusions

This study leads to the conclusion that the more flexible the fuzzy signal controller the better the traffic control it provides. For control based on the dynamic selection of fuzzy signal controllers to be practical and effective, the generation of the different controllers off line is required. This occurs by means of a technique that takes into account the fuzzy signal controller’s main characteristics other than the partition of the universe of discourse for the fuzzy sets’ input. The genetic algorithm associated with a suitable simulation program is a promising possibility.
The continuation of this present research so as to allow for the use of a greater number of alternative controllers as well as the inclusion of other variables to the decision-making process such as average delay per vehicle depends on the insertion of the approaches proposed for dynamic selection into a robust simulation program.

Aknowledgements

The authors wish to thank CNPq and CAPES, the Brazilian agencies for R&D that provided financial support for the project, and FINATEC for supporting its presentation.

References


PHYSICAL AND OPERATIONAL DESIGN IN TRANSIT NETWORK

Luigi DELL’OLIO¹, Jose Luis MOURA², Angel IBEAS²

Abstract. The problem of designing an urban public transport system is renowned for being one of the most complex to solve in the transport sector. This paper proposes a model of physical and operational design in congested local public transport networks. We approach and solve the problem of reducing the total social cost of operating the transport system, including the costs of providing services, traveling and bus stop construction by using a bi-level mathematical programming method. The suggested model mixes optimization and simulation, and allows us to link supply and demand. The decision variables of the problem to be solved are: frequency of each line and distancing between bus stop.

1. Introduction

This paper applies a bi-level optimisation model aimed at estimating the optimum frequencies of fixed routes and finding an ideal location for the bus stops. The bibliography contains a series of studies with regard to the design of transport networks using bi-level programming techniques (1, 2, 3, 4). The main hypotheses for this model are described below:

1. The bus routes have been defined a priori (they will not be subject to change).
2. The mean speeds in the sections into which the public transport network are divided are known.
3. The demand for journeys using public transport between each point of origin and destination in the time interval modelled (O-D matrix) is known.
4. A series of possible locations for bus stops are established following standard practice.

¹ Universidad de Burgos, Campus San Amaro Edificio "La Milanera" C/ Villadiego s/n 09001 Burgos, Spain, ldellolio@ubu.es
² Universidad de Cantabria, Avda. De los Castros s/n, Código 39005, Santander, Spain, mourajli@unican.es
5. A series of network section classes are defined, each having specific common characteristics and are therefore assigned a specific distancing value between bus stops.

As far as the demand for journeys is concerned, it is considered opportune to use the case of peak times since the frequencies could easily be readjusted by applying the same model during off-peak times. The distances between bus stops cannot be easily changed throughout the day, therefore, it is important that the location of bus stops is as suitable as possible for the most critical times of the day.

2. Modeling methodology

The problem of designing an urban public transport system can be conceived as a non-cooperative game on two levels (Stackelberg’s game). The players in this game are the planner, who determines the characteristics of the transport system and the system users, who try to minimise their overall personal cost for the journey, thus producing a flow pattern over the system. As already commented, the problem of locating bus stops and the optimisation of public transport network frequencies can be considered as a bi-level mathematical programming problem. On the upper level, the total social cost involved in the operation of the transport system is minimised ($\text{Min } Z$). Particular attention will be paid to the total travel cost ($\text{CTV}$) and the cost of the company running the service ($\text{CE}$) which depend on the location of the stops and the current average distancing between the stops of each macro group ($DG_i$), and on the service frequency of each line ($f_i$) as well as the construction cost ($CC$) for the stops which depends on the number of stops required. All are subject to the capacity restrictions on the buses (2) and at the bus stops (3).

Upper Level:

$$\text{Min } Z = \text{CTV}(DG_i, f_i) + \text{CE}(DG_i, f_i) + CC(n)$$  \hfill (1)

Subject to

$$f_i \geq \frac{\text{NTP}_l}{K_b} \quad \forall l$$  \hfill (2)

$$f_s \leq \min \left( \frac{3600}{\text{TO}_s} \right) \quad \forall s$$  \hfill (3)

Where:

$\text{NTP}_l = \text{Total number of passengers in the busiest route section } l \text{ during the given time interval.}$

$\text{TO}_s = \text{Time the bus stop } K' \text{ is occupied expressed in seconds.}$

$K_l = \text{Capacity of route } l$

Lower Level:

$$\text{Min } \sum_{s \in S} \int_0^{V_s} c_i(x)dx$$  \hfill (4)

Subject to
\[
\sum_{r \in R_w} h_r = T_w \quad \forall w \in W
\]
\[
\sum_{r \in R} \delta_{sr} h_r = V_s \quad \forall s \in S
\]
(5)
\[
v_i^* = \frac{f_l \cdot V_i}{f_s} \quad \forall l \in B_s, \forall s \in S
\]
\[
h_r \geq 0 \quad \forall r \in R
\]

Where:
- **W**: Group of origin-destination pairs O-D.
- **w**: Element in the \( W \) group, in which \( w = (i, j) \) with \( i, j \) centroids.
- **\( T_w \)**: Total number of journeys between the O-D \( w \) pair for users of public transport.
- **l**: Index for designating a public transport route.
- **R**: Group of available routes for public transport users.
- **r**: Index for designating a public transport route.
- **\( R_w \)**: Group of public transport routes associated with the O-D \( w \) pair.
- \( \overline{h}_r \): Public transport passenger flow on route \( r \).
- **s**: Index for designating a public transport route section.
- **S**: Group of available route sections for public transport users.
- **c_s**: Journey cost for public transport users over the route section \( s \).
- **\( \delta_{sr} \)**: Element of the incidence matrix section of route-route: it assumes value 1 if the route \( r \) goes past \( s \) and 0 in other cases.
- **\( V_i \)**: Transport passenger flow on route section \( s \).
- **\( V_l \)**: Transport passenger flow on route section \( s \) using route \( l \).
- **f_l**: Frequency of route service \( l \).
- **\( f_s \)**: Total frequency in the route section \( s \).

### 2.1. Solution algorithm

In order to solve the bi-level mathematical programming problem, a heuristic algorithm is proposed as described below:

**Step 1**: A feasible frequency vector \( f_l \) is generated and a distance \( d_l \) is established between stops by sector which satisfy the restrictions of the upper level problem.

**Step 2**: The lower level problem of optimisation is solved; i.e., the public transport origin-destination matrix is assigned obtaining the equilibrium flows \( V_i^* \).

**Step 3**: Feasible frequencies \( f_l \) and the equilibrium flows \( V_i^* \) are inserted in the target function of the upper level and the target function \( Z \) is evaluated.

**Step 4**: The Hooke-Jeeves (5) algorithm is used to assess new frequencies \( f_l \) and the new distancing between stops \( d_l \) and a return is made to step 2.
The computing implementation of the Hooke-Jeeves algorithm was performed based on a modified version of it during the exploratory search stage of the algorithm.

3. Case study

The practical application of the model designed was tested on various bus routes in the city of Santander (Spain). The routes taken used (routes 1, 2, 5C2, 7C2) are those with the greatest demand in the urban public transport network. Between them, they cover approximately 50% of all journeys on public transport (during peak times).

Moreover, by using a suitable survey, the data of 811 bus journeys with various origins and destinations made between 1.30pm and 3.30pm were compiled, this time interval is the peak midday time for both private vehicles and public buses.

<table>
<thead>
<tr>
<th>Interval</th>
<th>Existing Situation</th>
<th>Situation Optimising Frequencies</th>
<th>Global Optimisation</th>
<th>Variation with regard to the current situation</th>
<th>Variation with regard to the situation Optimising Frequencies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interval L1</td>
<td>18 minutes</td>
<td>11 minutes</td>
<td>11 minutes</td>
<td>-59%</td>
<td>0%</td>
</tr>
<tr>
<td>Interval L2</td>
<td>20 minutes</td>
<td>12 minutes</td>
<td>11 minutes</td>
<td>-45%</td>
<td>-8%</td>
</tr>
<tr>
<td>Interval L5C2</td>
<td>10 minutes</td>
<td>16 minutes</td>
<td>17 minutes</td>
<td>-70%</td>
<td>6%</td>
</tr>
<tr>
<td>Interval L7C2</td>
<td>12 minutes</td>
<td>13 minutes</td>
<td>13 minutes</td>
<td>-8%</td>
<td>0%</td>
</tr>
<tr>
<td>Mean interval</td>
<td>14.5 minutes</td>
<td>13 minutes</td>
<td>13 minutes</td>
<td>-10%</td>
<td>0%</td>
</tr>
<tr>
<td>Passengers L1</td>
<td>432</td>
<td>524</td>
<td>544</td>
<td>26%</td>
<td>4%</td>
</tr>
<tr>
<td>Passengers L2</td>
<td>400</td>
<td>490</td>
<td>490</td>
<td>23%</td>
<td>0%</td>
</tr>
<tr>
<td>Passengers L5C2</td>
<td>359</td>
<td>232</td>
<td>186</td>
<td>-48%</td>
<td>-20%</td>
</tr>
<tr>
<td>Passengers L7C2</td>
<td>449</td>
<td>394</td>
<td>420</td>
<td>-6%</td>
<td>7%</td>
</tr>
<tr>
<td>User cost</td>
<td>8118 €</td>
<td>7290 €</td>
<td>7230 €</td>
<td>-11%</td>
<td>-1%</td>
</tr>
<tr>
<td>Company cost</td>
<td>891 €</td>
<td>982 €</td>
<td>988 €</td>
<td>11%</td>
<td>1%</td>
</tr>
<tr>
<td>Construction cost</td>
<td>185 €</td>
<td>185 €</td>
<td>203 €</td>
<td>10%</td>
<td>10%</td>
</tr>
<tr>
<td>Total cost</td>
<td>9194 €</td>
<td>8466 €</td>
<td>8421 €</td>
<td>-8%</td>
<td>-1%</td>
</tr>
<tr>
<td>Access time</td>
<td>9.19 minutes</td>
<td>8.98 minutes</td>
<td>8.88 minutes</td>
<td>-3%</td>
<td>-1%</td>
</tr>
<tr>
<td>Waiting time</td>
<td>6.15 minutes</td>
<td>4.78 minutes</td>
<td>4.42 minutes</td>
<td>-28%</td>
<td>-8%</td>
</tr>
<tr>
<td>Travelling time</td>
<td>12.41 minutes</td>
<td>29.36 minutes</td>
<td>29.66 minutes</td>
<td>-8%</td>
<td>1%</td>
</tr>
<tr>
<td>DG1</td>
<td>346 m.</td>
<td>346 m.</td>
<td>323 m.</td>
<td>-7%</td>
<td>-7%</td>
</tr>
<tr>
<td>DG2</td>
<td>316 m.</td>
<td>316 m.</td>
<td>322 m.</td>
<td>2%</td>
<td>2%</td>
</tr>
<tr>
<td>DG3</td>
<td>350 m.</td>
<td>350 m.</td>
<td>292 m.</td>
<td>-17%</td>
<td>-17%</td>
</tr>
<tr>
<td>DG4</td>
<td>480 m.</td>
<td>480 m.</td>
<td>420 m.</td>
<td>-13%</td>
<td>-13%</td>
</tr>
<tr>
<td>DGmean</td>
<td>373 m.</td>
<td>373 m.</td>
<td>339 m.</td>
<td>-9%</td>
<td>-9%</td>
</tr>
<tr>
<td>N. of bus stop</td>
<td>71</td>
<td>71</td>
<td>78</td>
<td>9%</td>
<td>9%</td>
</tr>
</tbody>
</table>

Table 1. Results.
As a first step in this study, the base situation (current situation) was simulated. An assignment was carried out to public transport and it was verified whether this assignment reproduced the current situation with some reliability.

As a second step, the company costs were calculated as a result of the assignment.

The assignment was obtained using a model of assignment to public transport with capacity restriction. This model considers the capacity restriction in vehicles (buses) and *models the interaction with private transport* considering the mean speeds covered in the sections making up the network. The total result (Table 1) obtains a reduction of 9% in the distancing between stops in the study area. The increase in the number of stops and the variation in frequencies causes the total cost to drop with regard to the solution obtained by only optimising the frequencies.

4. Conclusions

The model proposed solves the problem examined with a mixed approach of optimisation and simulation. An assignment to public transport is made between one iteration and another of the modified Hooke-Jeeves algorithm. This approach allows the link between supply and demand to be considered. In this way, the consistency between equilibrium flows and the frequencies and distancing calculated between stops in each iteration of the algorithm is guaranteed. This method should be used because it is reasonable and fair. The passengers using the system try to maximise it for their own benefit, as do the drivers and the transport companies. The problem is that passengers, drivers and businessmen don’t necessarily have the same interests in common. The model makes it possible to find an optimum social solution for all parties involved. In the practical application it was found that passengers reduced their access times, journey times and waiting times. The companies, with a small increase in operational costs, gave a much better service more in line with the existing demand and obtained a medium term increase in the demand for their services. The model helps in analysing how different bus routes interact. The user can choose between a set of routes which allows him to minimise the overall travel cost. Clearly, this cost is a function of bus frequencies and distances between stops. From a theoretical point of view (following the method used for solving the model put forward in this paper) it is very easy to change the distances between bus stops in different zones ($D_G$), given that all the possible bus stop locations were carefully chosen following the previously stated criteria. However, from a practical point of view it is much more complicated, given the relatively high costs associated with taking down and putting up a bus stop. This is why the model has more validity when putting into practice a completely new public transport service. The study performed in the city of Santander, which has an existing public transport service, is useful because it demonstrates that by following scientific criteria (the applied model) for locating bus stops, rather than purely practical criteria (real situation), important cost savings can be made. This methodology is shown to be suitable for solving these types of problems. Moreover, this model can be applied using a very limited amount of data similar to that compiled in traditional transport studies. Furthermore, the practical application of the model demonstrated that considerable savings can be made to the overall cost of the system as well as providing a substantially improved service to the user at a very low cost.
References


Abstract This paper presents the an attempt to conceptualise Decision Support System (DSS) for motorway emergency system as the most important basis for the design of new emergency management system based on ICT tools, geographic information systems and systems for operative procedures.

1. Introduction

The most developed countries of the world focus basic attention to the protection and operation of the critical infrastructure. Namely, it has been noticed that particular segments of national infrastructure are very vulnerable and exposed to the permanent threat. Even the superficial analysis point out that within critical infrastructure the transportation is particularly critical. Namely, state motorway systems due to the permanent exposure to the risky events i.e. accidents should be especially concerned. For example, Europe has an average of 1.7 million of car accidents per year assisted by emergency services, which includes the medical emergency service. In the year 2004, 43 000 people lost their lives in the car accidents in the EU member states. The straight consequences of car accidents are higher costs of health insurance systems, and national economies are burdened with less productivity and bigger range of material goods damage. In order to maximise road safety and efficacy of emergency support, along with faster remediation of car accidents as to make interruptions of the transportation flows as short as possible, Europe initiated several projects to increase safety on the roads such as E-Merge [8], eSafety [9] and eCall [7].

Regarding the Pan-European In-Vehicle Emergency Call (eCall) Service a key recommendation of the first eSafety Communication was the establishment of a harmonised, pan-European in-vehicle emergency call (eCall) service that builds on the single European emergency number 112.

Concerning its efficiency, the eCall system could be of the special interest. After the crash, sensors placed in the cars send both data and oral call to the nearest 112 PSAP, as

---

1 University of Split, Faculty of Civil Engineering and Architecture, mladineo@gradst.hr
2 University of Split, Faculty of Civil Engineering and Architecture, knezic@gradst.hr
3 University of Split, Faculty of Civil Engineering and Architecture, njajac@gradst.hr
well as to eCall operators. Individual call contains the following data: location, time, vehicle identification and eCall status. eCall system should lower the intervention time for 50% in the rural and 40% in the urban areas. This obviously could mean decreasing the mortality rate and the severity of the injuries happened during the accident. Regarding the fact that car accidents lead to the traffic overload and could cause secondary accidents, by lowering the intervention time, the traffic overload would be reduced for 20%, thus enabling introduction of interrupted or slowed transportation flows.

Acceptance of the eCall technology means that European countries have to improve their PSAP by the end of 2007. Moreover, the whole emergency management system should be improved with necessary ICT tools which enable quick and reliable response to the car accidents and fully utilize the advantages of eCall technology. Regarding the promotion of eCall at national level the Commission strongly recommends that the Member States set up national platforms for promoting eCall [7]. They should have participation from relevant ministries including the authorities responsible for emergency services, as well as private industry and service providers.

Accepting the fact that Croatia is host for numerous tourists and the fact that in the last year a contemporary motorway Zagreb-Split was completely constructed, a project concerning novel concept of emergency help on the new motorway has been launched. The main objective of the project is to use both technical and technological solutions proposed by eCall project in order to decrease the intervention time (50% in rural and 40% in urban regions). Until complete introduction of the eCall system a part-time solutions will be implemented according to the technology principles proposed by eCall. The main partial goal will be to locate the accident spot as precise as possible.

The main obstacle of the system could be the ability to identify accurately the accident location, because participants of the accident or perhaps, direct witnesses who call emergency services usually are not able to make precise estimation of the accident location. By introduction of eCall technology, each car will have embedded GPS device, which will define the precise accident location. As the introduction of eCall technology is expected not earlier than the year 2009, when the eCall units will be built in the cars during the production, at this moment the accident location, regarding present technology of 112 call systems, should be the main issue.

There is also the question about the capability of the motorways and other road infrastructure to handle rescue operations triggered by eCall. As it is already mentioned there is certain responsibility by the mobile phones providers to improve their systems as well. There is also lots of organisational problems such as finding the optimal route for the emergency vehicles in the means of satisfying several criteria such as: the best access route, fastest route, most convenient route regarding some special rescue vehicles etc. So far, the GIS technology could help to solve such problem.

2. **Pilot-project for Split-Dalmatia County**

Taking into account that the eCall Driving Group that was established to foresee the implementation has set 2009 as a target year for full roll-out, a stepwise action plan for the Croatian motorways has been developed. As the launching of the eCall technology has to improve performance of the PSAP by the end of the year 2007, the intention has been to
improve whole emergency system in the meantime with necessary ICT tools that will enable fast and reliable car accident response and fully utilize the benefits of the eCall technology. Moreover, the emergency system should be adjusted to the existing motorway system. It is not as easy task as it probably seems because motorway networks are closed systems and entering or exiting them is strictly regulated. Additionally, high speed of the vehicles could endanger whole emergency operation. Therefore, using system analysis, the preliminary plans for the emergency management should be developed.

The development of eCall system was the reason to both motivate and pursue all relevant ministries (transportation, communication, internal affairs), as well as telecommunication providers, according to the EC objective, to reduce the traffic accident rate for 50% by 2010.

Action plan will have two main phases. **The first phase** will include the following actions:

- Establishment of common protocol for all emergency services that serve motorway (Emergency plans methodologies and processes);
- Establishment of temporary caller location system for mobile phones;
- Optimisation of the capacity and schedule of the motorway emergency resources;
- Dissection of the motorway in the optimal number of segments according to the power of the caller location system;
- Dispatching the emergency resources to the each motorway segment and determination of the fastest “approaching route” to the accident location using so called “Blue corridor system” which warns road users of approaching emergency vehicle;
- Introduction of AVL (Automatic Vehicle Locating) system for all motorway emergency vehicles [5];
- Determination of the optimal evacuation route for injured people and selection of medical emergency resources in the hospitals according to the degree and type of injuries;
- Establishment of the Coning system, namely warning approaching road users of accident.
- Conceptualisation of DSS which would comprise the data and procedures.

**The second phase** will involve stepwise introduction of eCall technology in collaboration with the telecommunication providers, as well as telemedicine system which involves the possibility to send medical data using any form of communication interface, from the patient to any other site. This typically includes biomedical information such as ECG, Pulse, Oxygen Saturation, Blood Pressure, Video Imaging.

The analysis of the possibility to implement the first phase pointed out to the several critical points responsible for the success of the whole project. Besides other things, the establishment of temporary emergency caller location from mobile phones is defined as crucial technical problem. Namely, efficacy and quickness of emergency help increase significantly if the caller location could be determined precisely during the call. Unfortunately, telecommunication providers are not ready to invest into new technologies that will improve mobile phone caller location, so it is necessary to obtain mobile phone caller location as precise as possible. One technical solution providers are ready to implement is determination of the cell which receives the signal from the mobile phone. Based on the direction of the cell’s axis and amplitude of the signal providers can approximately determine the caller. However, the denser is the GSM station network along
the motorway the higher is the precision of caller location. Additional problem arises from the fact that in Croatia exist three mobile telecommunication providers that work in the separate base stations systems, so two distinct networks having different locations of the base stations has to be analysed.

Relatively unfavourable spatial distribution of emergency services, their significant distance from the motorway, demand very precise emergency plans. The first phase of the Pilot project has begun with structural definition of the traffic accidents in accordance with the evaluated parameters. As indicated by the parameters, the intervention procedures for the emergency plans are defined. Plans are evaluated by using very useful “decision table”, (Figure 1.). It enables straightforward assessment of the car accident complexity, and according to the previously described procedure the resources could be easily selected. In accordance with the decision table the valuation of the available resources is being performed, in order to determine real available resources.

<table>
<thead>
<tr>
<th>Intervention type</th>
<th>Emergency degree assessment</th>
<th>Number of vehicles involved in the accident</th>
</tr>
</thead>
<tbody>
<tr>
<td>Technical services</td>
<td>0,2</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Extra 1.5</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>3 to 5</td>
<td>More than 5</td>
</tr>
<tr>
<td>Police</td>
<td>0,5</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Medical emergency service</td>
<td>0,75</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Extra 3,0</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>More than 9</td>
</tr>
<tr>
<td>Fire brigades</td>
<td>1,5</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>Fire brigades and Medical emergency service</td>
<td>1,5</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Extra 5,5</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>Special technical squads</td>
<td>+0,4</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>+2,0</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>+8</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>+9</td>
<td>More than 9</td>
</tr>
<tr>
<td>Special requirements</td>
<td>+0,5</td>
<td>Out of category</td>
</tr>
<tr>
<td></td>
<td>+2,0</td>
<td>Out of category</td>
</tr>
</tbody>
</table>

**Figure 1: Decision table for determination of the traffic accidents complexity**

In order to establish coordination between all emergency services, as well as other services which are closely connected, the functional analyses was performed for the area near the motorway. Three important emergency levels for the system elements which are in the function for the intervention on the motorway are defined.

Criteria for the classification of all participants of the emergency system were:
- extent of the territory from which the calls for emergency are being received both from fixed and mobile phone network,
- operational time and number of employees who receiving emergency calls,
- number of the available on-duty brigades prepared for the emergency action,
- frequency of the emergency calls in the past two years.
The three defined levels in the certain way represent “availability level”, meaning that they are indirect indicators of both the response and mobility degree for each unit of the emergency system.

The next step deals with an optimisation of both capacity and schedule of the emergency resources according to the characteristics of the each motorway segment. As it is emphasized in the action plan, segmentation of the motorway into the optimal number of the segments is performed according to the possibility to locate emergency calls. It is in directly depends on the number of the GSM stations that receive from the motorway.

Figure 2 show the segmentation procedure according to the GSM stations’ positions of providers.

For each of the motorway segment, previously defined in the Emergency plans, available intervention resources are given according to the defined “availability level” and shortest path methods:

- the units that are being directed to the accident spot for the accident degree I;
- the routes that vehicles should take during transfer to the motorway (depending on traffic conditions at the time);
- the motorway entrances that vehicles should take and the direction of the route (to avoid bottleneck which could occur after the accident);
- the evacuation routes (towards defined motorway exits) and at the hospitals in which injured people should be transferred.

In the case of complex accident with greater number of the victims, units that are accredited for the intervention degree II, or perhaps degree III in the situation of accident with numerous victims, are charged by applying abovementioned procedure.
3. Conclusion

Improving the emergency services that can save lives and reduce the injuries caused by traffic accidents is posed as imperative. Croatia, like EU countries, is willing to cope with such challenge. Introduction of eCall which is fully connected with 112 emergency system requires good emergency plans adjusted to the existing motorway system. The conceptualised decision support system could give basis for the further implementation of the emergency system on motorways. The study of an introduction of eCall technology in motorway network and its connection with 112 system demands an intensive research and application of both decision support methodology and technological possibilities.

References


USING MECHANISTIC DESIGN PROCEDURES OF PAVEMENTS TO OPTIMIZE TRANSPORTATION PLANNING

Swetha KESIRAJU and Hussain BAHIA

Abstract. Optimization of transportation facilities for capacity and pavement condition could be achieved with mechanistic analysis of pavement structures. This paper is a call for integrating the soft (operations) side of transportation and the hard (pavement and structures) side of transportation. It is focused on using a tool developed recently in the USA (the AASHTO M-E Design Guide) to show how traffic volumes and traffic loadings can be optimized for better overall pavement management. It is intended to show planners what happens in pavements when traffic volumes, traffic speeds, and total transported loads change. It shows quantitative sensitivity analysis of typical pavement structures (rigid and flexible pavements) to highlight the main factors that affect pavement availability to carry traffic. The sensitivity analyses were conducted using the Mechanistic – Empirical Pavement Design Guide software to study the effects of traffic design input parameters on pavement performance. Pavement performance included specifically faulting, transverse cracking, and smoothness for rigid pavements. It also included smoothness, longitudinal cracking, alligator cracking, transverse cracking, and permanent deformation for flexible pavements. Based on the sensitivity results, the input parameters were ranked and categorized from most sensitive to insensitive to help pavement designers to identify the level of importance for each input parameter and also identify the input parameters that can be modified to satisfy the predetermined pavement performance criteria. It is expected that ranking could also help planners to determine how traffic of heavy vehicles could be directed to enhance the service life and ensure better maintenance strategies.

1 University of Wisconsin at Madison, 3350 Engineering Hall, 1415 Engineering Drive, Madison, WI, 53706, USA. Bahia@engr.wisc.edu
1. **Introduction**

The National Cooperative Highway Research Program (NCHRP) sponsored research projects 1-37 and 1-37a, resulting in the development of a Mechanistic-Empirical (M-E) pavement analysis system. This computerized system is expected to replace the currently used AASHTO design guide, most recently revised in 1993. There have been many concerns expressed regarding the applicability and effectiveness of the 1993 AASHTO Guide for considering current traffic loads, advances in materials design and characterization, and lack of consideration of the effect of environmental factors on pavement performance. The newly released Mechanistic-Empirical Design Guide (MEPDG) predicts pavement distresses using calibrated distress prediction models. It is based on using principles of mechanics in which critical material characteristics, climatic conditions, and detailed traffic conditions are used to estimate pavement response during various stages of service life. The responses are used to predict critical pavement distresses using damage accumulation models that include coefficients determined through national calibration efforts using databases for major full scale projects such as the Long-Term Pavement Performance (LTPP) project.

2. **Objective**

This paper is intended to suggest integrating the soft side (operations) of transportation and the hard side (pavement materials and structures) of transportation using the new design system which allows detailed evaluation of effects of changes in traffic conditions on pavement performance. The study is focused on using MEPDG which has been recently introduced in the United States as a tool for more accurate and effective pavement analysis system. One of the main features of this new system is to allow detailed evaluation of effect of traffic variables including volume of traffic, axle load distribution, speed and others on the various performance indicators of pavements. The objective of this study is to specifically show how traffic volumes, traffic loadings, and pavement conditions can be optimized for better overall transportation management. It includes analyses to show planners what happens in pavements when traffic volumes, traffic speeds, and total transported loads change. It includes a quantitative sensitivity analysis of typical pavement structures (rigid and flexible pavements), used in the state of Wisconsin, to highlight the main factors that affect pavement capability to carry traffic. The sensitivity analyses were conducted to study the effects of traffic design input parameters on pavement performance.

3. **Traffic Design Inputs**

In the MEPDG, 4 basic types of traffic input data are required for pavement design:
- Traffic volume – base year information.
- Traffic volume adjustment factors- Monthly adjustment; Vehicle class distribution; Hourly truck distribution; Traffic growth factors.
- Axle load distribution factors.
General traffic inputs- Number axles/trucks; Axle configuration; Wheel base; Traffic Speed

The sensitivity analysis of traffic was performed for 3 rigid and 2 flexible pavement sections. The rigid pavements include 1. Conventional Rigid Pavement; 2. Stabilized Base Pavement; 3. An overlay of (Jointed Plain Concrete Pavement. The two flexible pavements include 1. Conventional Flexible Pavement used in Wisconsin; and 2. An Overlay of Hot Mix Asphalt (HMA) on HMA pavement. The design input parameters were divided into two groups: fixed input parameters and varied input parameters, as shown in Table 1.

Table 1 Fixed traffic input parameters of the rigid pavement sections

<table>
<thead>
<tr>
<th>Fixed Input Parameter</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Traffic General</strong></td>
<td></td>
</tr>
<tr>
<td>No. of lanes in design direction</td>
<td>2</td>
</tr>
<tr>
<td>% of trucks in design direction (%)</td>
<td>50</td>
</tr>
<tr>
<td>Mean wheel location(in)</td>
<td>18</td>
</tr>
<tr>
<td>Design lane width(ft)</td>
<td>12</td>
</tr>
<tr>
<td><strong>Traffic Volume Adjustment Factors</strong></td>
<td></td>
</tr>
<tr>
<td>Hourly truck distribution</td>
<td>Default</td>
</tr>
<tr>
<td>Traffic growth factor</td>
<td>1.5%</td>
</tr>
<tr>
<td>Axle load distribution factors</td>
<td>Default</td>
</tr>
<tr>
<td><strong>Axle Configuration</strong></td>
<td></td>
</tr>
<tr>
<td>Average axle width(ft)</td>
<td>8.5</td>
</tr>
<tr>
<td>Dual tire spacing(in)</td>
<td>12</td>
</tr>
<tr>
<td>Axle Spacing – Tandem, Tridem, Quad</td>
<td>51,6,49,2,49,2</td>
</tr>
<tr>
<td>Tire Pressure(psi)—Single and Dual tire</td>
<td>120, 120</td>
</tr>
</tbody>
</table>

The values for the other traffic parameters used in the analysis, and the pavement structure that were defined by the pavement designers are shown in Table 2, for the rigid pavement, and in Table 3, for the flexible pavement.

Table 2 Traffic and Structure input parameters of the rigid pavement sections

<table>
<thead>
<tr>
<th>Conventional Pavement</th>
<th>Stabilized Base</th>
<th>Overlay</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traffic</td>
<td>Traffic</td>
<td>Traffic</td>
</tr>
<tr>
<td>AADTT</td>
<td>1178</td>
<td>AADTT</td>
</tr>
<tr>
<td>% trucks in the design lane</td>
<td>85</td>
<td>% trucks in the design lane</td>
</tr>
<tr>
<td>Operational Speed(mph)</td>
<td>70</td>
<td>Operational Speed(mph)</td>
</tr>
<tr>
<td>Traffic Wander SD</td>
<td>10</td>
<td>Traffic Wander SD</td>
</tr>
<tr>
<td><strong>Structure</strong></td>
<td><strong>Structure</strong></td>
<td><strong>Structure</strong></td>
</tr>
<tr>
<td>Layer 1: JPCP</td>
<td>10.0&quot;</td>
<td>Layer 1: JPCP</td>
</tr>
<tr>
<td>Layer 2: A-1-a</td>
<td>3.0&quot;</td>
<td>Layer 2: Cement Stabilized</td>
</tr>
<tr>
<td>Layer 3: Crushed Gravel</td>
<td>8.0&quot;</td>
<td>Layer 3: A-1-a</td>
</tr>
<tr>
<td>Later 4: SP</td>
<td>NA</td>
<td>Layer 4: A-2-4</td>
</tr>
<tr>
<td>Layer 5: A-6</td>
<td>NA</td>
<td>Layer 5: A-6</td>
</tr>
</tbody>
</table>
Table 3 Traffic and Structure input parameters of the flexible pavement sections.

<table>
<thead>
<tr>
<th>Conventional Pavement</th>
<th>Overlay</th>
<th>Traffic</th>
<th>Traffic</th>
</tr>
</thead>
<tbody>
<tr>
<td>AADTT</td>
<td>AADTT</td>
<td>1178</td>
<td>2000</td>
</tr>
<tr>
<td>% trucks in the design lane</td>
<td>% trucks in the design lane</td>
<td>85</td>
<td>95</td>
</tr>
<tr>
<td>Operational Speed(mph)</td>
<td>Operational Speed(mph)</td>
<td>70</td>
<td>60</td>
</tr>
<tr>
<td>Traffic Wander Standard Deviation</td>
<td>Traffic Wander Standard Deviation</td>
<td>10</td>
<td>10</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Structure</th>
<th>Structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Layer 1: Asphalt Concrete</td>
<td>Layer 1: Asphalt Concrete</td>
</tr>
<tr>
<td>Layer 2: A-1-a</td>
<td>Layer 2: Asphalt Concrete(existing)</td>
</tr>
<tr>
<td>Layer 3: Crushed Stone</td>
<td>Layer 3: Crushed Stone</td>
</tr>
<tr>
<td>Later 4: SP</td>
<td>Layer 4: A-6</td>
</tr>
</tbody>
</table>

4. Traffic Sensitivity Analysis

In this analysis, climate and structure are kept fixed for all sections. Four input parameters of traffic i.e., Annual Average Daily Truck Traffic, Percent of trucks in the design lane, Operational Speed and Traffic Wander Standard Deviation were varied while the rest of traffic inputs were kept fixed. The ranges of magnitude for the varied input parameters were selected based on the recommendations of MEPDG and experts in the pavement section of Wisconsin DOT.

To define the level of sensitivity of pavement performance to each of the traffic parameters, a ratio was used to determine the relative change in a response parameter relative to the total range used in the MEPDG. Depending on the value of the ratio the level of sensitivity is defined as shown in Table 4.

\[
\text{Ratio} = \frac{\text{Maximum Value of Range} - \text{Minimum Value of Range}}{\text{Target change}} \times 100
\]

Table 4 Sensitivity level definition

<table>
<thead>
<tr>
<th>Sensitivity Level</th>
<th>Abbreviation</th>
<th>Ratio (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very Sensitive</td>
<td>VS</td>
<td>&gt;50%</td>
</tr>
<tr>
<td>Sensitive</td>
<td>S</td>
<td>25%-50%</td>
</tr>
<tr>
<td>Insensitive</td>
<td>I</td>
<td>&lt;25%</td>
</tr>
</tbody>
</table>

As an example of the sensitivity analysis Table 5 shows the results for the three main performance indicators (IRI(in/mi), % Cracking, and Faulting(in)) as a result of changing the Average Annual Daily Truck Traffic (AADTT) for the conventional rigid pavement.
Table 5  Sensitivity of the performance indicators of rigid pavement to changes in AADTT.

<table>
<thead>
<tr>
<th>Distress Target Limit</th>
<th>IRI(in/mi)</th>
<th>% Cracking</th>
<th>Faulting(in)</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;172</td>
<td>&lt;172</td>
<td>&lt;15</td>
<td>&lt; 0.12</td>
</tr>
<tr>
<td>1000</td>
<td>77.6</td>
<td>0.3</td>
<td>0.004</td>
</tr>
<tr>
<td>2500</td>
<td>82.4</td>
<td>1.4</td>
<td>0.012</td>
</tr>
<tr>
<td>5000</td>
<td>90.7</td>
<td>4.4</td>
<td>0.023</td>
</tr>
<tr>
<td>10000</td>
<td>107.8</td>
<td>13.1</td>
<td>0.042</td>
</tr>
</tbody>
</table>

\[
\text{Ratio} = \frac{107.8 - 77.6}{172 - 0} = 17.5\% \text{(Insensitive)}
\]

Table 6 shows a similar example for the conventional flexible pavement. In this case the performance indicators include IRI, longitudinal cracking (L/C)(%), alligator cracking (A/C)(%), transverse cracking (T/C)(%), and permanent deformation (P/D)(in) in the top pavement layer.

Table 6  Sensitivity of flexible pavement performance to changes in AADTT

<table>
<thead>
<tr>
<th>Distress Target Limit</th>
<th>IRI(in/mi)</th>
<th>L/C(ft/500)</th>
<th>A/C (%)</th>
<th>T/C (ft/mi)</th>
<th>PD(AC)(in)</th>
<th>PD(TP)(in)</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 172</td>
<td>&lt; 1000</td>
<td>&lt; 25</td>
<td>&lt; 1000</td>
<td>&lt; 0.25</td>
<td>&lt; 0.25</td>
<td>&lt; 0.75</td>
</tr>
<tr>
<td>1000</td>
<td>139.4</td>
<td>25.2</td>
<td>0.1</td>
<td>1</td>
<td>0.12</td>
<td>0.27</td>
</tr>
<tr>
<td>2500</td>
<td>139.4</td>
<td>101</td>
<td>0.4</td>
<td>1</td>
<td>0.19</td>
<td>0.35</td>
</tr>
<tr>
<td>5000</td>
<td>139.6</td>
<td>284</td>
<td>0.8</td>
<td>1</td>
<td>0.27</td>
<td>0.44</td>
</tr>
<tr>
<td>10000</td>
<td>139.9</td>
<td>905</td>
<td>1.9</td>
<td>1</td>
<td>0.39</td>
<td>0.57</td>
</tr>
</tbody>
</table>

Using the sensitivity criterion the MEPDG was used to calculate the changes in all performance indicators and to determine their sensitivity to changes in traffic conditions. Table 7 and Table 8 show the traffic & structure information, and the sensitivity analysis results of the rigid pavements respectively.

Table 7  Sensitivity analysis results of the three rigid pavement sections.

<table>
<thead>
<tr>
<th>Conventional</th>
<th>Investigated Values</th>
<th>IRI</th>
<th>Cracking</th>
<th>Faulting</th>
</tr>
</thead>
<tbody>
<tr>
<td>AADTT</td>
<td>1000, 2500, 5000, 10000</td>
<td>S</td>
<td>VS</td>
<td>S</td>
</tr>
<tr>
<td>% Trucks</td>
<td>65, 75, 85, 95</td>
<td>I</td>
<td>I</td>
<td>I</td>
</tr>
<tr>
<td>Traffic SD</td>
<td>7, 9, 11, 13</td>
<td>I</td>
<td>I</td>
<td>I</td>
</tr>
<tr>
<td>Op. Speed</td>
<td>45, 60, 75, 90</td>
<td>I</td>
<td>I</td>
<td>I</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Stabilized Base</th>
<th>Investigated Values</th>
<th>IRI</th>
<th>Cracking</th>
<th>Faulting</th>
</tr>
</thead>
<tbody>
<tr>
<td>AADTT</td>
<td>1000, 2500, 5000, 10000</td>
<td>VS</td>
<td>VS</td>
<td>VS</td>
</tr>
<tr>
<td>% Trucks</td>
<td>65, 75, 85, 95</td>
<td>I</td>
<td>I</td>
<td>I</td>
</tr>
<tr>
<td>Traffic SD</td>
<td>7, 9, 11, 13</td>
<td>I</td>
<td>S</td>
<td>I</td>
</tr>
<tr>
<td>Op. Speed</td>
<td>45, 60, 75, 90</td>
<td>I</td>
<td>I</td>
<td>I</td>
</tr>
<tr>
<td>Investigated Values</td>
<td>IRI</td>
<td>Cracking</td>
<td>Faulting</td>
<td></td>
</tr>
<tr>
<td>---------------------</td>
<td>-----</td>
<td>----------</td>
<td>----------</td>
<td></td>
</tr>
<tr>
<td>AADTT 1000, 2500, 5000, 10000</td>
<td>VS</td>
<td>VS</td>
<td>VS</td>
<td></td>
</tr>
<tr>
<td>% Trucks 65, 75, 85, 95</td>
<td>I</td>
<td>S</td>
<td>I</td>
<td></td>
</tr>
<tr>
<td>Traffic SD 7, 9, 11, 13</td>
<td>I</td>
<td>S</td>
<td>I</td>
<td></td>
</tr>
<tr>
<td>Op. Speed 45, 60, 75, 90</td>
<td>I</td>
<td>I</td>
<td>I</td>
<td></td>
</tr>
</tbody>
</table>

Note: VS - Very Sensitive, S - Sensitive, I - Insensitive.

Table 8 shows the Flexible Pavements Sensitivity

<table>
<thead>
<tr>
<th>Investigated Values</th>
<th>IRI</th>
<th>L/C</th>
<th>A/C</th>
<th>T/C</th>
<th>P.D(AC)</th>
<th>P.D(TP)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AADTT 1000, 2500, 5000, 10000</td>
<td>I</td>
<td>VS</td>
<td>I</td>
<td>I</td>
<td>VS</td>
<td>S</td>
</tr>
<tr>
<td>% Trucks 65, 75, 85, 95</td>
<td>I</td>
<td>I</td>
<td>I</td>
<td>I</td>
<td>I</td>
<td>I</td>
</tr>
<tr>
<td>Traffic SD 7, 9, 11, 13</td>
<td>I</td>
<td>I</td>
<td>I</td>
<td>I</td>
<td>I</td>
<td>I</td>
</tr>
<tr>
<td>Op. Speed 45, 60, 75, 90</td>
<td>I</td>
<td>I</td>
<td>I</td>
<td>I</td>
<td>I</td>
<td>I</td>
</tr>
</tbody>
</table>


From the sensitivity results and for the above investigated values, it can be seen that for rigid pavements the IRI, Cracking and Faulting are most sensitive to AADTT. It is also observed that Cracking is most sensitive to the % Trucks and Traffic Wander Standard Deviation. None of the rigid pavement performance indicators are sensitive to the Operational Speed. The results of the flexible pavements (Table 8) show that IRI, A/C, and T/C are not sensitivity to any of the input traffic parameters. It is however observed that L/C (%), P.D is sensitive or very sensitive to AADTT. Also P.D for the overlay is sensitive to Traffic SD and to Op. Speed (mph).

5. Effect of Pavement Structure:

The results of the sensitivity analysis for traffic inputs were not expected and somewhat not clear since most of the performance measures were not found to be sensitive to traffic input changes. One possible explanation is that the pavement sections selected are over designed for the traffic inputs. To further study sensitivity to traffic conditions, the thickness of the top layer of rigid and flexible pavements was varied for both the rigid and flexible pavements. Only two traffic variables (AADTT and Operational Speed) were also varied for all layer thicknesses considered.
Figure 1 depicts the results for the rigid pavement conventional section. Some interesting trends are observed.

- It can be seen that IRI is very sensitive to layer thickness. For example, a 2” increase in the pavement thickness (From 8” to 10”) would allow increasing traffic from 2000 to 20000 AADTT while keeping the terminal IRI at 200. This is a 10 folds increase in traffic by only a 2” increase in thickness. This type of analysis would allow planners to predict how much more truck traffic can be added without deteriorating the smoothness of pavements.

- Similar trends can be observed for faulting. The trends for cracking, on the other hand show a slightly different trend. The change from 8” and 10” pavements are not important when AADTT is above the value of 10000.

Figure 1. Results of Sensitivity Analysis of Rigid Pavements
Observing the plots for the effects of Operational Speed, it can be seen that there is no effect on any of the performance indicators regardless of the thickness of the layer. It is shown that the terminal IRI, faulting and cracking are highly affected by thickness but not by the operational speed.

The results for flexible pavements are shown in Figure 2. They indicate:

- IRI and PD are sensitive to AADTT only for the thin layer (4”). In the case of 6” and 8” pavement layer, there appears to be very little effect of increasing traffic or increasing traffic speed.
- Fatigue cracking (A/C)(%) is also very sensitive to AADTT but it is not as sensitive to traffic speed. The sensitivity is similar for all layer thicknesses.
- Rutting in the asphalt layer is also very sensitive to AADTT. It is also affected by traffic speed. Rutting estimated at 20 miles per hour for a 6.0 “ layer is equivalent to rutting of a 4.0” layer at 90 miles per hour.

Figure 2. Results of Sensitivity Analysis of Flexible Pavements
6. Concluding Remarks

The sensitivity of typical pavement performance sections used in the U.S.A. to various traffic and structural parameters was studied using the Mechanistic-Empirical Pavement Design Guide (MEPDG). Pavement performance included specific distresses as shown in Table 9. Based on the sensitivity results, the input parameters were ranked and categorized from most sensitive to insensitive to help pavement designers identify the level of importance for each input parameter and also identify the input parameters that can be modified to satisfy the predetermined pavement performance criteria. The following table provides the summary of trends observed.

<table>
<thead>
<tr>
<th>Traffic</th>
<th>Rigid Pavement</th>
<th>Flexible Pavement</th>
</tr>
</thead>
<tbody>
<tr>
<td>AADTT</td>
<td>VS-cracking; S-faulting</td>
<td>S-IRI; VS-(faulting, cracking)</td>
</tr>
<tr>
<td>% Trucks</td>
<td>I I</td>
<td>S-cracking</td>
</tr>
<tr>
<td>Traffic Speed</td>
<td>I</td>
<td>I</td>
</tr>
<tr>
<td>Layer thickness</td>
<td>VS- IRI, Cracking and Faulting</td>
<td>NA</td>
</tr>
</tbody>
</table>

Note: VS – Very Sensitive; S – Sensitive; I – Insensitive

Based on the results of this study, pavement designers can identify the level of importance for each input parameter and also identify the input parameters that can be modified to satisfy the predetermined pavement performance criteria. It is expected that ranking could also help planners to determine how traffic of heavy vehicles could be directed to enhance the service life and ensure better maintenance strategies.

References:

7. Federal Highway Administration, Long Term Pavement Performance Study.
Abstract. The polarisation phenomenon has impacts on the spatial distribution of mobility and, therefore, on the transportation systems asset. From the analysis of where and how these polarities have been developed it is possible to understand how the functional organisation of urban areas is changing. Moreover, in this paper, it is highlighted the consistency between the morphologies of the activity settlements and the railways’ networks.

1. Introduction

The focus of this paper is on the monitoring of both the polarisation phenomena for the Italian metropolitan areas and of the morphology of local railways’ systems. The Italian metropolitan areas are the ones reported in the law 142 of the year 1990 and they are Turin, Milan, Genoa, Bologna, Venice, Rome, Florence, Naples, Bari, Palermo, Catania and Cagliari.

The work is still in progress and the methodology and preliminary results are here reported only for three areas, representative of three different urban systems typologies, chosen as examples. The latter are the areas of Rome, Naples and the Veneto region [2] [3].

The main objective of this analysis, carried out through the definition of some ad hoc indicators, is to provide the consistency between urban and rail system developments and to demonstrate, at the same time, how the rail system supports polarisation phenomena and how the urban form modifies railway lines.

2. Methodology description

The polarisation phenomena have been analysed through four steps using both socio-economic data and mobility information.

* Department of Transportation Engineering, University of Naples Federico II, via Claudio 21, 80125 Naples, Italy, {madeluca|fpagliar}@unina.it
a) **Distribution of activities.** For all the areas the spatial distribution of residents and activities has been considered in both the decades 1981-1991 and 1991-2001 in order to analyse their trends.

b) **Identification of attractive municipalities.** The indicator adopted to define the attractive municipalities has been the *attractive capacity*, computed as the ratio between the “destinated” and “generated” flows, i.e. in-coming and out-coming systematic relationships from each municipality. In Italy, Census data provides information concerning the places of work and study of the residents of each municipality. By number of in-coming systematic relationships in a given municipality $i$ it is meant the number of residents of other municipalities who have their places of work or study in municipality $i$. On the other hand, by number of out-coming systematic relationships from a given municipality $i$ it is meant the number of residents of that municipality who have their places of work or study in another municipality. All the information refers both to the 1991 and 2001 Census. The ratio between the “destinated” and “generated” flows indicates a measure of the prevailing capacity of each municipality of attracting people for work and study purposes over the capacity of generating flows towards other municipalities for the same purposes.

c) **Identification of polarised systems.** For the municipalities identified in the previous step, it has been defined the indicator “intensity of total systematic relationship” as:

\[
\text{Intensity}_{\text{tot syst rel mun}} = \frac{\sum_j \frac{\text{Gen}_{ij}}{\text{Res}_j}}{n} + \frac{\sum_i \frac{\text{Dest}_{ij}}{\text{Res}_j}}{n}, \quad i \neq j
\]

where:

- $n$ is the number of municipalities of the region to which municipality $i$ belongs;
- $\text{Gen}_{ij}$ are the generated systematic relationships from municipality $i$ to municipality $j$;
- $\text{Dest}_{ij}$ are the destinated systematic relationships to municipality $i$ from municipality $j$;
- $\text{Res}_j$ are the residents of municipality $j$.

This indicator measures the entity of connections between one municipality and the surrounding area.

From the analysis of the indicator of the intensity of total systematic relationships it has come out a polarisation of the examined urban areas. The comparison of the indicators relative to the years 1991 and 2001 shows that the polarisation phenomenon is present in both the years and, rather, it increases in the decade. The polarisation phenomenon is revealed through Second Level Urban Systems (SLUSs), i.e. sets of neighbouring municipalities with a reference pole. All the SLUSs are reciprocally integrated into a First Level Urban System. Three different urban systems typologies can be identified: many municipalities with high attractiveness with one prevailing over the others; many municipalities with high attractiveness of the same level; only one big municipality with high attractiveness with the remaining ones gravitating around it. The three different urban phenomena have been named polarised, multi-pole and monocentric respectively [3].

d) **Endowment and development of regional rail systems.** Local railway lines have been examined. For each of them the route, the stations, the municipalities served by the lines, the timetable have been considered to test whether it exists a consistency between transport supply and urban form.
3. Three examples

In the case of the area of Naples, the analysis of the different indicators highlights the existence of a main pole, that of Naples, and of five poles belonging to Second Level Urban Systems (highlighted in black in Fig. 1). The Neapolitan area can be considered as an example of a polarised structure.

For the Veneto area it has come out the presence of poles of same relevance (highlighted in black in Fig. 2): this area represents an example of a multi-pole urban structure.

In the case of the area of Rome, no other pole can be identified except the capital, whose attractiveness is extended on a wide and populated area (see Fig. 3). This is a proof of a monocentric structure.

In all the three cases, it is shown that the poles, when they exist, coincide with the stations of the local railway systems.

Figure 1. Polarised structure: the case of the city of Naples.
Figure 2. Multi-pole structure: the case of the Veneto region.

Figure 3. Monocentric structure: the case of the city of Rome.
4. Territorial dynamics and impacts on the transportation system

The research proposed focuses on the relationship between the urban transformation of the Italian metropolitan areas and the morphology of the local rail networks. In the last decades activities have been mainly located over space according to two overlapping models:

− one, which can be considered traditional and which represents the “background”, characterized by a density decreasing with distance from the main centre;
− the other, which has come out only recently, characterised by the display of Second Level Urban Systems (SLUSs) around the main and more ancient centres.

Both phenomena are supported by rail networks. The latter have first allowed the development of the monocentric metropolitan areas; then, in the last decades, they have supported the growth of the SLUSs, making them accessible through the main rail links. This phenomenon is of great interest for transportation planners as it gives directions for strategic choices. In particular, it highlights that, in these circumstances, public transport can be used in a condition which makes it more competitive with respect to the private system. Rail systems, from tram to light rail, from regional to fast rail can be destined to the main connections and to link the poles of the SLUSs [1]. Buses, on the other hand, able to satisfy a limited travel demand on short distances and in areas at low densities, are suitable for minor urban systems and can provide the connection within each SLUS as well as among the SLUSs themselves. As a whole, it is easier to work in the direction of a re-equilibrium of the transportation system towards the public system.

References


TOWARDS A SECURE RAILWAY TRANSPORT SYSTEM

Giuseppe SCIUTTO¹, Mirco BERRUTTI¹, Diego MAZZINI, Cristian VEGLIA.

Abstract. The problem of security in railway transport system is becoming a concern of high priority due to the large penetration and large use by million of people. Railway transport system is defined as a critical infrastructure and the security issues require an integrated approach between what are considered the best practices for reduction of vulnerability and a scientific and methodological approach for risk management. The author aims not only at identifying the steps and provisions already developed but also highlights in this publication the research made in this domain and proposes a methodological approach tailored for the railway system in order to assess the efficiency of specific countermeasures adopted and the related residual risk.

1. Introduction

The events to which we have assisted in the last years have underlined, on world staircase, the fact that any Country today can feel to the sure one from terrorist acts.

After events such as the Tokyo Sarin Gas subway attack of March 1995, the Madrid train bombings of March 2004 or London suicide bombings of July 2005 and the continuous threats to which the whole occidental world is submitted, transportation security has become a concern of high priority.

Transport systems are increasingly targeted because they are used by millions of people a day and are easily penetrated; substantial casualties can be inflicted with a relatively small device, given the typical passenger density in a railway or subway station.

The problem has to be approached in terms of vulnerabilities of the system, assessment of risk through risk analysis and risk management, evaluating security threats, intended as an indication of the likelihood that a specific type of attack will be initiated against a specific target or class of target; threat may include any indication with the potential to cause the loss or damage to an asset; it can also be defined as an adversary’s intention and capability to undertake actions that would be detrimental to a valued asset.

According to leading US experts in rail security, attacks against railways are more numerous and deadly than those on airports and planes, and bombs are the most frequently

¹ Consorzio Nazionale Interuniversitario per i Trasporti e la Logistica Via Opera Pia 11a Genova Italy
used weapons in these attacks. At a recent Rail Industry Safety Conference held in the US, experts noted that there had been more than 181 attacks on trains and related rail targets worldwide between 1998 and 2003, in such countries as Colombia, India, Spain, Pakistan, the UK, the US, and Venezuela, an average of 30 per year (these numbers do not include the 2004 Madrid bombing). Despite such events, public transit is still an extremely safe form of travel. The traffic fatality rate per passenger-kilometre is less than one-tenth that of automobile travel, as indicated in Figure 1. Even including terrorist attacks and other crimes against transit passengers, transit is far safer than private vehicle travel.

Figure 1: UK Death Rate By Mode (Steer Davies Gleave, 2005)

Figure 2 shows that traffic fatality rates decline with increased transit ridership; London has one of the lowest fatality rates of all cities listed.

Figure 2: International Traffic Deaths (Kenworthy and Laube, 2000)

The extreme thoughtfulness of the matter is evident as you notice that the happening of a disastrous event by hand of terrorists causes direct and indirect effects, more and more wide and complexes, not only relatively to the loss of human life, but also to the psychology of the people with a progressive conditioning of the life up to the paralysis of the system, with devastating repercussions on the whole economy.
There are several reasons that people react particularly strongly to terrorist attacks (Adams, 2005). Such attacks are designed to be highly visible, producing intense media coverage. The fact that the harm they cause is intentional rather than accidental makes them particularly tragic and frustrating. And they raise fears that such attacks may become more frequent or severe, so risks may increase in the future.

Travellers would increase their total risk if they shift from transit to driving in response to terrorist threats. Analysis by Gigerenzer (2004) and Sivak and Flannagan (2004) indicate that in the three months after the 11 September 2001 terrorist attacks, shifts from air to automobile travel caused several hundred additional roadway traffic fatalities. Because of actions by governments and the airline industry to increase air travel security, these travel shifts have been reduced, reducing excess traffic deaths. For these further reasons strong action is justified to protect transit users’ safety and sense of security.

2. Railway system vulnerabilities

Therefore new issues concerning security has risen, assuming more and more importance and underlining new aspects of vulnerability in a delicate system and complex as that of the transports both it of people that of commodities.

Vulnerability can be defined as a "weakness that can be exploited to gain access to a given asset"[2]. The API/NPRA expands this definition to include "... and subsequent destruction or theft of [the] ... asset"[1], whereas criticality is typically defined as a measure of the consequences associated with the loss or degradation of a particular asset.

Weaknesses, like criticality, can be categorized in a number of ways:

- physical (accessibility, relative locations, visibility, toughness, strength, etc.)
- technical (susceptible to cyber attack, energy surges, contamination, eavesdropping, etc.)
- operational (policies, procedures, personal habits)
- organizational (e.g. would taking out headquarters severely disrupt operations), etc.

Vulnerability assessment is a systematic process which consists on the identification of these weaknesses in physical structures, personal protection systems, processes or other areas that may be exploited. A vulnerability assessment identifies inherent states and the extent of their susceptibility to exploitation relative to the existence of any counter-measures.

Vulnerabilities are assessed by the analyst against specific attacks. API/NPRA (American Petroleum Institute (API) and the National Petrochemical and Refiners Association (NPRA)) in their Security Vulnerability Assessment Methodology (SVA) identifies three steps to assessing vulnerabilities:

- Determine how an adversary could carry out a specific kind of attack against a specific asset (or group of assets);
- Evaluate existing countermeasures for their reliability and their effectiveness to deter, detect, or delay the specific attack
- Estimate current state of vulnerability and assign it a value

Rail transportation system is very difficult to protect from security threats because of its nature: although the system is remarkably resilient, it is also open and decentralized: the
volume of ridership, the number of kilometers of tracks and of access points make it impractical to subject all rail passengers and their baggage to the type of screening of airline passengers: aviation-style security is impossible to viable on it, for the reason that railway system has to be open, economic and accessible and queues and delays would bring the system to a halt.

Nevertheless steps can be taken to reduce the risks of an attack as focusing attention on the need to address whatever security vulnerabilities may exist in the railroad system, making a systematic analysis of transportation assets, the risks to those assets after an exhaustive evaluation of all possible security threats for the rail. On January 11 2005, the 9/11 Commission Report and Congressional Response mentioned the Aviation and Transportation Security Act, which established the Transport Security Administration (TSA) and assigned the TSA the responsibility of developing strategic plans to provide security for critical parts of the USA transportation system. The Commission expressed concern that 90% of the annual federal investment made in transportation security goes toward commercial aviation security without a systematic risk assessment to determine if this is the most cost-effective allocation of resources. The Commission noted that “major vulnerabilities still exist in cargo and general aviation, and that the security improvements in commercial air traffic may shift the threat to ports, railroads and mass transit system.

3. Risk

3.1. Definition

Risk implies uncertain consequences. Risk is defined as the “... probability of loss or damage, and its impact ...” or essentially the expected losses should a specific target/attack scenario occur”[4]. “Expected” loss is determined by multiplying the estimated adverse impact caused by a successful threat/attack scenario by the probabilities associated with threat and vulnerability. API/NPRA defines risk as “a function of:
• consequences of a successful attack against an asset;
• likelihood of a successful attack against an asset”.
“Likelihood” is defined as “a function of:
• the attractiveness of the target to the adversary (based on the adversary’s intent and the target’s perceived value to the adversary)
• degree of threat (based on adversary’s capabilities)
• degree of vulnerability of the asset”

3.2. Risk reduction activities

Risks can be reduced in a number of ways:
• by reducing threats (e.g. through eliminating or intercepting the adversary before he strikes);
• by reducing vulnerabilities (e.g. harden or toughen the asset to withstand the attack);
• by reducing the impact or consequences.
For each potential countermeasure, the benefit in risk reduction should also be
determined. More than one countermeasure may exist for a particular asset, or one
countermeasure may reduce the risk for a number of assets. Multiple countermeasures
should be assessed together to determine their net effects. The analyst should also assess
the feasibility of the countermeasure.

Examples of sample protective measures are:
- Perimeter barriers, high-tech fencing and lighting.
- Intrusion detection equipment
- Alternative external communications capability for continuity of operations
- Increased number of uniformed and undercover patrols
- Training for personnel
- Increased number of inspections of trash receptacles and other storage areas -- this also
  included securing, closing off, or removing some of these areas where things may be
  stored and removing some trash receptacles
- Increased number of bomb detecting canine teams (presently, trained dogs represent an
  excellent broad-spectrum, high-sensitivity sensory system for investigating bomb
  threats)
- Continued broadcast of public announcements to alert riders and citizens to be aware,
  watch their surroundings, and report any suspicious activities or abandoned property
  such as back packs, garbage bags, etc.
- Increased video surveillance and review of such materials
- Procurement of personal protective equipment for emergency responders
- Modern rail station design by eliminating columns, dark corners, and other areas where
  criminals can lurk and that can minimize the effect of potential attacks.
- Random passenger bag checks.

### 3.3. Best Practice

#### 3.3.1. A new IP-based CCTV surveillance system

Two examples of best practises are the new CCTV systems recently installed (by Bosch) on
lines in the Aachen area of the German Rail (DB) network and the Lisbon Metro in
Portugal. Security systems must include CCTV surveillance cameras, public announcement
and alarm systems, and public information and help points. With the legacy of past
developments, these systems, often difficult and expensive to install, mostly operate on
separate networks and are based on different technologies, resulting in decentralization
and inefficiency. The ready availability of Local Area Network (LAN) connectivity based on
the Internet Protocol (IP) represents something of a revolution. In this system CCTV
cameras, info points, alarm and announcement systems are integrated into a LAN, each
with its own IP address and all overseen and managed from a central control room. Each
camera is connected with a single-channel digital video recorder/transceiver unit and all
these units send security camera footage to the control room via an optical-fibre LAN. The
units also feature a new codec with the capability of offering both MPEG-2 and MPEG-4
video compression, that allows DVD-quality video to be transmitted live via the IP network
to the control room while recording the video, introducing a delay of just 100 milliseconds to the video transmission against 500/1000 of legacy systems. This means that it allows for easy zooming and panning to follow events in real-time captured on camera.

3.3.2. K9 Program

In the USA, the DHS (Department of Homeland Security) has planned to develop a rapid deployment Mass Transit K-9 program by utilizing existing Homeland Security explosive K-9 resources. These mobile response teams would be specially trained to work in the undergrounds and tunnels unique to some transit and rail environments and would be used primarily in order to detect explosives. For the K-9 program, the TSA (Transportation Security Administration) primarily uses sporting breeds--such as Labradors, Chesapeake Bay Retrievers and Golden Retrievers--that usually are obtained from breeders and canine vendors from across the United States and Europe. These breeds were chosen for their gentle temperament and keen sensory capabilities.

The Department also plans to implement a pilot program: Transit Inspection Pilot (TRIP), to test the feasibility of screening luggage and carry-on bags for explosives at rail stations and aboard trains. Unlike airports, passengers can board trains at several stops, some as simple as a lone platform. As a result, the pilot program would not resemble an aviation-type solution to transit and rail, but rather provide the Department with a venue to test new technologies and screening concepts.

4. Methodology

Many models/methodologies have been developed by which threats, vulnerabilities, and risks are integrated and then used to inform the cost-effective allocation of resources to reduce those risks.

A scientific risk-based approach, managing risks in terms of threat, vulnerability and consequence is needed, including the following steps:

Assessments
- identify assets (specific sites and facilities) and identify which are most critical;
- identify, characterize, and assess threats, from deliberate events;
- identify and assess hazards/vulnerabilities of critical assets to specific threats and their potential consequences;
- determine the risk (i.e. the expected consequences of specific types of attacks on specific assets), assessing the likelihood and magnitude of risks.

Using Assessments to Identify and Prioritize Risk Reduction Activities
- identify and characterize risk reduction activities, developing a strategy and action plan to reduce risks establishing a series of preventive and protective steps that would increase security at multiple levels;
- define processes and measures for verifying and evaluating the management of risk.
RISK MANAGEMENT

RISK ANALYSIS

SECURITY MANAGEMENT PROCESS

Define scope
Threat ID
Vulnerability Assessment
Risk likelihood
Threat Impacts
Risk Determination
Documentation

RISK MANAGEMENT

Identify Potential Controls
Cost-Benefit Analysis
Prioritize Controls
Select Controls

Assign Responsibility
Develop Implementation Plan for Risk Mitigation
Implement Controls
Testing & Validation
Address Residual Risk

Control review
5. **Conclusions**

This paper has defined a methodology for the risk assessment of a railway transport system as far as the security is concerned, including countermeasures and their effectiveness. In the full paper also a detailed taxonomic classification of the system and related vulnerabilities, criticalities and threats will be provided.

6. **References**


Abstract. This paper suggests an application for using prescriptive roadside route guidance as an ATMS measure during incident situations. The guidance is calculated in an optimal control setting and intends to optimize the system performance given an incident. In order to do so, a simulation of the accident is made in terms of the moment it starts, the effect on road capacity and the moment it ends. Results indicate that dependent on the type, location and duration of the accident in this case a compliance rate of only 30% was able to allow the route guidance to maximize network efficiency which would make it applicable in practice.

1. Introduction.

Route guidance in general can be used to inform drivers on network conditions be it pre trip, roadside or in car [1], more specifically it can be used to inform travelers about an accident on the network. This paper focuses on roadside route guidance of a prescriptive nature [2]. The effectiveness of route guidance is limited by the number of people who actually use / follow it. When presenting travelers with roadside prescriptive information, e.g. a VMS stating: “for direction 1 follow road B”, the number of people who follow this advice can be expressed by the compliance factor. The compliance factor is the fraction of people for whom a road side advice contains valuable information, e.g. drivers actually heading to direction 1, and who follow the advice. This compliance factor is expressed as a number between 0 and 1 and its influence on the effectiveness of the route guidance for an accident situation is the focus of this paper. In section 2 the DTA model DSMART is briefly discussed which is used to simulate the traffic and incident. In section 3 a sample network and accident description is given. In section 4 the optimization technique for generating route guidance is briefly discussed and the results are given for different levels of compliance. The last section discusses the optimization results.

1 Delft University of technology, PO BOX 5048, 2600 GA Delft. F.s.zuurbier@tudelft.nl, h.vanlint@citg.tudelft.nl
2. DSMART Traffic Model performance measure

For this study, a Dynamic Simple Macroscopic Road Traffic assignment model is used (DSMART) [2]. A traffic network is discretized into time and space (a time step of 20 seconds, and links are divided into cells based on this time step and maximum link speed). Traffic flow is modeled based on the kinematic wave model (LWR model [3]) and numerically approximated by the Godunov scheme as done by Lebaque [4]. The simulation duration is split up into a number of periods and for every period a new set of route choices can be calculated / used. The model is implemented in Matlab.

All assignments in DSMART are evaluated by means of an assignment score and the score used in this study can be calculated by means of (1) and resembles a system optimum (SO).

\[ F_{OBJ} = \frac{\sum_{t=0}^{T} tQ_{o}^{Total}}{\sum_{d} D_{d} - \sum_{d} A_{d}} \]  

(1)

The total OD demand is cumulatively summed for the total number of time steps in the simulation (the divided), whereas the divisor is calculated as being the difference between the cumulative summed number of departures and arrivals for all origins and destinations. These latter can be calculated by means of (2) where L indicates origin or destination link i.

\[ \int_{0}^{T} L_{i,d,t}^{Travellers} dt, \forall i \in \text{Origin, Destination} \]  

(2)

By evaluating the difference between the cumulative number of departures and arrivals, at a certain time step one in fact calculates the generalized gross travel time (GGTT). To optimize the score in (1) is to minimize this GGTT and comes down to: "getting as much people to their destination as soon as possible".

3. Experimental setup: Sample network

The sample network and OD demand pattern is described below. The network resembles a two directional two, lane ring way network around a city.
Table 1 Size of the routable flows

<table>
<thead>
<tr>
<th>Origin</th>
<th>Total flow [veh/hr]</th>
<th>Destination</th>
<th>Routable flow [veh/hr]</th>
<th>%</th>
<th>Direction</th>
</tr>
</thead>
<tbody>
<tr>
<td>18</td>
<td>1992</td>
<td>36</td>
<td>255</td>
<td>12%</td>
<td>East – West</td>
</tr>
<tr>
<td>6</td>
<td>1825</td>
<td>24</td>
<td>336</td>
<td>18%</td>
<td>South - North</td>
</tr>
<tr>
<td>23</td>
<td>1812</td>
<td>5</td>
<td>336</td>
<td>19%</td>
<td>North – South</td>
</tr>
<tr>
<td>35</td>
<td>1685</td>
<td>14</td>
<td>217</td>
<td>13%</td>
<td>West – East</td>
</tr>
</tbody>
</table>

Saturday scenario

The demand pattern is chosen to create a scenario which resembles a Saturday morning traffic load where there are no signs of congestion. The simulation is carried out between 08:00 and 10:00 and is divided into 8 periods. Every period the route choices are updated,
but since there are no real changes in link travel times, all route choice is based on the actual shortest travel time paths.

3.1. The accident

In figure 1, the red star indicates the position in the network where an accident occurs (link 19, heading south on 3/6 of the link length). The accident is a non lethal simple head to tail collision. In the picture below the effect of this collision is described in terms of a multiplication factor for the total available link capacity with values between [0..1] as a function of time between 08:00 and 10:00 hours.

Figure 2 Link capacity availability as a function of time due to the accident

Given this accident scenario, an assignment is made without changing the route choices, e.g. the travelers stay at their original intended routes as they would do in the reference situation. The resulting assignment is visualized in the picture below.
In this assignment, all travelers are assumed to use their original route as they would normally do. Clearly this is sub optimal and the therefore corresponding score to this assignment has the low value of 1502.8371. An assignment has also been made given this accident with the probit route choice model active. This would mean that every period route choices would be recalculated and updated based on user perceived shortest time paths and the assignment looks more like a user equilibrium assignment. This type of assignment scored 1822.6358.

Both assumed assignments during the incident are based on unrealistic assumptions. The first assumes no knowledge of the current network (accident) at all, whereas the other assumes full network knowledge (be it user perceived) and in the latter case a change in actual OD demand is not taken into account. If something exceptional happens in the network, non-informed drivers often stay at the routes which they have chosen for normal conditions [6]. These drivers and the system could benefit from appropriate rerouting using prescriptive route guidance and therefore the first type of assignment is used as the reference assignment and will be optimized using route guidance.

4. Optimized routing

The optimization technique is based on an optimal control loop and an evolutionary algorithm (EA) [7]. The latter is a chance based search process which is used in order to find the optimal destination specific split fractions at the VMS locations which maximize the objective function. Every split vector at the VMS locations can have a real value
between [0 1] and [1 0] (Vector has two elements corresponding to two downstream connected links). To keep computation tractable a significance of 0.1 per element is chosen., which means that one split vector can have 11 combinations. Since there are 4 VMS’s installed and the simulation is divided into 8 periods a total number of 11^4*8 combinations exists. The EA searches this space for an optimal VMS scenario which maximizes the objective function. Such a search process takes time and the implemented prototype can not yet be used online since it would take about 250 minutes to evaluate 500 generations implemented parallel on 8 standard PC’s (generation size of 8). After these 500 generations, the objective function is maximized to the best ability of the EA and an optimal form of control is calculated. However; since the search process is chance based it can not be guaranteed that the absolute global is found, yet it is likely to be very near.

This prototype can however be used to assess the influence of the exogenous determined compliance factor, in terms of how this relates to the effectiveness of the calculated guidance during such an accident situation. In order to do so, the accident situation described is optimized using different levels of compliance and the results are presented in the next section.

4.1. Optimization results

In total 10 different levels of compliance have been investigated and their values are given in the table and figure below.

<table>
<thead>
<tr>
<th>Compliance</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>100%</td>
<td>1694.3801</td>
</tr>
<tr>
<td>90%</td>
<td>1694.3801</td>
</tr>
<tr>
<td>80%</td>
<td>1694.3801</td>
</tr>
<tr>
<td>70%</td>
<td>1694.3801</td>
</tr>
<tr>
<td>60%</td>
<td>1694.3801</td>
</tr>
<tr>
<td>50%</td>
<td>1689.5202</td>
</tr>
<tr>
<td>40%</td>
<td>1692.1274</td>
</tr>
<tr>
<td>30%</td>
<td>1691.9413</td>
</tr>
<tr>
<td>20%</td>
<td>1671.1475</td>
</tr>
<tr>
<td>10%</td>
<td>1644.935</td>
</tr>
<tr>
<td>5%</td>
<td>1550.982</td>
</tr>
<tr>
<td>0%</td>
<td>1502.8371</td>
</tr>
<tr>
<td>Normal</td>
<td>2049.4315</td>
</tr>
<tr>
<td>UO</td>
<td>1822.6358</td>
</tr>
</tbody>
</table>

Figure 6. Score values for different compliance rates

It can be seen that the maximum effect of the achievable route guidance in this situation is achieved at a compliance rate of 60% only. The slightly lower value of the score found at
50% can be attributed to the stochastic nature of the EA optimization algorithm. The algorithm is likely to find the global or near global optimum but can not guarantee it. However, the most interesting result is that for this accident situation at only 30% compliance rate, the maximum achievable network optimization is already almost found.

5. Conclusions

This paper investigated the influence of different levels of user compliance on the effect of partial network optimization using route guidance during incident situations. It was found that for a simple incident and network, a compliance rate of only 30% was enough to maximize the effect of the route guidance in terms of a lower generalized gross travel time. This percentage relies heavily on the characteristics of the accident, and the size of the routable flows. In both cases it was attempted to use realistic values. For real life application, achieving a compliance rate of 30% does not seem unreasonable which makes the proposed type of routing strategy feasible from that point of view.

6. References

A HYPER-NETWORK MODEL FOR SIMULATING PARK-AND-RIDE TRIPS

Luca D’ACIERNO†, Mariano GALLO‡, Bruno MONTELLA†

Abstract. This paper proposes a hyper-network approach for simulating park-and-ride trips assuming elastic travel demand, as regards the mode choice; this model estimates user flows on both road and transit networks by an equilibrium assignment algorithm, which simulates modal split and transfers between road and transit systems (park-and-ride trips) on the network. The hyper-network model is based on dummy links’ cost functions that are able to simulate “on-network” modal split and transfers. An algorithm for solving the equilibrium assignment problem on the hyper-network is proposed and tested on a trial network.

1. Introduction

In this paper a hyper-network model for simulating park-and-ride trips is proposed. The model allows user flows to be calculated on both road and transit networks by a rigid demand assignment algorithm, simulating the modal split and the transfers between road and transit systems (park-and-ride trips) on the network. We propose cost functions that have to be adopted on the dummy links for simulating “on-network” the modal split and the transfers. Moreover, with these functions it is possible to prove the mathematical equivalence of the results obtained with a Multinomial Logit route choice model on the hyper-network with those obtained by applying a Nested Logit model (mode choice and route choice) on monomodal networks.

The proposed approach is a method for solving the elastic demand assignment model ([1]), assuming that only mode choice is elastic and the available modes are road, transit and park-and-ride (road-transit combined trips).

A literature review on elastic demand assignment models can be found in [2], where a hyper-network approach (without combined modes) is also proposed. [5] and [10] studied...
some hyper-network approaches. Simulation of the combined modes was studied among others by [4], [6], [7] and [9]. [12] proposed the Cross Nested Logit model that allows mode choice to be simulated when combined modes, such as park-and-ride, are available.

2. The proposed model

Generally, a park-and-ride transportation system can be viewed as consisting of three kinds of elements: road network elements; transit network elements; interchange elements.

For simulating trip choices on a multimodal system that also allows park-and-ride trips, we propose a model that assigns the all-mode OD matrix to a hyper-network. The assignment model assumes that demand is rigid, that is the all-mode matrix is not influenced by the assignment results. Moreover, mode choice is included in route choice on the hyper-network.

Besides, in the following we hypothesise that users can choose among three transportation systems: road, transit and park-and-ride; the first two systems are monomodal, while the third is a combined mode.

For simulating the mode choice where there are combined modes, the Cross Nested Logit model ([12]) can be adopted; it considers the park-and-ride mode belonging to both road and transit systems. In this paper we propose splitting the park-and-ride (P+R) modes into two virtual modes: P+R belonging to the road system and P+R belonging to the transit system; operating in this way, we may state that the Cross Nested Logit is equivalent to a Nested Logit, defining a rate of belonging of P+R to the transit system (the rate 1 – αm is the rate of belonging to the road system). We also prove this equivalence and propose a simple method for estimating the value for αm.

For each od pair, the probability of choosing path k (on the hyper-network), belonging to mode m (road mode, transit mode, P+R road mode or P+R transit mode), belonging to system s (road system or transit system) can be written as:

\[ P(\text{smk}) = P(\text{sm}) \cdot P(k/\text{ms}) \]

Using a Multinomial Logit model for route choice and a Nested Logit model for mode choice, it is possible to write:

\[ P(k/\text{ms}) = \frac{\exp(V_{k/\text{ms}}/\theta)}{\sum_{i \in I_{\text{ms}}} \exp(V_{i/\text{ms}}/\theta)} \]

\[ P(\text{ms}) = \frac{\exp(\theta_{1} \cdot Y_{1} + SE_{1}/\theta)}{\sum_{i \in I_{\text{ms}}} \exp(\theta_{1} \cdot Y_{1} + SE_{1}/\theta)} \]

\[ \sum_{i \in I_{\text{ms}}} \exp(\theta_{i} \cdot Y_{i} + SE_{i}/\theta) \]

\[ \frac{\exp(\theta_{i} \cdot Y_{i} + SE_{i}/\theta) + \exp(\theta_{m} \cdot Y_{m} + SE_{m}/\theta + \ln \gamma_{m,i}/\theta)}{\sum_{i \in I_{\text{ms}}} \exp(\theta_{i} \cdot Y_{i} + SE_{i}/\theta) + \exp(\theta_{m} \cdot Y_{m} + SE_{m}/\theta + \ln \gamma_{m,i}/\theta)} \]

with:

\[ Y_{m,i} = \ln \sum_{i \in I_{\text{ms}}} \exp(V_{i/\text{ms}}/\theta) \]

\[ Y_{i} = \ln \sum_{w \in I_{\text{ms}}} \exp((\theta_{i} \cdot Y_{i} + SE_{i}/\theta) + \ln \gamma_{w,i}/\theta) \]

where:

\[ V_{k/\text{ms}} \] is the systematic utility of path k, of mode m belonging to system s;

\[ I_{\text{ms}} \] is the set of paths [systems, modes].
\( \theta_0 \) is the Gumbel parameter for the system [mode, path] choice Logit model; 
\( \theta_1 Y_{m/s} \) is the inclusive utility of mode \( m \) [system \( s \)]; 
\( SE_{m/s} \) represents the socio-economic and specific attributes of mode \( m \) [system \( s \)]; 
\( \gamma_{m/s} \) is a parameter that depends on the rate of belonging of a mode \( m \) to a system \( s \): 
\( \gamma_{m/s} = 1 \) if the mode is “pure” (only road or only transit); \( \gamma_{m/s} = \gamma_{m,s}(\alpha_m) \) if the mode is combined. For \( P+R \) modes: \( \gamma_{m/s} = \alpha_m \) if \( s \) is transit and \( \gamma_{m/s} = 1 - \alpha_m \) if \( s \) is road.

With these assumptions the probability \( P(smk) \) can be written as:

\[
P(smk) = \frac{\exp((\theta_1 Y_{m/s} + SE_{m/s})/\theta_0)}{\sum \exp((\theta_1 Y_{m/s} + SE_{m/s})/\theta_0)} \cdot \frac{\exp((\theta_1 Y_{m/s} + SE_{m/s} + \theta_1 \cdot \ln \gamma_{m/s})/\theta_1)}{\sum \exp((\theta_1 Y_{m/s} + SE_{m/s} + \theta_1 \cdot \ln \gamma_{m/s})/\theta_1)} \cdot \frac{\exp(V_{m,s}/\theta_1)}{\sum \exp(V_{m,s}/\theta_1)}
\]

(6)

In this paper we propose to adopt the hyper-network model depicted in figure 1.

---

**Figure 1. Hyper-network supply model.**
Let $V_{smk}$ be the systematic utility of a path on the hyper-network. The route choice Logit model can then be written as:

$$P(smk) = \frac{\exp(V_{smk}/\theta_k)}{\sum_{s} \sum_{m} \sum_{i} \exp(V_{ism}/\theta_k)}$$

(7)

In the paper, we prove that the model (7) is equivalent to the model (6) if the following cost function is adopted for system links:

$$c'_{s/out} = -\theta_s \left( \frac{\theta_s}{\theta_i} - 1 \right) Y_s - \frac{\theta_s}{\theta_i} \cdot SE_s \quad \forall s \in I_s$$

(8)

and the following cost function is adopted for mode links:

$$c'_{m/out} = -\theta_s \left( \frac{\theta_s}{\theta_i} - 1 \right) Y_m/s - \frac{\theta_s}{\theta_i} \cdot SE_m/s - \theta_s \cdot \ln \gamma_m/s \quad \forall m \in I_m/s, \forall s \in I_s$$

(9)

3. Solution algorithm and preliminary results

The calculation of user flows on the hyper-network can be obtained by solving a (rigid demand) fixed-point assignment problem (see for instance [1]), adopting the Multinomial Logit route choice model (7).

As regards theoretical properties, we may state the existence and uniqueness of the equilibrium solution; the uniqueness cannot be stated if we consider that the transit costs depend on road costs in the case of non-exclusive facilities (crossed congestion).

In order to apply the Multinomial Logit model also on real dimension networks, we propose the implicit enumeration of paths by the Dial algorithm ([3]) and its adaptation to transit systems ([8]).

For solving the assignment problem we use the MSA-FA algorithm (see [1] and [11]). The details of the implemented solution algorithm are reported: calculation of costs on dummy links requires a backward approach, while the network loading requires a forward approach. The proposed algorithm is implemented in Visual Basic code.

The proposed model and algorithm were applied on a trial network, considering three user categories (with different car availabilities and parking times at destination) and assuming that the centre is parking priced. Initial results show the applicability of the model and the equivalence of models (6) and (7).

References


RESEARCH ON USERS’ BEHAVIOR CHANGES
DURING THE TRAM STOP-SHARING

Shuichi MATSUMOTO¹, Hironao KAWASHIMA²

1. INTRODUCTION

In Japan, where motorization has been realized, car has become essential for today’s society and economy as a tool to lead a daily life and as a transportation to support the industries. In addition, as automobile transportation has become more developed and increased, the road facilities have been improved actively.

However, motorization has not always provided good aspects. Since automobile transportation has increased excessively, road conditions, regulatory developments and automotive technologies cannot keep up with the change. It is no exaggeration to say that the whole automobile society lost its balance.

Especially in the local cities, public transport network is not developed as much as that of the large city areas. The development of motorization is causing significant decline in the number of public transportation users. Moreover, heavy traffic caused by the increase in car users is lowering the service level of public transport system, leading to further decrease in users of the system.

¹ Regional ITS Infrastructure Research Center, Kochi University of Technology, Japan
E-mail: matsumoto.shuichi@kochi-tech.ac.jp
² Faculty of Science and Technology, Keio University, Japan. E-mail: kawashima@ae.keio.ac.jp
The heavy traffic and the environmental contamination resulting from these things have been observed even in the local cities. Therefore, it is now necessary to promote and activate utilization of public transportation.

Based on the above background, this report has focused on the Tram Stop-Sharing, which is the transportation measure aiming at reducing the traffic/environmental loads and at improving the level of the public transportation service at urban area where both tram and bus are traveling side by side. It also has aimed at the quantitative evaluation on the effect of its introduction at the tram stop “Prefectural Museum Street” and the bus stop “Prefectural Museum” in Takasu-shinnmachi in Kochi-city, by the before-and-after evaluation on the change in traveling time, waiting time, and consciousness of the users through person probe, bus probe using GPS terminal and a monitoring questionnaire. Figure 1 shows the before-and-after image of the Tram Stop-Sharing at the relevant stop.

2. OUTLINE OF THE EXPERIMENT

The experimental schedule is shown in Table 1.
<table>
<thead>
<tr>
<th>Item</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subjects recruitment</td>
<td>Sep. 1, 2005 (Thu) ~ Oct. 18, 2005 (Tue)</td>
</tr>
<tr>
<td>Orientation</td>
<td>Nov. 4, 2005 (Fri)</td>
</tr>
<tr>
<td>Evaluation experiment</td>
<td>Nov. 7, 2005 (Mon) ~ Dec. 16, 2005 (Fri)</td>
</tr>
<tr>
<td>Follow-up hearing survey</td>
<td>Dec. 8, 2005 (Thu) ~ Dec 16, 2005 (Fri)</td>
</tr>
</tbody>
</table>

Table 1. Experimental Schedule

In this experiment, the introduction effects of the Tram Stop-Sharing on the subjects’ behavior during commuting to school have been verified based on the determination of the required time for schooling and transferring and the subjective evaluation of the transfer service and the public transport. The difference between the experimental condition of the before evaluation session and that of the after evaluation session is summarized in Figure 2 and the definitions of the commuting time and the transferring time in Figure 3.

Figure 2. Experimental conditions for the before and after evaluations
3. Estimation of introduction effects

3.1. Traveling time, transferring time and transferring time ratio

The change of the commuting time, transferring time, defined in the chapter 3, and the transferring time ratio, as an indicator of the transfer service, between the before evaluation and the after evaluation have been examined. The transferring time ratio is defined as Formula (1).

\[ S = \frac{\sum_i \sum_j t_{ij}}{\sum_i \sum_j t'_{ij}} \]  

\( S \) is the transferring time ratio. \( t_{ij} \) and \( t'_{ij} \) are the transferring time and the commuting time for the subject called \( i \) on the day called \( j \), respectively. Table 2 shows the change in *the traveling time, the waiting time and the transferring time ratio. *the transferring time, commuting time and transferring time ratio.
As shown in Table 2, the commuting time and transferring time are reduced after the Tram Stop-Sharing. Moreover, it is observed that the transferring time ratio was improved substantially by 20.5%. From the point of the transportation behavior, these facts have suggested that the Tram Stop-Sharing is very effective for the required time reduction.

3.2. Subjective evaluation

We have used a questionnaire by e-mail to survey the degree of congestion, waiting time at the stop and satisfaction level with public transportation/transfer services on a daily-basis. The results are summarized in Figure 4

<table>
<thead>
<tr>
<th>Table 2. Change in the transferring time ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Before evaluation</td>
</tr>
<tr>
<td>Transferring time</td>
</tr>
<tr>
<td>After evaluation</td>
</tr>
<tr>
<td>Transferring time</td>
</tr>
<tr>
<td>Improvement ratio</td>
</tr>
</tbody>
</table>

Figure 4. Results of subjective evaluation
In Figure 4, the subjective evaluation scores show that the subjects may notice the waiting time reduction at the bus stop during the after evaluation. Moreover, this has led to the rise in the scores of the whole transportation service. The intentions to paying the public transportation fares for commuting have also been increased. Therefore, it is suggested that the Tram Stop-Sharing and the loop bus service can enhance the service quality of public transportation.

3.3. Follow-up survey

The results of the follow-up hearing performed after the experiment with the subjects are summarized below.

- During the experiment, it was impossible to arouse the interests of non-daily users in the utilization of the public transportation. On the other hand, we found out that more than half of the subjects were interested in participating the similar experiment in the future.
- Almost all subjects answered that the loop bus is better than the route bus for commuting.
- Near 70 percent of the subjects felt less worried about what time to leave home after the Tram Stop-Sharing compared with before it.

4. Conclusions and future tasks

4.1. Results

The summary of the results obtained from this research is as follows.

The Tram Stop-Sharing and the loop bus service can shorten the traveling time and the transferring time and the transferring time ratio.

The Tram Stop-Sharing can reduce the subjects’ stresses during the transit.

The Tram Stop-Sharing leads to the improvement of the intermodal capacity and raises the users’ intention of paying the public transportation fares.
4.2. Future tasks

Only young people were included in the experiment of this research. It is necessary to perform experiments on more general type of users.

Regarding the method to analyze the transportation behaviors of pedestrian and bus using portable GPS terminal employed in the experiment, further validation to correct the measuring errors is necessary.

In the future, in addition to the improvement of the hardwares used in this research for determining the effectiveness of the Tram Stop-Sharing, it is required to incorporate the measures of the software, which meet the local needs, into this ITS service to make it more practical.

Acknowledgement

The authors would like to thank the many people for the support and advice which provided us with valuable findings. In particular, we very much appreciate National Institute for Land and Infrastructure Management, Kochi prefectural government, Tosa National Road Office of Ministry of Land Infrastructure and Transport, Tosa Electrical Railway Co., Ltd., Time Agent Co., Ltd., Oki electric Industry Co., Ltd., and Tosaden Dream Service Co., Ltd. who gave us numerous materials. Also we would like to thank Associate Professor Uno, Kyoto University, Professor Kumagai and Associate Professor Terabe, Kochi University of Technology, and Mr. Sadahiro, ProbeCar .net for the advice at various stages of this research which made our research smoothly.

References

MAXIMIZATION OF ACOUSTICAL AIRPORT CAPACITY: PARAMETERS AND VARIABLES

Nicola GUALANDI, Luca MANTECCHINI

Abstract. Airport environmental capacity is going to pose a serious threat to the growth of aviation. It has been proved that environmental airport capacity is essentially driven by acoustical criteria, which implies, with the introduction of the EU Regulation 30/2002 a de-facto constraint on the air traffic growth. An airport operator must learn how to manage acoustical capacity in order to accommodate future traffic. It emerges the need to investigate deeply airport acoustical capacity and understand how it is affected by variables such fleet mix, number of movements and time of the day. It is used the Airport G. Marconi of Bologna as a practical case study to investigate the relation between fleet mix and width of the noise contours. It is investigated as well, the relation between an increase in the length of the runway and a reduction in the noise exposition.

1. Noise pollution and air traffic growth

Noise pollution is likely to affect soon the growth of the entire aviation market in Europe and world wide.

It’s widely accepted that annoyance caused by air transport is higher than the annoyance caused by other mode of transport and this is due to the presence of low frequency components and the separation from the background noise that makes this kind of noise well perceptible in comparison with other environmental noise.

According to IATA in the 2003 there were between 1.63 and 2.21 millions of people exposed to aircraft noise in Europe. A survey conducted by the European Environmental Agency, published in 2001, pointed out that around the 10% of the EU population may be highly annoyed by air transport noise.

The EU Transport Commission has stated: “it’s the environmental capacity of airports, much more than physical or financial constraints, which cause imbalances and impediments to growth” [13].

1 DISTART – Transportation, University of Bologna.
n.gualandi@unibo.it; luca.mantecchini@unibo.it
The cause of the problem can be traced back in a lack of planning in the growth of the suburbs during the sixties and the seventies that led to urbanize areas close to the existing runways. The growth of aviation market and the consequential expansion of airports in terms of number of flights and dimensions did the rest. It’s nevertheless true that the last two decades have witnessed tremendous improvements in the construction of aeronautic engines in regard of noise emissions and consequently noise emissions per airplane have been declining in most of the airports, but the steady growth of the aviation market has offset the overall noise performances of the airports.

Many aviation experts believe that environmental impacts constitute the single most important impediment to the future growth of the air transport industry [7], in addiction EUROCONTROL has recognized that unless airport environmental capacity is effectively managed, environmental issues will form an increasingly significant constraint on European airport development and operations [8]. The IATA Environmental Reporting Survey (Dobbie and Hooper 2001) concluded that for airlines environmental issues are of similar importance to commercial factors such as employment generation, building new alliances, and developing air routes network at their home airport Noise seems not only to represent an annoying consequence of aviation for the people who live nearby the airports but also constitutes a serious threat to aviation itself. The adoption of the EU directive n. 30/2002 embodies the will of the Commission in reducing acoustic pollution and links the exceeding of the noise limits with operating restrictions. This directive adopts the so called ICAO Balanced Approach to noise management that consists in four principal elements and requires a careful assessment of all different options to mitigate noise, including: reduction of airplane noise at source, land use planning and management measures, noise abatement procedures and operating restrictions.

From an airport manager prospective the concept of acoustical capacity grows in importance since it represents the number of aircrafts that an airport can handle within the limits of acoustic zoning. In addiction in a regime of concurrence of airport infrastructures a good way of managing capacity becomes necessary in order to gain market share. Since airport operators cannot deal with the reduction of noise at source and with the land use planning they need to learn how to manage and maximize airport acoustical capacity in
order to avoid operating restrictions. In essence Europe is running out of airport capacity [2] and environmental capacity seem to represent the main constraint.

2. Environmental acoustic capacity

Airport capacity takes several forms: terminals and buildings capacity, runways and aprons capacity, ATM (Air Traffic Management) capacity and environmental capacity. Conventionally the term airport capacity has always been referred to the processing capability of a service facility over a period of time [4], and mainly it has been applied to operational conceptions as number of runways, slots terminals and aprons [13].

European airports have long been perceived as having capacity problems although these have been offset by innovations in air traffic management and by the construction of terminals [2]. By the way of example it’s worth mentioning the cap on movements that the Dutch Government posed in 1997 at Amsterdam Schiphol Airport for the 1998, for environmental reasons, limiting the theoretical capacity of around 650,000 slots per year to only 360,000. At the Airport Council International Meeting (ACI) European Environmental Strategy Committee, held in Brussels in 2002, a total of 15 airport representatives out of 20 reported that environmental capacity constraints already affected the growth of their airports. It emerges clearly that constraints to growth stem from runways and aprons capacity and from environmental capacity and since the last one, in Europe, seems to be more stringent than physical capacity, it emerges the need to investigate environmental capacity and understand how it is affected by variables such as number of movements and time of the day, fleet mix and runway length.

Several definitions are reported in technical literature for airport environmental capacity. On June 2000 environmental capacity was the theme of a workshop held at Heathrow Airport (SCAN-UK 2000) the aim of which was to attempt to agree on a working definition of environmental capacity. The meeting came up with the following definition: “environmental capacity represents the extent to which the environment is able to receive and tolerate, assimilate, or process outputs derived from air activity”.

Thomas (2001), [8] with a more pragmatic approach, gave a definition where environmental capacity is not defined directly, although the environmental component constraining capacity is recognized as: “the impact of airport operations upon the local environment and upon the lives of residents of local communities”. An interesting definition of environmental capacity is the one elaborated by the Amsterdam focus group (2002), composed of the stakeholders of the main airports in Europe, by which environmental capacity is defined as: “the traffic volume possible within environmental load limits, depending on aircraft environmental characteristics (noise, emissions) and time distribution of traffic” [8].

Upham et al. (2001) have defined the environmental capacity of an airport: “as and equated with the capacity of the receiving environment both human and non human, to tolerate the impacts of airport activities”[12].

The literature search confirmed both the scarcity of aviation related works on airport environmental capacity and the importance to redress that deficit.

A different approach is used in defining the environmental acoustical capacity and consists in linking the number of movements with the amount of soil affected by a
particular sound level. In this sense acoustical capacity of an airport is the number of movements per direction of operations that generates a noise pollution that is below the limits imposed by local legislations for the different areas nearby the airports. On the other hand acoustical capacity can be represented by the distance between the runway and the dwellings. Generally speaking the more distance is available the higher number of aircraft movements can be accommodated.

Airport capacity is thus a function of operational scale, management ad environmental constraints [12]. The areas nearby the airports are indeed characterized by different acceptable theoretical limits for noise and those values change in relation with the different types of dwellings: residential and industrial. If the noise generated by air operations oversteps these limits fixed by law, airport operators must take measures in order to decrease noise pollution and a traffic cut off is a likely solution, as a consequence of the Directive number 30 of the 2002 and this is the point that links the exceeding of airport capacity the with a decrease in traffic.

Several factors influence airport acoustical capacity, generally speaking the can be grouped in two categories with respect to the airport environment. Internal factors refer to the airport itself and they can be identified with the geometric characteristics of the infrastructure (runway length), with the characteristics of the traffic (fleet mix, time of the movements) and ATM characteristics. The airport authority, generally has some degree of influence in the internal factors so that it makes possible the likely application of measures to reduce noise pollution modifying such factors. On the other hand external factors refer to the environment is located and, generally, they are difficult to modify in order to gain acoustical capacity. Meteorological conditions such temperature and humidity, the distance between the runway and the suburbs and the position between flight paths and dwellings are some of the external factors.

<table>
<thead>
<tr>
<th>Internal Factors</th>
<th>External Factors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Runway/s length</td>
<td>Meteorological conditions (Humidity)</td>
</tr>
<tr>
<td>Apron and taxiways geometry</td>
<td>Distance runway/s-suburbs</td>
</tr>
<tr>
<td>Airport territory surface</td>
<td>Position of the dwellings</td>
</tr>
<tr>
<td>Traffic</td>
<td>Type of the dwellings</td>
</tr>
<tr>
<td>Types of the carriers (cargo passenger)</td>
<td>Legislative scenario</td>
</tr>
<tr>
<td>Fleet mix</td>
<td></td>
</tr>
<tr>
<td>Load factor</td>
<td></td>
</tr>
<tr>
<td>Time of the day</td>
<td></td>
</tr>
</tbody>
</table>

The adoption of the EU directive 2002 n.49 changes the scenario in most of the European countries introducing the acoustic metric Lden, a unified noise indicator to classified the noise exposure of the territory and to measure the noisiness of the mode of transport. The indicator is based on the A-weighted long-term average sound level as defined in ISO 1996-2: 1987 and is determined over all the day period of the year. The single events have a different weight in the formula to take count of the differences in perceived noise in the different periods of the day; a 10dB penalty is given for the events during the night time and a 5dB penalty for the evening events and this will penalize airport operators limiting the existing airport capacity.
In case of exceeding the limits two types of action can be individuated in order to decrease airport noise pollution: short term mitigations measures and long term measures. Short term mitigation measures are those which allow a reduction of noise exposition in the short term, and they can permit to achieve good results, but they do not guarantee the environmental compatibility in the long run. On the other hand they are less costly and easier to put in practice. Noise restrictions, noise taxes, noise monitoring systems and departure and arrival trajectories designed to avoid the flyover of noise sensitive areas are some of the short term measures widely used to reduce noise pollution around airports.

Long term measures are those which guarantee environmental compatibility also in the long run. They require a much higher amount of investment and a long time realization and can be divided in airport measures as the construction of new runways oriented to minimize the flyover of noise sensitive areas or lengthening of the existing runways so that airplanes flyover communities at an higher altitude; and airline measures such as fleet renewal in order to replace older and noisier aircraft with newer and more quite ones.

3. Case study

A practical case study is conducted in collaboration with the airport of Bologna G. Marconi. The airport that has been chosen represents an interesting case study for several reasons: the traffic and the fleet mix are similar with most of the medium size airport in Europe and the proximity with the city of Bologna makes some areas that lie in the main take off direction noise sensitive.

The fleet mix operating in the 2005 is divided in categories of similar airplanes in regard with noise emissions.

It is evaluated the influence of humidity and temperature in acoustical capacity through the acoustic data measured by the noise monitoring system that the airport G. Marconi is provided.

The relation between number of passengers per single type of aircraft and noise generated is analyzed and it shows a linear trend between the two variables.

It is also studied the influence of fleet mix and the period of the day of the movements in determining acoustical capacity. It is forecasted a traffic growth and it is studied the variation of the capacity over time and it is evaluated the increasing in environmental capacity due to a reduction in the noisier airplane (MD-80 series) replaced by less noisy aircrafts.

A further aspect is analyzed: an increasing in acoustical capacity determined by a lengthening of the runway. An increase in the length allows indeed aircrafts to fly over the dwellings with an higher altitude determining a reduction in noise exposition.

It is examined the case of a runway with a traffic composed of the same aircraft type and it is calculated the number of movements necessary to fill acoustical capacity. It is studied next the same situation but with a longer runway and it is analyzed and discussed the reduction in noise exposition due to the lengthening of the runway.

It is used INM 6.1 (Integrated Noise Model) to simulate the sound level and consequently the noise pollution due to aircraft operations. INM is a widely used tool for such analysis, developed by the Federal Aviation Administration, it requires data input such as types of aircraft and number of movements per period of the day along with data input
for the airports such as runways length and take off and landing procedures and it gives a graphic input representing the noise contours for a particular sound level.

Figure 2. INM simulation of the acoustic climate for a given number and type of aircrafts. The different layers represent a particular sound level measured in Lva.

References


A NEW METHOD FOR OFFSET OPTIMIZATION IN URBAN ROAD NETWORKS

Bernhard Friedrich 1, Essam Almasri 2

Abstract. Despite the fact that manual optimisation approaches represent good engineering practise, their performance is restricted due to complexity of the problem. But also numerical methods suffer from complexity of the NP-hard optimisation problem which is discrete in time and space. Therefore, this study presents a new offset optimisation method employing the Cell Transmission Model, which is a 1st order discrete traffic flow model. The proposed method further particularly cares about optimisation strategies and the respective running time requirements. In this context an optimisation algorithm based on Genetic Algorithm is developed, which uses a decomposition of the problem.

1. Introduction

For traffic signal control, manual optimization approaches (e.g. Schnabel [8]) represent good engineering practise, however, their performance is restricted due to complexity of the problem. Therefore, model-based optimization methods are needed. For that purpose, either equilibrium queuing models or discrete models can be employed. Whereas equilibrium models can only be applied for macroscopic adaptation of split and cycle time at single intersections, discrete models also allow for microscopic adaptation and in particular for offset optimisation.

However, the employment of models, discrete in time and space, is constrained by the run-time requirements of the solution of the NP-hard optimisation problem. In this context the interaction of suitable heuristic optimisation algorithms and simple discrete queuing models is of particular interest. TRANSYT [7], as the first and most well known approach, consequently uses a simple traffic model and a basic solution heuristic. However, none of

1 Institute of Transport, Road Engineering and Planning, University of Hannover, Appelstr. 9A, 30167 Hannover, Germany, friedrich@ivh.uni-hannover.de
2 Institute of Transport, Road Engineering and Planning, University of Hannover, Appelstr. 9A, 30167 Hannover, Germany, almasri@ivh.uni-hannover.de
the known model-based optimization methods—including TRANSYT— is intended for the entire range of the fundamental diagram derived from a flow-fluid analogy (Feldman et al. [3]).

In order to overcome deficiencies of known methods with respect to traffic theory issues, this paper presents a new approach for offset optimisation which is based on the Cell Transmission Model (CTM) which is a 1st order traffic flow model proposed by Daganzo [1] and [2]. Consequently the model is able to predict build-up, propagation and dissipation of queues for both under- and oversaturated conditions. With respect to optimisation strategies and running time requirements, the proposed approach applies two optimisation algorithms based on genetic algorithm (GA).

2. The Cell Transmission Model

2.1. Basic Conditions and Flow Equation

As discretized version of the macroscopic modelling approach the CTM has been developed by Daganzo [1] and [2]. The CTM provides an approximation to the LWR model (abbreviated from Lighthill and Whitham [5]; and Richards [6]) and can be used to predict transient phenomena such as the build-up, propagation and dissipation of queues.

The CTM employs a simplified version of fundamental diagram based on a trapezium form (see Figure 1) assuming that a free-flow speed \( v_f \) at low densities and a backward wave speed \(-w\) for high densities are constant.

![Figure 1. The Flow Density Relationship for the CTM](image)

Dividing the road into homogeneous sections (or cells) and time into intervals of duration \( T \) such that the cell length \( L \) is equal to the distance travelled by free-flowing traffic in one time interval, the LWR results are approximated by a set of recursive equations Daganzo [1] and [2]:

\[
n_j(t+1) = n_j(t) + q_j(t) - q_{j+1}(t)
\]

\[
q_j(t) = \min \left\{ n_{j-1}, Q_j(t), \left( \frac{w}{v_f} \right)(N_j(t) - n_j(t)) \right\}
\]
The subscript \( j \) refers to cell \( j \), and \( (j+1), (j-1) \) represents the cell upstream (downstream) of \( j \) (see Figure 2). The variables \( n_j(t), N_j(t), q_j(t), Q_j(t) \) denote the actual number of vehicles, the maximum number of vehicles (or holding capacity), the actual inflow, and the inflow capacity that can be present in cell \( j \) at time \( t \), respectively. The variable \( q_j(t) \) is also outflow for cells \( j-1 \).

To simulate the effect of a traffic signal, it is assumed that for a signalised cell \( j \) the inflow \( q_j(t) \) is determined by (2) if the signal is green, or is equal to zero if the signal is red.

### 2.2. Delay Calculation

For delay calculation at cell level, an equation is derived based on CTM. In the flow density diagram shown in Figure 1, the slope of the line drawn from the origin represents the actual speed \( v_j(t) \) in cell \( j \) at time \( t \) which is equal to the outflow divided by the density as follows:

\[
v_j(t) = \frac{q_{j+1}(t)}{k_j(t)}
\]  

(3)

where:

\[
k_j(t) = \frac{n_j(t)}{L}
\]  

(4)

The relationship between the delay of one vehicle in cell \( j \) at time \( t \) and the actual speed \( v_j(t) \) is assumed to be linear as shown in Figure 3. Based on this relationship, delay ranges from 0 when the actual speed is equal to the free flow speed \( v_f \) to the duration of \( T \) (time step) when the vehicles are not moving. This delay times the number of vehicles \( n_j(t) \) determines the total delay of all vehicles in one time step in cell \( j \) at time \( t \):

\[
d_j(t) = \left[ T - \left( \frac{T}{v_f} \times v_j(t) \right) \right] \times n_j(t)
\]  

(5)

where:

\[
v_f = \frac{L}{T}
\]  

(6)

By substituting (3), (4), and (6) in (5), the following equation results:

\[
d_j(t) = T \left[ n_j(t) - q_{j+1}(t) \times T \right]
\]  

(7)
Once delay has been calculated at the cell level, the total delay $D_{\text{link}}$ can easily be calculated at the link level by summing up all delays in all cells during the cycle time as follows:

$$D_{\text{link}} = \sum_{i} \sum_{j} d_j(t)$$  \hspace{1cm} (8)

3. Model Validation

To evaluate the delay calculation based on the CTM a virtual environment was prepared by means of microscopic simulation (AIMSUN2) and was used for reference measurements of total delay. The model validation was conducted at a one-way street with two intersections to investigate the delay calculation of the developed model when changing the offsets. The offset at the first intersection was fixed while the offset at the second intersection was varied from 0 to the cycle time (120s) with a step of 10s.

Figure 4 shows the relationship between offset and delay for the analyzed link as estimated by AIMSUN2 and the developed model. The Figure shows that the delay values of the CTM are similar to AIMSUN2. A high regression coefficient $R^2$ of 0.98 was obtained in this comparison.

Figure 4. Delay comparison between CTM AIMSUN2
4. Optimization Algorithms

The objective for the optimization is the minimization of the total network delay, where delay is calculated using the analysis module. The proposed objective function is as follows:

\[ f = \min \sum_i \sum_j d_j(t) \]  \hspace{1cm} (9)

where \( f \) is the sum of delays in all the cells of the network throughout the planning horizon. The objective is to select the offset values for all intersections such that \( f \) is minimized.

Since the objective function has an irregular shape in the solution space and therefore the classical search methods cannot be employed, two heuristic approaches based on genetic algorithm (GA) were developed.

The two approaches are different in their choice of the search direction. In the first approach, referred to in this research as parallel genetic algorithm (PGA), a simultaneous search over all offsets (the entire chromosome) is performed using the reproduction-crossover-mutation process of variation. A convergence criterion or the computing time available ends the search. In the second approach, which is called serial genetic algorithm (SGA), a group of offsets and therefore only a part of the chromosome is varied until the best solution is found. In the next step, the offsets of the next group of intersections are optimized. In a serial search such as this, the order in which the intersections are treated significantly influences the optimization results. A method has been developed for the determination of the search order. Further description of both algorithms will be presented in the paper.

5. Laboratory Test and Method Comparison

A microscopic simulation (AIMSUN2) is used for comparison purposes and in order to compare the new method against the state of the art method TRANSYT-7F [4] and an established manual optimisation approach [8]. Additionally, as the absolute reference, the global optimum was determined by a full enumeration of all variations of offsets.

The comparison was carried out for two case studies, one representing a small real network with 6 intersections; and the other a relatively large grid network consisting of 12 intersections. Applying the new optimization method to the small network, as presented in Table 1, the measured delay is close to the absolute optimum and it outperforms the existing time plan and both the TRANSYT-7F solution and the manual optimisation.
Table 1. Comparison between different methods for the realistic network

<table>
<thead>
<tr>
<th>Method</th>
<th>Mean of 30 replication AIMSUN2 delay (veh.s)</th>
<th>Relative changes %</th>
<th>CPU time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enumeration</td>
<td>38530</td>
<td>0.0%</td>
<td>24 hr</td>
</tr>
<tr>
<td>SGA</td>
<td>38747</td>
<td>0.6%</td>
<td>175 s</td>
</tr>
<tr>
<td>PGA</td>
<td>39923</td>
<td>3.6%</td>
<td>700 s</td>
</tr>
<tr>
<td>TRANSYT-7F</td>
<td>44331</td>
<td>15.1%</td>
<td>1800 s</td>
</tr>
<tr>
<td>Manual</td>
<td>45568</td>
<td>18.3%</td>
<td>-</td>
</tr>
<tr>
<td>Existing</td>
<td>44952</td>
<td>16.7%</td>
<td>-</td>
</tr>
</tbody>
</table>

Figure 5. Delay versus CPU run time of a PC Athlon 1 GHz using different methods for the grid network case study

For the larger network, the full enumeration was impossible even if one has a 10s time step; the calculation time needed for iterating 11 offset variables is estimated to be 18.69 years. Three time steps (1, 2.5 and 5s) are studied in this case study. Figure 5 shows a comparison between the methods cited, and plots the delay (veh.s) versus the CPU run time of a PC Athlon 1 GHz. The results show that the quasi absolute optimum could be found after 230s using the SGA with a 5s interval and after 900s which was not possible using PGA even after 2200s. Therefore, the SGA with the help of the proposed search order determination is applicable, and not only could find the quasi absolute optimum but could also shorten the computation time which allows for online optimization. However, when the developed search order is not used, the SGA failed to find this quasi absolute optimum.

Table 2 presents a comparison between the developed SGA using a 2.5 time step. The results show that a satisfying solution could be found, which is clearly superior to both
TRANSYT-7F (+10% delay) and the manual optimisation (+8 % delay). A more description of the two case studies will be presented in the paper.

<table>
<thead>
<tr>
<th>Method</th>
<th>Mean of 30 replication AIMSUN2 delay (veh.s)</th>
<th>Relative changes %</th>
</tr>
</thead>
<tbody>
<tr>
<td>SGA (2.5s)</td>
<td>157394</td>
<td>0.0%</td>
</tr>
<tr>
<td>Dominance</td>
<td>170202</td>
<td>8.1%</td>
</tr>
<tr>
<td>TRANSYT-7F</td>
<td>172878</td>
<td>9.8%</td>
</tr>
</tbody>
</table>

Table 2: Summary of results of 12-intersection grid network case study

References


VALIDATING NESTED LOGIT MODELS SUBJECT TO PROGNOSIS CONSTRAINTS

Francisco G. Benitez¹, Javier Vazquez¹

Abstract. A formulation of the problem using a nested logit model, with the use of Box-Cox transformations, is presented which allows obtaining in an operative way solutions that fulfill certain prognosis criteria. The classic objective function that uses the maximum likelihood method is completed with some additional terms that include restrictions in the probabilities the model would forecast for variations and perturbations in the values of the attributes defined in the model.

1. The non-linear utility multinomial logit model formulation

The random utility function of the discrete choice scheme assumes that an individual $q$, from a set $N$ confronted to $I$ alternatives, chooses the one with the highest utility. The utility associated with each alternative $i$ and each individual $q$ is represented by the expression:

$$U_{qi} = V(i, q; \Theta, x, \epsilon)$$  \hspace{1cm} (1)

where $x$ is a vector of explanatory variables measured by the analyst, and the components $\Theta$ and $\epsilon$ are parameters which can not be directly observed by the analyst. Assuming the non-observable components are of non-deterministic form (i.e.: random type) distributed iid extreme value, the choice probabilities are given by the very well know logit model [1,5].

Representative utility is usually specified to be linear in parameters, which results in a log-likelihood unimodal function in parameters $\Theta$ guaranteeing a unique local maximum (coincident with the global optimum).

In some contexts it is useful not to limit the functionality of utilities to linearity. For non-linear utility functions estimation of parameters is more difficult, since unimodality of the log-likelihood function is not globally warranted, and non-convex optimization schemes are needed to reach the optimum (the global one, or a quasi-global estimate). The most

¹ Transportation Engineering, Department of Engineering, University of Sevilla. Camino de los Descubrimientos, Sevilla-41092, Spain. Phone: +34-954487315; Fax: +34-954487316; e-mail: benitez@esi.us.es
spread out non-linear utility function comes from considering an additive function of nonlinear multinomial terms of the form:

\[ V^q = \sum_{k=1}^{K} \theta_{jk} \cdot z_{jk}^q \]  

(2)

where parameters \( \theta \) include the distribution parameter \( \beta \), \( z_{jk}^q \) stands for a nonlinear transformation of attribute \( x_{jk}^q \) proposed by Box and Cox [4]. This transformation defines a family of functions that includes the linear and the logarithmic transformation as particular cases. The maximization is based on the general choice model log-likelihood function. Under the above hypothesis, the nonlinear utility logit model is defined by the following problem:

\[
\begin{align*}
\text{Maximize } l(\theta, \tau) &= \sum_{q=1}^{N} \sum_{j=1}^{I} g_j^q \cdot \left[ \sum_{i=1}^{K} \theta_{jk} \cdot z_{jk}^q - \ln \sum_{i=1}^{K} \exp \left( \sum_{k=1}^{K} \theta_{jk} \cdot z_{jk}^q \right) \right] \\
\text{subject to } \\
&\quad z_{jk}^q = \begin{cases} 
(x_{jk}^q)^{\tau_k} - 1, & \text{for } \tau_k \neq 0 \\
\ln x_{jk}^q, & \text{for } \tau_k = 0
\end{cases} \\
&\quad a_k \leq \tau_k \leq b_k \\
&\quad c_k \leq \theta_{jk} \leq t_k \\
&\quad \theta_{ml} = \alpha_{pk} \cdot \theta_{pk}, \quad \forall mp \in l' \subseteq \{1,2,...,l\}, \forall lk \in K' \subseteq \{1,2,...,K\} \\
&\quad g_j^q = \begin{cases} 
1 & \text{if individual } q \text{ chooses alternative } j \\
0 & \text{otherwise}
\end{cases}
\end{align*}
\]  

(3)

where parameter’s constraints have being taken into account through i) upper and lower bounds \( a_k, b_k, c_k \) and \( t_{jk} \) in the inequality constraints, ii) fixed value constraints by equating \( a_k = b_k \) or \( c_k = t_{jk} \) in certain inequality constraints, and iii) linear constraints of form \( \theta_{ml} = \alpha_{pk} \cdot \theta_{pk} \).

2. The predefined prognosis constraint model (PDPC)

The purpose of developing a model is twofold: the first one is to describe disaggregate (to the level of individual) and aggregate (to the level of market share) behaviour, hidden by the huge amount of collected data, by expressions and algorithms easier to manipulate and use; the second and most important one is for the purpose of forecasting. In this last sense using models to predict disaggregate and aggregate effects due to changes in the variables is of uttermost importance. The validity and acceptance of a model depend on its sensitivity to variable change policies. Thus unreasonable forecast leads to model rejection, forcing to
infer a new model. Assuming the data base and decision structure are acceptable, the
modification of parameter bounds permits to the analyst a control of the solution derived
from problem (5). This can be done by constructing the problem, which infers the model, in
such a manner that the expected model prediction share is predefined in its formulation
[2,3]. Following these ideas we can formulate the new problem as:

**Problem (5) subject to additional constraints:**

- relative to prediction shares for certain alternatives:

\[ l_j \leq \frac{1}{N} \sum_{i=1}^{N} \sum_{q=1}^{Q} P^q_{ij} \bigg|_{x_q \cdot \alpha_k} \leq u_j, \quad \forall j \in \{1, 2, ..., I\} \]

\[ \frac{1}{N} \sum_{j=1}^{J} \sum_{i=1}^{I} \sum_{q=1}^{Q} P^q_{ij} \bigg|_{x_q \cdot \alpha_k} = 1 \]

- relative to right prediction share for certain alternatives:

\[ l'_j \leq \frac{1}{N} \sum_{q=1}^{Q} P^q_{ij} \bigg|_{x_q \cdot \alpha_k} \leq u'_j, \quad \forall j \in \{1, 2, ..., I\} \]

with \( \sum_{j=1}^{J} l'_j \leq \sum_{j=1}^{I} u'_j \leq 1 \)

- relative to global right prediction share:

\[ l \leq \frac{1}{N} \sum_{q=1}^{Q} P^q \bigg|_{x_q \cdot \alpha_k} \leq u \]

with \( l \leq u \leq 1 \)

where \( P^q_{ij} \bigg|_{x_q \cdot \alpha_k} \) denotes the probability prediction of individual \( q \) choosing alternative \( j \)

for a change in attribute \( x_k \) given by factor \( \alpha_k \); the rest of scalars stand for allowed
bounds of probability forecast in market segments and global prediction shares. The second
restriction in each set of constraints in (4) assumes the non-violation of fundamental rule of
probability.

For the sake on conciseness no additional constraints are included in formulation (4) as
the relatives to right prediction and global right prediction indexes (which might be more
difficult to define *a priori* by the modeller).

The problem defined above might physical be over-specified as it needs, for its proper
formulation, a very good knowledge by the modeller of the tendencies of market segment
probabilities. In most practical cases the modeller’s knowledge is restricted to global
tendencies, making feasible only the last set of constraints formulated in (4). From now on
we focus our attention in the study of the scheme (3) subject to the last constraint set
formulated in (4).
3. Application example

Data source
To test the methodology, a synthetic data bank on the choices of modes for journeys inside a conurbation generated in a 24 hour time period is generated.

The modal split model chosen has the nesting structure depicted in figure 1. The first level of the model evaluates public transportation –alternative 3- usage, with the private car being the competing mode of transport –alternative 4-. The public transport nest comprises two public mode alternatives, 1 and 2. The variables assumed to influence the alternative choices are individual age, purpose of the journey, time and cost of the trip. The values of these attributes follow a discrete random variable for the purpose, an integer random variable for the age, and a normal distribution $N(\mu, \sigma^2)$ of mean $\mu$ and variance $\sigma^2$, with a restricted ratio $\sigma / \mu \leq 0.25$. The travel time for alternative 3 is always chosen as the minimum of both times for alternatives 1 and 2 of the lower nest. The travel time for alternative 4 was defined as of alternative 3 affected by a random factor in the range (0.25, 0.65) for each observation.

Figure 1. Nested tree

The deterministic parts of the utility functions of the synthetic, generator, model have the following structure:

\[ V_1 = 1.4 + 1.5 \cdot \text{purpose} - 3 \cdot \cos t_1^{0.3} - 0.3 \cdot \text{time}_i^{0.125} - 1.2 \cdot \text{time}_i^{0.125} \]
\[ V_2 = 1.85 \cdot \text{purpose} - 2 \cdot \cos t_1^{0.3} - 0.3 \cdot \text{time}_i^{0.125} - 1.4 \cdot \text{time}_i^{0.125} \]
\[ V_3 = 0.14 + 0.32 \cdot \text{age} - 0.1 \cdot \text{time}_i^{0.1} - 1 \cdot \text{EMU}_{1-2} + 0.5 \cdot \text{EMU}_{1-2} \]
\[ V_4 = -0.12 \cdot \text{age} - 0.9 \cdot \text{time}_i^{0.1} - 1 \cdot \text{EMU}_{1-2} \]

From the observations generated a set of 194 individuals was chosen, all of them with a choice probability predicted by a logit model greater than 95%.
The Hessian of the log-likelihood objective function of (3), evaluated for the observation set of 194 extracted individuals, has negative and positive eigenvalues, characterizing a non-definite negative nature and a non-concave function, which proves the existence of multiple local maxima.

Model specification
The first objective of the model specification is to determine whether the observation set replicates the synthetic model. The estimated model NL, corresponds to the inferred model obtained after the optimization process by applying an Exhaustive Search scheme to ensure that the global optimum is reached. Table 1, lists the results for the model obtained; the t-statistic of significance is given in parenthesis below each coefficient estimate. The performance of this model is compared in terms of their ability to reproduce the observed market share.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Estimated Parameters (t-statistics)</th>
<th>Nesting level 1</th>
<th>Nesting level 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>dummy</td>
<td>0.137256 (-0.23)</td>
<td>-</td>
<td>1.306046 (2.47)</td>
</tr>
<tr>
<td>purpose</td>
<td>-</td>
<td>-</td>
<td>1.386588 (4.47)</td>
</tr>
<tr>
<td>cost</td>
<td>-</td>
<td>-</td>
<td>-3.92600 (19.85)</td>
</tr>
<tr>
<td>time</td>
<td>-0.113857 (-0.61)</td>
<td>-0.887704 (-4.23)</td>
<td>-1.483213 (-8.69)</td>
</tr>
<tr>
<td>age</td>
<td>0.331833 (12.12)</td>
<td>-0.131833 (-4.81)</td>
<td>0.121085 (8.69)</td>
</tr>
<tr>
<td>EMU</td>
<td>0.922406 (22.35)</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Likelihood (at zero) -134.47 -135.33
Likelihood (share) -65.155 -79.99
Likelihood (convergence) -2.73 -0.126
R² (0) 0.980 0.998
R² (share) 0.980 0.998
Likelihood (share) -135.39 -65.098
Likelihood (convergence) -1.01 -16.95
R² (0) 0.992 0.739
R² (share) 0.992 0.739
**PDPC model estimation**

In order to investigate the consequences of adopting a PDPC scheme versus a NL model, we test their performance in terms of their ability to reproduce the observed market shares, and also in terms of their prediction capabilities for hypothetical scenarios.

The scenarios evaluated represent different policy changes which affect to particular changes of attributes involved in the definition of the utility functions. The results obtained by executing the inferred models, for the simulated future scenarios, are compared in terms of market share recovery. Changes ranged from slight to strong policies affecting attributes of time and cost. The scenario modeled by the pre-defined prediction criteria scheme represents a policy change that affects variations of time and cost attributes involved in the definition of the utility functions. This scenario constraints forecast of alternative 1 share, to be at least 35% of the public share –alternative 3- for a decrease of 25% in the net trip time for alternative 2. The model inferred is listed in table 2. The estimation results and expected shares of the NL and PDPC models are listed in table 3. The third row gives the correct expected shares, expressed as percentages. The last row gives the expected share, for each alternative segment. From the results one can observe that the predictions replicate the database and they suggest very acceptable models have been obtained. By comparing NL and PDPC models one observes that the two models have a similarly high level of significance. All models present statistically high values of the goodness-of-fit, although based in rho-squared the PDPC model performs worse than NL model as it was expected since PDPC model is more constrained than NL models.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Estimated Parameters (t-statistics)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Nesting level 1</td>
</tr>
<tr>
<td></td>
<td>Altern. 3</td>
</tr>
<tr>
<td>dummy</td>
<td>0.142674 (0.32)</td>
</tr>
<tr>
<td>purpose</td>
<td>-</td>
</tr>
<tr>
<td>cost</td>
<td>-</td>
</tr>
<tr>
<td>time</td>
<td>-0.107769 (-0.77)</td>
</tr>
<tr>
<td>age</td>
<td>0.299518 (14.95)</td>
</tr>
<tr>
<td>EMU</td>
<td>0.855408 (26.81)</td>
</tr>
</tbody>
</table>

Likelihood (at zero) -134.47 -65.155
Likelihood (share) -135.39 -65.068
Likelihood (convergence) -2.07 -18.59
R²(0) 0.984 0.714
R²(share) 0.984 0.714
Table 3. Reproduction of the market share by NL and PDPC models (NL, PDPC)

<table>
<thead>
<tr>
<th>Observe numbers and predictions</th>
<th>Nesting level 1</th>
<th>Nesting level 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observed</td>
<td>Altern. 4</td>
<td>Altern. 3</td>
</tr>
<tr>
<td></td>
<td>100, 100</td>
<td>94, 94</td>
</tr>
<tr>
<td>Predictions</td>
<td>100, 100</td>
<td>94, 94</td>
</tr>
<tr>
<td>Correct expected share (%)</td>
<td>100, 100</td>
<td>100, 100</td>
</tr>
<tr>
<td>Expected share inside the nesting level</td>
<td>0.520, 0.518</td>
<td>0.480, 0.482</td>
</tr>
</tbody>
</table>

3.1. Prediction results

For most practical purposes and policy implications one would predict the market share of the available alternatives for simulated scenarios. As a matter of exercise three simulated strategies are designed. These strategies are summarized below:

i. Scenario A: 25% reduction in total time and 50% increase in cost for alternative 2.
ii. Scenario B: 25% reduction in cost for alternative 2.
iii. Scenario C: 25% increase in total time cost for alternative 1.

To evaluate the predictions, in each scenario, the data observations were altered to reflect the impact of the strategies. The results corresponding to the expected shares yielded by the different models are pointed by their names and they are summarized in table 4.

Table 4. Expected share yielded by models NL and PDPC

<table>
<thead>
<tr>
<th>scenario</th>
<th>model</th>
<th>Alternative 1</th>
<th>Alternative 2</th>
<th>Alternative 3</th>
<th>Alternative 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>original</td>
<td>NL</td>
<td>0.550</td>
<td>0.450</td>
<td>0.480</td>
<td>0.520</td>
</tr>
<tr>
<td></td>
<td>PDPC</td>
<td>0.527</td>
<td>0.473</td>
<td>0.432</td>
<td>0.518</td>
</tr>
<tr>
<td>A</td>
<td>NL</td>
<td>0.532</td>
<td>0.468</td>
<td>0.432</td>
<td>0.568</td>
</tr>
<tr>
<td></td>
<td>PDPC</td>
<td>0.680</td>
<td>0.320</td>
<td>0.487</td>
<td>0.513</td>
</tr>
<tr>
<td>B</td>
<td>NL</td>
<td>0.203</td>
<td>0.797</td>
<td>0.484</td>
<td>0.513</td>
</tr>
<tr>
<td></td>
<td>PDPC</td>
<td>0.393</td>
<td>0.607</td>
<td>0.487</td>
<td>0.513</td>
</tr>
<tr>
<td>C</td>
<td>NL</td>
<td>0.530</td>
<td>0.470</td>
<td>0.480</td>
<td>0.520</td>
</tr>
<tr>
<td></td>
<td>PDPC</td>
<td>0.483</td>
<td>0.517</td>
<td>0.487</td>
<td>0.513</td>
</tr>
</tbody>
</table>

From the results one can appreciate the fulfillment of the predefined prognosis criteria modeled; the expected share by the PDPC model for alternative 1 under scenario B
surpasses the lower bound of 35% imposed (0.393). For scenario C the expected shares by all models follow the reasonable tendencies of decrease in the use of alternative 1 in parallel to an increase in the competing public transportation mode.

Overall, the results indicate that differences between the inferred models, being negligible from an estimation point of view, may lead to substantial differences in subsequent forecasts.

4. Conclusions

An alternative NL model has been developed and used in research and applications. The model developed is a classic non-linear utility NL model subject to additional constraints. These constraints force the model to provide prognosis proportions inside a predefined range.

Our application demonstrated the feasibility of PDPC estimation. The application showed that the PDPC model produces results of similar level of fitness that a classical non-linear utility NL model but with a controlled forecasting behavior. The PDPC model adds useful flexibility to the family of logit models by providing control of market share forecast.

The results provided by this work recommend the use of this type of constraints and suggests its investigation in other closed-form and open-form discrete-choice models.

References

Abstract. When vehicle tracking technologies are widely in practice, crime investigation services benefit more provided they have certain tools. The paper explores the detection of certain movements of vehicles for two possible types of situations, i.e. was a particular vehicle followed by any vehicle, and did a particular vehicle follow any vehicle. These algorithms assumed that every link in the network is equipped with some sort of vehicle identification or tracking devices and the identities of all vehicles, such as their number plates, are fed into the program. A simulation program was developed to implement the first algorithm, as an example, to visualise the concept introduced.

1. Introduction

The concept of Car Following in traffic engineering deals with analysis of how one vehicle follows the one in front. Basically these theories are based on microscopic modelling of two consecutive vehicles moving in the same lane without overtaking. The speed of the follower and hence its longitudinal location is determined by the stimulus caused by its leader travelling in front. When the leader accelerates/decelerates the follower, too, changes speed to adjust its safe following distance. This adjustment takes place after a time delay which is a function of the reaction time of the following driver and the vehicle’s accelerating/decelerating capability. Most famous car following models are known as GM, Forbes, Pipes and Gipps models. See [1], [2] and [4] for a review of these models. This type of following is random in a sense that the follower-leader pairs are formed randomly without any particular intention. Any vehicle that happens to be travelling in front of the follower becomes its leader. Also, all the existing car following models are interested in microscopic interactions between vehicles travelling in a link. Therefore in simulation studies, after every overtaking or lane changing and after every junction, new leader-
follower pairs are introduced and all the parameters are updated accordingly. The present paper, however, will introduce a different fashion of following, which needs to be distinguished from the former. We will use the term “intentional car following” for the latter, which will be treated in more macroscopic fashion.

Vehicle tracking technologies, such as automatic number plate recognition (ANPR), have been advancing rapidly. In the near future it will be possible to track nearly all the vehicles on a network by means of various methods available at that time. Because of the small percentage of vehicles which are fitted with on-board vehicle identifications devices, ANPR currently continues to dominate the available vehicle tracking technology. ANPR is a mass surveillance method that uses optical character recognition on images to read the licence plates on vehicles. Systems can now scan number plates at around one per second on cars travelling up to 160 km/h [7]. Over the past a few years, vendors, technologies and even number plate designs have evolved considerably. Various companies and governmental organisations have been using ANPR in various areas, such as speed cameras, car parks, gated communities, airport security, traffic monitoring, and CCTV systems. Research has mainly concentrated on the accuracy of the recognition and the technological improvements of these systems (see [6], for a review of these developments). Also, [3] looked at developments of algorithms for optimally locating surveillance technologies with an emphasis on ANP readers by maximizing the benefit that would accrue from measuring travel times on a transportation network. [5] used information gathered from automatic vehicle identification systems to help estimate short-term trip origin–destination (OD) matrices in an urban environment. They proposed a method for using sample link choice proportions and sample OD matrix information derived from AVI data sampled from a portion of vehicles to estimate population OD matrices with these data collection points acting as the origins and the destinations. Based on a comparison of the OD estimate accuracies, with and without these data, they found that using vehicle identification data in OD estimation improves significantly the accuracy of the OD matrix estimates.

In order for a vehicle to follow another one intentionally over a period of time and over a number of links on a network, this following act should possess certain characteristics: (1) The travelled path from an origin to a destination will be the same. (2) The follower will not pass the leading vehicle, even if there are opportunities. (3) If the leader stops for a reason, including traffic lights, the follower will stop too. (4) If the leader follows a non-shortest route, the follower too will take that more expensive route. (5) If the gap between the leader and the follower widens, the follower will speed up to close it. (6) Most of the time both vehicles should be in the same link or the follower should be in the previous link (depending on link length) so that the leader can be seen. If, however, there is a middle link between the leader’s and the follower’s current links it will be impossible to see and follow the leader. (7) When the leader leaves a link by making a turn at a junction, where both left and right turns are possible, in order for the follower to see the leader leave, the follower must be on the link which leads to that particular junction. (8) The distance between the leader and the follower will be short enough, i.e. to be on the same link is not sufficient for long links.
Networks can be classified into two as those that are assumed to be fully covered with vehicle recognition systems and those that are partially covered with these systems. In other words, the former assumes that every link in the network has some sort of vehicle (or number plate) recognition device, meaning that the path of every individual vehicle is known without any missing link. In the latter, the links equipped with these systems will be called “a smart link” and the ones without any recognition camera, etc. will be called “a blind link”. There are certain (major) factors affecting the magnitude of the probability f intentional following. These are: (i) Attraction values of destinations, (ii) Heterogeneous link travel times, (iii) Overtaking opportunities, (iv) Variation in the desired speeds of drivers, (v) Driver information about the network and traffic conditions, (vi) Network size, and (vii) Junction operations. These factors will be detailed in the final paper. Analytical analysis of these factors is almost impossible mainly because of the large number of combinations of these factors and the stochastic nature of some of the factors, such as the desired speed distributions of drivers. Therefore, a basic simulation methodology will be provided to demonstrate the best approach to determine a threshold between random and intentional following for the filtering process.

2. Detection

There are three possible intentional following situations. These are (a) was a particular vehicle followed by any vehicle? (b) did a particular vehicle follow any vehicle, and (c) was there any pair in an intentional following act? Figure 1 shows the flow chart of Type (a), where the following abbreviations are used. This algorithm assumes that the intentional follower, if any, starts this following act on link “a”.

NP = Number plate
Δt = max time span for the intentional follower to lag behind the leader but still able to see the leader
tE = expected travel time from link “a” to link “b” or from link “bn” to “bn+1”
SV = subject vehicle

In Type (a), the identity of one of the two vehicles (i.e. the leader) is known. Whenever this subject vehicle is seen by the system, the detection process is triggered (Box 3 of the flow chart shown in Figure 1). Then a number of vehicles behind the subject vehicle are monitored (Box 4). These vehicles arrive at a link after the arrival of the subject vehicle at t1. The time (t1+Δt) is the cut-off time for recording the number plate of these potential intentional followers. In this paper, we are not interested in the determination of Δt. But its magnitude can be up to 10-20 seconds. Boxes 5 and 6 are for finding the next smart link visited by the subject vehicle (i.e. leader). This is done by checking every link, other than “a”, which are within the proximity of link “a”. In order to reduce the computing time for this checking process, the scan time is limited by (t1+tE+Δt), where tE is the expected travel time between link “a” and this link, “b”, or between the two links that are consecutively visited. In order to ensure this, tE must be long enough to cover the arrival of very slow moving vehicles at link “b”. In Box 8, similar to Box 4, a second group of potential followers are recorded for link “b”. These two groups (pools) are compared to eliminate those vehicles that are not detected on link “a” during (t1+Δt). If this comparison gives no common vehicle, the algorithm is ended, stating that this subject

161
vehicle was not intentionally followed by a vehicle from “a” to “b”. If, however, there are common vehicles, the algorithm moves to the next smart link to carry out the same procedures until all smart links are visited. Box 12 lists these potential intentional followers. It will then be up to the user of the tool to filter this list to distinguish the less and more likely followers.

Figure 1. Flow chart of Type 1 following detection.

In (b), however, the identity of the potential follower (i.e. the subject vehicle) is known. Whenever this subject vehicle is seen by the system, the detection process is triggered. Then a number of vehicles in front of the subject vehicle are monitored. These vehicles arrive at a link before the arrival of the subject vehicle at t1. The time (t1-Δt) is the cut-off time for recording the number plate of these potential leaders. Similar to Type (a), above, we are not interested in the determination of Δt, but its magnitude can be up to 10-20 seconds. Finding the next smart link visited by the subject vehicle is done by checking
every link, other than “a”, which are within the proximity of link “a”. In order to reduce the computing time for this checking process, the scan time is again limited by \((t_1 + t_E - \Delta t)\), where \(t_E\) is the expected travel time between the two links consecutively visited. Then, a second group of potential leaders are recorded for link “b”. These two groups (pools) are compared to eliminate those vehicles that are not detected on link “a” during \((t_1 - \Delta t)\). If this comparison gives no common vehicle, the algorithm is ended, stating that this subject vehicle did not intentionally followed any vehicle from “a” to “b”. If, however, there are common vehicles, the algorithm moves to the next smart link to carry out the same procedures until all smart links are visited. Finally, these potential leaders are listed. It will then be up to the user of the tool to filter this list to distinguish the less and more likely followers.

In (c), both vehicles are not known, hence all the vehicles in the entire population on the network have to be monitored. This is the most difficult detection type as it requires huge amount of resources of time and computing capacity, hence it will not be treated in the present paper.

In small networks, these algorithms can be implemented manually. For example, the user can look at the lists of number plates obtained form few points on the network and check to see if there is any suspicious following. However in large networks, with many ANPR points, a piece of software can quickly identify these potential followers. Therefore we developed a package called IntCarFol and will be introduced at the conference. The program is a deterministic approach and does not simulate the stochastic nature of the phenomenon discussed in Section 2. But further improvements will include this dimension as well. The software is able to read up to 10 smart link files. These files are called SL1.txt to SL10.txt, and stored in the same folder as the IntCarFol.exe file. Each file consists of two sets of artificial data: Column 1 is an array of integers representing number plate readings of vehicles and Column 2 is the time of their arrival (in seconds) at the point where the ANPR camera is assumed to be present. To run the software, the user first chooses the number of smart links that the ANPR data are obtained. Then the number of vehicles in the longest SL file is entered. This is the number of lines in the longest of these .txt files. After that the uniform link lengths in metres are chosen. Then the user enters the number plate of the subject vehicle that might be intentionally followed by another vehicle. When all the Smart Link buttons are pressed the table displays the number plate of each car in the order of their arrival in the first column of each Smart Link and the time of arrival in the second column. Finally, an average speed for the traffic on the network is chosen. When the Run button is clicked, the program does the required calculations outlined in the flow chart given in Figure 1. As an output, the program lists the links that are visited by the subject vehicle and the IDs of all the other vehicles that were always reasonably behind the subject vehicle along this path. The current version of the program makes the following assumptions: (i) the subject vehicle is first seen in Smart Link 1, (ii) all the links are in the same length, (iii) there are maximum of 10 smart links in the network, (iv) all links in the network are smart (full coverage), and (v), the software works on off-line basis, handling of real time will be subject to future research.
3. Concluding remarks and research direction

The paper introduced the issue of potential intentional vehicle following. The stochastic nature of the factors affecting the variability of the phenomenon was also discussed. Three types of detections were introduced and two of them were discussed in detail by providing algorithms. Furthermore, a software package was developed for the Type 1 detection to illustrate the potential possibility of developments to serve this purpose. Main limitations of the software at the moment are that it can handle only a limited number of files (assumed to represent smart links) in an off-line fashion. Also the program does not include most of the factors mentioned in Section 1 and their stochastic effects for simplicity reasons. Further improvements of the software will be able to deal with more links for various detection types, ideally on-line, with more factors involved. It will then become possible, assuming a network covered with these vehicle recognition systems, to capture the registration plates of every vehicle, their link arrival times, routes, OD patterns, etc. This will be a valuable database if these pieces of information are stored continually. However to store such a large amount of data requires huge resources. Thus, more algorithms will be needed, in the light of the discussions of Section 2, to be able to filter the data before their storage. Depending on the user’s preference on the filtering percentage, the system will be able to eliminate probable unintentional following, and store only likely intentional following cases. Another benefit of this study will be to provide increased accuracy to the existing registration plate recognition systems. Although the available registration plate reading technology is quite good, due to the problems like poor visibility, dirt on the plate, alike letters such as D and O, S and 5, etc., image processing algorithms may sometimes make mistakes. However, if the readings are handled in pairs of vehicles, the next camera downstream can be used to verify the reading of the previous camera. This aspect of the research will be subject to a separate publication.

REFERENCES

PATH CHOICE MODELLING FOR FREIGHT ROAD TRANSPORT: A MODEL FOR NATIONAL LEVEL

Francesco RUSSO, Antonino VITETTA, Agata QUATTRONE*

Abstract. In this paper a specification of a path choice behavioural model for freight transport at national level is proposed. The path choice model is simulated in two phases: 1. choice set generation, that is the possible alternatives; 2. path choice among alternatives included in the choice set. The choice set generation is realized with selective multi-criteria path generation, while the path choice is simulated with a Logit and C-Logit model.

1. INTRODUCTION

Freight Transport plays a fundamental role in the economy of every country. The performance of the European transport sector has been in line with the expanding economy. From 1970 to 2000 total European freight transport in the present 15 Member States grew from 1,407,000 to 3,078,000 million tkm (119%). Considering only inland transport, it appears that the considerable growth has been almost entirely realised by road transport (177% Road, -12% Rail, 23% Inland waterways, 33% Pipelines, 169% Sea intra EU).1

In particular, commercial vehicles used for freight transport on road in Italy registered in the 2004 were approximately 3,987,785 units.2

The study, therefore, of the behaviour of truck-drivers’ path choice is very topical. In literature, few behavioural path choice models are specified, calibrated and validated with real experimentation for freight transport on road. The problem of path choice was treated in two phases:

1. the generation of choice set, that is the possible alternatives ([1]; [3]; [14]; [16]);

1 Sources: DG Energy and Transport, Eurostat (pipelines), ECMT, UIC, national statistics.

1 Sea transport 1999: gross estimate subject to revision.

2 Source: ACI – Automobile Club d’Italia

* Mediterranea University of Reggio Calabria, Italy Dept. of Computer Science, Mathematics, Electronics and Transportation

francesco.russo@unirc.it, vitetta@unirc.it, agata.quattrone@unirc.it.

1 Sources: DG Energy and Transport, Eurostat (pipelines), ECMT, UIC, national statistics.

1 Sea transport 1999: gross estimate subject to revision.

2 Source: ACI – Automobile Club d’Italia
2. **PATH CHOICE MODELS**

In the specification of a transport system model the assignment models assume a very important role. They concur to simulate the interaction between supply and demand transport, to calculate the users’ flows and the performances for every element of the supply system. The flows are the result of the choice behaviours of the users who used the transport system and in particular they depend on the path choice in order to move from the zone of origin to the destination. The path choice models supplies the probability of every path from those perceived like admissible.

Given the travel characteristics as origin, destination, trip purpose, time period and mode of travel, the problem of path choice, for a user travelling between origin-destination pair \((o,d)\), consists in characterizing "the better" path with various criteria. The better path
minimize the disutility connected to the trip, that the existing assignment models in literature measure through single attributes, like travel time or distance, or through formulas of the generalized travel cost. The problem of path choice is much complex for the great number of existing alternatives between every \((o,d)\) pair, also on networks of modest dimensions, and their overlapping. These difficulties are treated in literature in various studies ( [3]; [9]; [16]; [17]) in which the complete model specification of path choice is articulated in two phases:

1. the generation of choice set, that is the possible alternatives;
2. the path choice from the alternatives belonging to choice set.

In section 2.1 the factors that influence the path choice behaviour are reported; the paragraphs in the follows treats the path choice models present in literature distinguishing the phase of generation of the alternatives available (section 2.2) from the phase of path choice among the alternatives characterized (section 2.3).

### 2.1. Factors that influence the path choice behaviour

The main factors that influence the users’ behaviour in path choice can be summarized in three categories:

1) users’ knowledge of the network and the available paths
   - users on long distances trips tend to restrict their route to the motorways, because they have insufficient acquaintance of the local roads ([16]; [17]; [20]);
   - users are not always allowed to estimate simple path characteristics, like time or distance, and therefore does not happen as expected to satisfy their objective to choose the faster or shorter path ([6]; [12]; [16]; [17]);
   - users follow a "hierarchical" path that is developed on local roads close to the origin zone, continue on higher level roads for the greater part of the trip and then is concluded newly on local roads [8].

2) decisional process
   - users choose their path before beginning the travel (preventive choice);
   - the decisions on the path to follow are taken during the travel and are independent from those previous (adaptive choice);
   - the decisional process is intermediate to the two previous cases.

3) path attributes and preferences of choice
   - travel time;
   - distance;
   - number of stops and traffic lights regulated intersections;
   - scenic path attraction;
   - safety;
   - commercial development along the path;
   - congestion;
   - quality of the road.
2.2. Generation of choice set

For the generation of path choice set, in literature, different approaches are present (figure 1):

- exhaustive approach, all the analytical paths loop less on the network are available and belonging to choice set that is only one for all the users;
- selective approach, only some available paths represent attractive alternatives of choice.

In the second case, the generation of available paths can be obtained following three different approach: with mono criterion approach, with multi-criteria approach and with probabilistic model of belongings to choice set.

Regarding the mono criterion approach the available paths are obtained by the satisfaction of a single criterion. In order to generate the minimum paths for the criterion is necessary to minimize the disutility connected to the trip measured with a single attribute (time, distance, number of traffic lights regulated intersections, etc.) or to construct a covered function relative to the criterion that is a weighed sum of the travel time and the descriptive attribute, in order to the generated paths do not differ considerably from minimum paths. The parameters of such covered function must be calibrated, as an example maximizing the level of covered of the chosen with generated paths, so the choice set constructed really includes path chosen from the users.

In alternative it is possible to use the multi-criteria approach according to which the available paths are obtained by the satisfaction of some criteria calibrated maximizing the degree of cover of the generated with the chosen paths, in analogous way to the mono criterion approach.

![Figure 1. Approaches for the generation of choice set](image-url)
The main objective of the multi criteria approach is to define choice set, for a specific origin-destination pair \((o,d)\), replacing the great number of physically available paths with few paths, called “label paths” [3], everyone of which represents the optimal respect to a determined criterion. The criteria used in literature ([1]; [3];[16]) for generation of choice set are:

- to minimize the travel time;
- to minimize the distance;
- to maximize the "view" along the path;
- to minimize the number of traffic lights regulated intersections;
- to maximize the motorways use;
- to minimize the use of congestion paths;
- to maximize the presence of commercial areas along the path;
- to travel on high quality roads;
- to travel on high capacity roads;
- to follow a "hierarchical" path that is developed on local roads close to the origin zone, continue on higher level roads for the greater part of the trip and then is concluded newly on local roads;
- to travel on sure roads.

The possible alternatives can also obtained with an approach more rigorous from the behavioural point of view, that demands the specification and calibration of a probabilistic model of belongings to choice set. Regarding the generation with probabilistic model of belongings to choice set it is possible to follow two different approaches:

- implicit approach: the perception/availability of every alternative is simulated with a model that included in the function of systematic utility the model used for path choice. It is assumed that every alternative can have various degrees of perception/availability in the interval \([0,1]\), for which the choice set is represented from a set of continuous variables that supply the degree of inclusion of every alternative in choice set.
- explicit approach: the generation of choice set is realized with an adjunctive model. Such approach is not enough used, because the models that derive from it are not so efficient in computational terms: the numerosity of the choice set exponentially grows with the increase of possible alternatives ([1]; [14]) making much difficult the calculation of the choice probabilities and the calibration of parameters.

### 2.3. Choice model

As regards the path choice, the greater part of the models proposed in literature belongs to the family of random utility model.

Based on assumptions on the random residual of the utility perceived from the users, the models can have different specifications, between which those more used for the path choice are the specification Multinomial Logit, the specification Probit ([4]; [10]; [15]; [16]; [19]) and the Fuzzy model [7].

In the Logit model the random residuals relating to the alternatives are assumed to be independently and identically Gumble variates of zero mean and parameter \(\theta\). This model has the benefit of a closed analitical structure allowing efficient calibration on disaggregate data. Usually the Logit is used in order to simulate the path choice starting from a choice
set generated with the explicit approach and implicit [12] or explicit enumeration of paths ([1]; [3]; [8]).

However the use of a Logit model can carry to choice probability not realistic for paths sharing a number of links ([11],[13]). That depends by the property of independence from the irrelevant alternatives that is at the base of the Logit models and that derives from the assumptions made on random residuals. The Probit model [19], that derives from the assumption that the random utility residuals of the alternatives are distributed as a Multivariate Normal of zero, takes into account the similarity among paths having links in common through the introduction of a random residual covariance proportional to the cost attributes of shared links [11]. That concurs to exceed the disadvantages of the Logit model increasing however the complexity of the problem from the analitical point of view. The Probit model can in fact imply, regarding the Logit, greater difficulties related to the explicit calculation of the choice probabilities.

For the calculation of the path choice has been proposed a modified specification of the MNL, the C-Logit model, that overcomes the main shortcoming of the Logit ([2]; [9]). The C-Logit model takes into account the similarity among the alternatives considering an additional cost attribute, called “commonality factor”, that reduces the systematic utility of a generic path in function of its degree of similarity (or overlapping) with the alternative paths. The C-Logit overcomes the main shortcoming of MNL, i.e. unrealistic choice probabilities for paths sharing a number of links, while keeping a closed analytical structure allowing calibration on disaggregate data and efficient path flow computations when paths are explicitly enumerated.

The Path Size model (PSL) [5] is similar to C-logit in that a correction term, the Path Size (PS) attribute, is added to the deterministic part of the utility. However, PSL has a different theoretical basis. The notion of size comes from the theory of aggregate alternatives, which was first employed for destination and residence choice. However, unlike destination choice, in which zones may have a size representing thousands of elemental destinations (e.g., jobs), the largest size a path may have is one. Such a path shares no links with other paths and may be called a distinct or disjoint path. The path-size term may be calculated on the basis of the length of links in a path and the relative lengths of the paths that share a link.

3. THE PROPOSED MODEL

The proposed model concurs to analyze the behaviour of a sample of truck-drivers in terms of path choice on a road network at national level.

The path choice model is simulated in two phases:
1. choice set generation, that is the possible alternatives;
2. path choice among alternatives included in the choice set.

Concerning to the generation of choice set, the followed selective approach with multi criteria, described in section 2.1, concurs to define for every (o,d) pair a choice set that consist of few paths everyone of which is generated optimizing a covered function associated to a determined criterion (i.e. minimum travel time, maximum motorway use, minimum travel cost etc.). The criteria used are function of the factors that influences the
truck-drivers’ behaviour. The covered function can be calibrated maximizing the degree of overlapping of the chosen paths by the sample with the generated paths.

Regarding the path choice among the alternatives belonging to the generated set the used model is the C-Logit, whose calibration can be carried out with the Maximum Likelihood Method that supplies the values of the parameters on which the utility of every alternative depends, maximizing the probability to observe the choices carried out from the users sample. In section 3.1 the factors that are assumed influence the users sample in the specific case are described; in section 3.2 the generation process of the set of the possible alternatives is defined; in section 3.3 the specification of the path choice model proposed is treats.

3.1. Factors that influence the path choice behaviour

In the phase of choice set generation and therefore of the alternatives available reference to the researches carried out previously in literature with respect to the behaviours assumed from the users in the path choice and the factors that determine them has been made.

In particular, has been taken into account the fact that travel time and distance are not the only determining factors in the choice process, the users often is induced to choose a path rather than from other reasons: the trip cost, the level of road safety (i.e. users can choose a determined path because characterized from low levels of road accidents), the scenic attraction of the path or the commercial development along it (i.e the users can prefer road in which are present service areas, bar and restaurants, etc.), the road quality (a path can be attractive because characterized by low levels of winding etc.).

Moreover user on the long distances is carried to prefer the use of motorways, regarding them in fact is easier that he has a greater amount of information deriving from the personal experience and from the eventual advanced information systems (i.e VMS – Variable Message Signal), regarding the local road network of which has insufficient knowledge.

Finally, the meteorological conditions can influence on the path choice. Users, in particular periods of the year, could in fact prefer coastal and level roads, rather than winding road that cross mountain zones interested by adverse meteorological conditions.

The main factors of which has been taken into account for the generation of choice set are:
- travel time;
- trip cost;
- total length of path and length on motorway;
- scenic attraction of path and commercial development along it;
- road quality;
- road safety;
- meteorological conditions.
3.2. Generation of choice set

3.2.1. Specification

In order to define the path choice set the selective multi criteria approach has been followed. Starting from the factors that have been assumed mainly influence the users behaviour in the dimension of path choice, the criteria have been defined through whose optimization to generate the paths candidates to constitute the set choice. Such criteria ($h$) are list in the follows:

1. minimum travel time;
2. minimum path length;
3. minimum winding;
4. minimum motorway path with bridges and viaducts;
5. maximum motorway path with service areas;
6. maximum motorway path with parking areas;
7. maximum motorway path with bar and restaurants;
8. minimum monetary cost;
9. maximum motorway path;
10. minimum path with high levels of road accidents;
11. minimum path with adverse meteorological conditions.

For all the criteria, relatively to every link $i$ of the network, a covered function equal to the weighed sum of the travel time and an attribute that characterizes the same criterion is defined. The function is dependent from unknown parameters that must to be calibrated in order to maximize the degree of cover of the chosen with the generated paths (label paths).

The general structure characterized for the covered functions for every criterion $h$ is:

$$
t_{ih} = \alpha_h A_i Z_{ih}$$

where

- $A_i$ = attribute of every criterion $h$
- $Z_{ih}$ = parameters calibrated for every criterion $h$
- $\alpha_h$ = parameters calibrated for every criterion $h$

In particular for some of criteria (3–7) the value of such attribute is estimated with the following relation:

$$Z_{ih} = \frac{V_i L_i}{L}$$

where:

- $V_i = \text{number of galleries (} N_{Gall} \text{), bridges and viaducts (} N_{PV} \text{), service areas (} N_{AS} \text{), parking areas (} N_{Parch} \text{) or bars and restaurants (} N_{Rist} \text{), present on the motorway which link } i \text{ belongs, with respect to the type of specific criterion}$
- $L = \text{total length of motorway which the link } i \text{ belongs}$
- $L_i = \text{length of link } i$

In the following the table 1 that describes the parameters relative to the covered functions those have the foretold general structure.
### Table 1. Parameters of the covered functions for the criteria

<table>
<thead>
<tr>
<th>Criterion (h)</th>
<th>( A_i )</th>
<th>( \alpha_{Ah} )</th>
<th>( \alpha_{Sh} )</th>
<th>( Z_{ih} )</th>
<th>( V_{ih} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimum travel time</td>
<td>( t_{0i} )</td>
<td>0</td>
<td>0</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Minimum path length</td>
<td>( L_i )</td>
<td>0</td>
<td>0</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Minimum winding</td>
<td>( t_{0i} )</td>
<td>to calibrate</td>
<td>( V_{ih} ) ( L/L )</td>
<td>( N_{Gall} )</td>
<td></td>
</tr>
<tr>
<td>Minimum motorway path with bridges and viaducts</td>
<td>( t_{0i} )</td>
<td>to calibrate</td>
<td>( V_{ih} ) ( L/L )</td>
<td>( N_{PV} )</td>
<td></td>
</tr>
<tr>
<td>Maximum motorway path with service areas</td>
<td>( t_{0i} )</td>
<td>to calibrate</td>
<td>( V_{ih} ) ( L/L )</td>
<td>( N_{AS} )</td>
<td></td>
</tr>
<tr>
<td>Maximum motorway path with parking areas</td>
<td>( t_{0i} )</td>
<td>to calibrate</td>
<td>( V_{ih} ) ( L/L )</td>
<td>( N_{Parch} )</td>
<td></td>
</tr>
<tr>
<td>Maximum motorway path with bar and restaurants</td>
<td>( t_{0i} )</td>
<td>to calibrate</td>
<td>( V_{ih} ) ( L/L )</td>
<td>( N_{Rist} )</td>
<td></td>
</tr>
<tr>
<td>Minimum monetary cost</td>
<td>( t_{0i} )</td>
<td>to calibrate</td>
<td>( C_{i,Carb} + Pedi )</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Maximum motorway path</td>
<td>0</td>
<td>1</td>
<td>( t_{0i} )</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Minimum path with high levels of road accidents</td>
<td>( t_{0i} )</td>
<td>to calibrate</td>
<td>( t_{0i} I_i )</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Minimum path with adverse meteorological conditions</td>
<td>0</td>
<td>0</td>
<td>to calibrate*</td>
<td>( t_{0i} )</td>
<td>-</td>
</tr>
</tbody>
</table>

*Note:*

\( \alpha_{Sh} = 1 \) for every link \( i \), with the exception of some links of the networks that cross mountain zones for which a much elevating time has been considered, placing \( \alpha_{Sh} = 100 \)

\( t_{0i} = \) travel time carried out by using functions reported in the literature. The functions proposed by the Italian National Research Council (1983) were used for motorways, while TRRL functions (1980) were used for extra-urban roads.

\( Pedi = \) monetary cost of toll estimated as product of the link’s length for the unitary specific cost for kilometre assumed equal to 0.05 €/km.

\( C_{i,Carb} = \) cost for the fuel consumption (Russo, 2005).

\( I_i = \) number of incidents taken place in year 2001\(^3\) (is the number of incidents for kilometre divided to the kilometric extended of the road which the link \( i \) belongs).

### 3.3. Choice Model

#### 3.3.1. Specification

The path choice is simulated with a C-Logit model (Cascetta et al., 1996).

---

\(^3\) Source: ACI – Automobile Club d’Italia
A generic user $n$, travelling between an origin-destination pair $(o,d)$, associates to each path $k$ belonging to the set $I_{od}$ of available paths connecting that $(o,d)$ pair a perceived utility $U_k^n$ which may be expressed as:

$$U_k^n = V_k^n + \varepsilon_k^n \quad \forall k \in I_{od}$$

where

$V_k^n$ is the systematic utility of path $k$

$\varepsilon_k^n$ is the random residual usually assumed to include perception errors of the decision makes as well as modelling approximation of the analyst.

The C-Logit model, with respect to Multinom ial Logit model, introduces a modified systematic disutility as:

$$\tilde{V}_k^n = V_k^n - CF_k \quad \forall k \in I_{od}$$

and consequently the path choice probabilities $p_{od}(k/n)$ can be expressed as:

$$p_{od}(k/n) = \frac{\exp[V_k^n - CF_k]}{\sum_{h \in I_{od}} \exp[V_h^n - CF_h]}$$

The “Commonality Factor” $CF_k$ of path $k$, is directly proportional to the degree of similarity (or overlapping) of path $k$ with other paths belonging to $I_{od}$.

Heavily overlapping paths have larger commonality factors and thus a smaller systematic utility (larger generalised cost) with respect to similar, but independent paths.

The commonality factor can be specified in different ways, giving rise to different path choice probabilities and, ultimately, different C-Logit model specifications.

The general structure of the commonality factor is:

$$CF_k = \beta_0 \cdot \phi(L_{hk}, L_h, L_k)$$

where

$\beta_0$ is a calibration parameter

$L_{hk}$ is the “length” (generalized cost) of links common to paths $h$ and $k$

$L_h$ and $L_k$ are the overall “lengths” (sum of link lengths) of paths $h$ and $k$ respectively

$\phi$ is a continue and monotone function that increases with $L_{hk}$ and decreases with $L_h$ and $L_k$

One possible way to specify the commonality factor is as follows:

$$CF_k = \beta_0 \ln \sum_{i \in k} w_{ik} N_i$$

where

$w_{ik}$ is the proportional weight of link $i$ for path $k$

$N_i$ is the number of paths, connecting the same $(o,d)$ pair, which share the link $i$.
Coefficients $w_{ik}$ can be specified in different ways expressing different hypotheses on the perceived relevance of an individual link in a path. One possibility is to assume $w_{ik}$ as the fraction of total path “travel cost” which can be attributed to each link $i$.

Terms $N_i$ can be seen as the summation of the 0/1 elements of the link-path incidence matrix relative to link $i$ and to all paths connecting the $(o,d)$ pair: $N_i = \sum_{h \in I_{od}} a_{ih}$

3.3.2. Calibration

The calibration of the model consists in obtaining the estimation of the parameters on which the same model depends starting from the choices carried out from the users’ sample.

The C-Logit model can be calibrated with the classic method of ML - Maximum Likelihood, (Ben Akiva and Lerman, 1985). Such method supplies the values of the unknown parameters that maximize the probability to observe the choices carried out from the users.

If link “lengths” or weights used in $CF_k$ specifications, are expressed through link generalised costs including unknown coefficients, the utility function is non-linear with respect to coefficients, and specialised ML models and algorithms have to be used.

4. CONCLUSIONS

In this paper the problem of truck-drivers’ path choice at national-scale transport was treated. In particular, path choice behavioural models were specified for road transport systems at national level.

Some preliminary calibration, that supplied some valid results and provide suitable indication on the user path decisional process in the extra-urban context, were carried out on a truck-drivers road-side survey. The complete results will be reported in a succeeding paper. In the future, in order to get more information about users’ path choice behaviour, the vehicles of some truck-drivers will be equipped with on board Intelligent Transportation Systems (ITS) for monitoring good transport. The utilization of ITS supplies new real time data about path choice. The database obtained by truck-drivers’ road-side survey will be updated with these new data. Finally, the proposed model and its ulterior advances will be compared with other similar models (i.e. path size models).

References


EXPLORING A GUIDANCE PROCESS FOR USING THE PUBLIC ADMINISTRATION GENOME: AN APPLICATION FOR TRANSPORTATION IN A NATURAL DISASTER SITUATION

John Dickey1

Abstract. The Public Administration (P. A.) Genome Project (PAGP) is a long term effort to map the set of P. A. “genes” (topics or variables) and relationships, including those in public sector transportation management. To date, about 11,000 variables and somewhat more relationships have been identified. Presented here is a first exploration of a process for using the Genome for guiding problem and strategy elaboration. The process is illustrated with an example for transportation in a natural disaster situation.

1. The Public Administration Genome Project

Work has progressed over the last six years on an endeavor known as the Public Administration Genome Project (PAGP). The basic premise is that public administration (including public sector transportation administration) behavior bears much similarity to the functions of genes in the human body. Further, since there are about 30,000 genes in each of the trillions of cells in the human body, there must be at least that many “genes” (referred to as single word “topics”) in public administration. The long term goal for the PAGP thus is to “map” these topics and clusters thereof (that is, “variables”), as well as the relationships between them. The result should be a better understanding of public administrative actors and actions.

It is not the purpose of this paper to describe the PAGP in depth. The interested reader is referred to the relevant website:

http://www.cpap.vt.edu/cyberquest/pag genome/

It should be mentioned, however, that the current total number of topics (P. A. “genes”) is over 5000; variables over 11,000; and bivariate relationships also over 11,000. All of these come from 46 sources, including textbooks, case studies, and analogies to the biochemical and genetic world.

1 Urban Affairs and Planning Department, Virginia Tech, Blacksburg, VA 24061-0520, USA, jdickey@vt.edu
2. The COMPASS Information and Guidance System

Although the PAGP still is far short of the goal of having a map of at least 30,000 topics, it appears that there is enough of a base to start thinking about the question of how the above-mentioned kinds of data might be employed most usefully. The purpose of this paper thus is to present a first concept of an information and guidance system which we are calling “COMPASS”.

As the name suggests, the system is intended as a steering assistance mechanism. It has some of the characteristics of an expert system but, as will be seen, the objective is much broader: to alert the public administrator to a wide range of strategy and impact considerations. These can be both internal to that administrator’s office and external. For example, the administrator may be contemplating a new transportation/communication system, say one to send a wireless message to all cell phones in the vicinity of a traffic accident. In addition to the technical difficulties involved, COMPASS may alert the user to be aware that fellow workers in his/her office may be opposed to the idea philosophically, and thus be uncooperative, as may be localities nearby who would be the recipient of additional diverted traffic.

3. The Predominant Case: Three Hurricanes

To help build the P. A. Genome (and subsequently COMPASS), we have been following the actions and impacts of three of the most devastating hurricanes ever in the United States – Katrina, Rita, and Wilma – along with the responses to them. As of this moment we have identified over 1400 variables, almost entirely from the newspapers, TV news networks, and interviews with two victims of the storms. Despite this relatively large number, the tally is not close to being a complete list, and only a relatively few (900) of the potential 1000’s of bivariate relationships have been documented. In any case, we have entered all these variables and relationships into COMPASS, and divided the former into 24 categories.

4. The General Search Process

At this point we have created a causal network linking most of the variables in the COMPASS system. Let us imagine that the variables are divided into four categories:
   - Strategies (options, decisions, alternatives, etc.)
   - Impacts (problems, goals, etc.)
   - Externals (exogenous or uncontrollable variables, etc.)
   - Intermediates (endogenous or internal variables, etc.)

Further imagine that the network has been laid out such that the strategy and external variables are at the top; the impacts are at the bottom; and the intermediates are in the middle.
The objective is to identify a set of strategies (which in actuality may number in the hundreds) for a given central aim or problem that addresses all the identified important impacts (which also may number in the hundreds).

A seven step, iterative process is involved.

1. Identify the aim (problem).
2. Pick the main topic ("gene") from this statement.
3. Do a search for variables that deal with the main topic
   (some of these variables will be strategies; some impacts).
4. Evaluate the strategies.
5. Pick an important strategy and follow it “down” the network to identify its impacts.
   Alternately, look for impacts within a particular category of variables (e.g., social, political, or finance).
6. Evaluate the impacts.
7. Pick an important impact and follow it “up” the network to identify associated strategies.
   The process continues until all important impacts are considered and there are corresponding strategies to address them.

The evaluations mentioned in step 6 above can be done very quickly and subjectively, especially when there are many alternatives, or else very formally and objectively, particularly on important and controversial items. The formal approach would involve answering questions like:

- Is the impact our responsibility?
  - If not, is it being addressed by another office/agency?
- Have we implemented a strategy for the impact before?
  - Currently being tried.
  - Tried before, unsuccessfully.
  - Tried before, successfully.
- If not, is the proposed strategy:
  - Definitely infeasible.
  - Requiring of more thought.
  - Requiring of more information.
  - Implementable, but only in a contingency situation.
  - Implementable immediately?

A decision then is made on whether the strategy is a “must” or a “want”. The former is one that is required (e.g., by statute or executive order). The latter does not have to be done, in which case the strategy subsequently is rated in terms of “importance” and “ease of implementation”.

5. Example: Transportation in a Hurricane Evacuation Effort

To illustrate the preceding process, let us imagine that the potential COMPASS user is a team working for a state department of transportation (SDOT) and responsible for overseeing a smooth evacuation process in front of an on-coming hurricane. The main focus is on resolving traffic jams (there was one 100 miles long from Houston to Dallas,
Texas before Hurricane Rita!). So the team does not want to be taken by surprise by some unforeseen factors contributing to, or caused by, a traffic jam.

The first step is to identify a main aim, which is taken to be:

“Reduce Traffic Congestion during the Evacuation”.

The next step (2) is to pick out the most central gene or topic in the aim statement (say, “traffic” or “evacuation”). Use of the former in a search (step 3) of COMPASS, results in 36 matches, which include:

- BOTF  Develop Ways to Optimize Traffic Flow
- CITME  Cater to Increases in Traffic
- EFMEN  Effective Traffic Management/Enforcement
- ETMRF  Electronic Traffic Management/Road Pricing
- FPGT   Fighting Pollution from Growing Traffic
- TRAFJAM  Traffic Jams on Evacuation Routes

(the first set of capital letters is a short identification code).

These representations of variables, although crude, still can be very useful in generating a whole set of thoughts. The first one, for instance, might engender the idea of setting traffic signal patterns on arterials to increase outward flow; the second of giving priority to buses and cars with more passengers; the third to making sure that police and SDOT workers are present to help (and do not evacuate themselves). The “Road Pricing” one is interesting, especially if a toll-road is involved. Should the toll be lifted temporarily or not?

The “Pollution” one (FGPT) also has some interesting ramifications. If there is a traffic jam, presumably roadside pollution levels will be higher, which may be caught by the monitors, which may contribute to Federal financial penalties. Should the monitors be turned off temporarily? Such a possibility (strategy) perhaps should be evaluated (step 4) using the formal methodology presented in the preceding section.

To continue the illustration, the last variable on the search results list, TRAFJAM, obviously is very close to the original aim statement and deserves more exploration (step 5) for both associated impacts and strategies. COMPASS can be employed to trace the linkages (if any) to variables within each of the 36 categories. On the impact side, for example, suppose the team looked within the “Finance” category (alternate step 5) and were quite surprised to find the related impact of:

$BANKAC  People Cannot Get Money Out of Bank

What is the possible connection to traffic jams? COMPASS could be employed to trace through the logic (step 5 again):

TRAFJAM  Traffic Jams on Evacuation Routes

INFLUENCE:

ATMRE  Bank Trucks Cannot Get Through to ATMs

WHICH INFLUENCES:

$BANKAC  People Cannot Get Money Out of Bank

---

181
This impact can be evaluated (step 6) using the methodology in the preceding section. Chances are that the impact would be considered to be beyond the scope of both SDOT’s statutory charge and capabilities. But SDOT may want to cooperate in whatever strategy the responsible agency develops.

In step 7 COMPASS can be employed to help identify some general strategies for reducing traffic jams (and subsequently their negative impacts). One strategy, for instance, is to reverse some lanes on that part of the expressway leading into the urban area from which the evacuation is being made. This rather obvious strategy is found in COMPASS by working “upwards”) from the variable in focus (TRAFJAM). The system subsequently presents the logic in this situation as:

```
THE VARIABLE:
============================================================================
REVLANE  Reverse Some Lanes on Inbound Freeway =
============================================================================
ALONG WITH:
============================================================================
ACCIDNT  Accidents on Roadways =
CAROWN   Motor Vehicle Ownership =
DAMAGEB  Damage to Highways/Bridges =
============================================================================
INFLUENCE:
============================================================================
TRAFJAM  Traffic Jams on Evacuation Routes =
============================================================================
```

A point to notice about this chart is that there can be many other factors that work in concert with the strategy under consideration to influence traffic jams. If, as one illustration, motor vehicle ownership is high, then, all other things being equal, there will be more vehicles to evacuate. On the other hand, if fewer people want to leave, there obviously will be less traffic.

6. Closing Thoughts

Through the process described, COMPASS is able to prompt the user to help produce a set of variables which, with some amount of translation, can be reformulated into relevant impacts and/or strategies. The process as it currently stands still depends on the experience and intuition of the user. This assumedly will lessen as the Genome on which COMPASS is based expands in the future and we gain more familiarity with the ways in which the system is utilized.
AN ADVANCED METHODOLOGICAL APPROACH FOR PERFORMANCE EVALUATIONS OF ALTERNATIVE FUNCTIONAL STRUCTURES OF A RO-RO TERMINAL

Domenico GATTUSO¹, Antonella POLIMENI², Paolo IARIA³

Abstract. This work focuses on the importance of developing an advanced methodological approach, based on the logical programming, in order to optimize the performances of a Ro-Ro port terminal, by evaluating different alternative functional structures.

1. Introduction

In order to increase sea transport competitiveness, it is necessary to improve maritime services not only “en route” (in terms of performances and costs), but also at the terminals where users can spend a great part of their total travel time. Storage areas of very receptive ports, which are able, during normal traffic conditions, to hold all the vehicles waiting for the loading operations, are seldom insufficient for certain situations like temporary demand excesses or capacity reductions. This involves considerable disadvantages not only for the service users, who have to face longer waiting times, but also for the community because, if these traffic conditions last a long time, the vehicles are forced to queue up on the access links to the terminal. For this reason, this work focuses on the importance of using modelling tools, that can simulate the operational aspects of these nerve centres of the network for different traffic conditions.

2. Modelling tools: the state of the art

The queueing problem at a gate is often studied by using the queueing theory, in order to analyse accurately a system in some specific nodes (Kleinrock, 1975). According to this theory, queue behaviour depends, first of all, on the user arrival rate distribution at the gate (ticket offices, loading points, etc.), on the service time (time required for payment of the fare, loading operations,

---

¹ University Mediterranea of Reggio Calabria (IT), DIMET, gattuso@ing.unirc.it
² University Mediterranea of Reggio Calabria (IT), DIMET, antonella.polimeni@ing.unirc.it
³ Consultant, paoloiaria@ilibero.it
etc.) and on queue discipline. The queueing theory is based on both stochastic and deterministic models. The first kind of model is used for systems in which the average service time is given; these models assume that the average arrival flow is always lower than the system capacity. On the other hand, the deterministic model can be used when, for a given reference period, the flow entering the system is greater than the system capacity (over-saturation conditions). Even though it is greatly used, the queueing theory is a static tool and does not allow dynamic simulation of the interaction between transport demand and supply.

In the same way, using a macroscopic approach for the simulation of a node does not allow the extraction of microscopic indicators, but only of global information.

Nowadays, in order to get round these difficulties, the use of micro-simulators (AA.VV., 1997; Barcelo, 2004; Ferrer J.L., 1996; PTV AG, 2003) to analyse the vehicle flow by representing dynamically the behaviour of the single vehicle is becoming widespread. On the market, there are many packages for analysing, simulating and evaluating different traffic scenarios, both in urban and extra-urban contexts, by representing in detail complex situations and providing macro and micro indicators of efficiency and effectiveness of the system. Even though these packages often use a logical programming, it is often difficult to combine them with ITS systems, as the dynamic logic has to be programmed and calibrated for each analysis context. A methodological approach based on a logical programming for the evaluation of Ro-Ro terminal performances is proposed below.

3. Proposed methodological approach

This work proposes an advanced methodological approach in order to evaluate the performances of alternative functional structures of a Ro-Ro terminal, from both the operator and the user points of view. The aim is, first of all, the estimate of the terminal capacity both on sea (depending on the loading/unloading operations) and on land (depending on the capacity of the storage areas and the ticket offices). If these elements are connected in series (fig. 1), as usual, the terminal capacity \( C \) can be calculated as:

\[
C = \min \{ C_s; C_t; C_m \}
\]

![Figure 1. Terminal capacity.](image)

Regarding the performances, this approach, based on the use of a micro-simulator, consists of a modelling representation of the transport supply and, then, of the analysis of the maritime transport, storage areas, ticket offices, system accessibility. To this end, this work presents, as a support system for microscopic analysis, a model for the topological-functional representation of
Regarding demand, it is fundamental to know not only the number of users, but also the temporal distribution of their arrivals, for which a Poissonian distribution has been chosen. It is also useful to formulate an assumption about user behaviour; in this research a psycho-physical driver behaviour model has been chosen (Wiedemann, 1974). The basic concept of this model is that the driver of a faster moving vehicle starts to decelerate as he reaches his individual perception threshold relative to a slower moving vehicle. Since he cannot determine the speed of that vehicle exactly, his speed will fall below that vehicle’s speed until he starts to accelerate again after reaching another perception threshold. This results in an iterative process of acceleration and deceleration, in a stochastic distribution of the travel speed (even though it can be fixed for some areas, like the loading/unloading points) and, therefore, in a realistic driving behaviour.

At this point, any other standard approach, after defining the supply and demand systems, would simulate their interaction. The proposed procedure provides for an intermediate step which translates, through logical programming, the management and control of some important activities which can influence the functionality and the performances of the terminal, as well as the dynamicity of the users’ behaviour and their path choices (fig. 2). By means of mathematical models, this logical programming allows dynamic management (in relation to the flows) of the conflict areas, access links, storage areas, communication between different infrastructural components, between vehicles and infrastructure, between different vehicles, and between vehicles and drivers.

To achieve this, the procedure proposes two different approaches: a logical implementation of some priority rules for non-signalised intersections and a logical implementation of coordinated
and actuated traffic signals. The former manages the conflict areas by introducing temporal ($h^*$) and space ($d^*$) gaps. A vehicle coming from a minor road can cross a conflict area only if it has enough time ($h$) and space ($d$) at its disposal:

$$\begin{align*}
    h &> h^* \\
    d &> d^*
\end{align*}$$

Concerning system management by means of traffic signals, fixed or variable (in relation to the flows obtained through ITS systems, based on loops, video cameras, push buttons, track circuits etc.) signal phases or cycle times can be defined.

After defining the centre of the procedure (and therefore the logical programming), the demand/supply interaction has to be simulated by means of a dynamic assignment model in which users choose their path according to the random utility theory:

$$U_k = 1/C_k = 1/\sum_{l} c_l$$

$$p(k) = U_k^\theta / \sum_{k'} (U_{k'}^\theta)$$

where $U_k$ is the utility of route $k$, $C_k$ the cost of route $k$, $c_l$ the general cost of link $l$ belonging to route $k$, $p(k)$ the probability of route $k$ being chosen and $\theta$ the sensitivity factor of the model ($\theta > 0$).

Finally, in order to evaluate and compare different scenarios, a set of macro and micro indicators is suggested, such as the number of served, delayed and refused vehicles, the total and average travel time, the total and average delay, the vehicles’ average speed.

4. Case study

The proposed procedure has been tested for the case study of the new Ro-Ro port terminal of Tremestieri (Messina, Italy), built to relieve freight traffic congestion in the city of Messina, in order to evaluate “ex-ante” its functionality, in relation to the number of the maritime companies operating at the port (C.I.Su.T., 2005).

As usual, the study starts from the definition of the actual supply and demand models; then, the attention is focused on the design of the alternative scenarios.

Planning of the simulation scenarios has been the most delicate step; they have been planned in relation to the following system variables: number of maritime operators, entering flows differentiated from one operator to another, vessel dwelling time for the loading/unloading operations, service frequencies, access links control system, logistic management of the wharfs, number of ticket offices, wharfs and storage areas managed by each operator (in proportion to the attracted demand). Regarding space organization, it has been decided that storage areas, ticket offices and wharfs cannot be shared by competing companies, being, therefore, for the exclusive use of each operator. This has required the logical programming of the access links control system, in order to meet the operators’ requirements (sprung from an apposite need analysis, carried out by means of a survey on the local maritime operators’ needs,) to separate constantly and physically the demand flows attracted by each operator, avoiding sudden changes of decision by users about the chosen company. The access links control system allows a continuous feed-back process between different areas of the port terminal; by using a “pulsating” system, the areas close to the wharfs can communicate with the storage areas, telling them if and when the stocked vehicles can be served, in relation to the number of vehicles and the service availability. Moreover, it has been useful to introduce a logical system in order to simulate the vehicles access/egress operations.
into/out of the vessels, by using tools simulating the priority rules (entering vehicles have to give way to vehicles getting off), without changing the assumptions about user behavior.

In detail, the basic assumptions are the following:
- three different supply scenarios, with one (scenario “a”), two (scenario “b”) and three (scenario “c”) competing maritime companies respectively. To each operator, his own storage areas, ticket offices and wharfs are associated (fig. 3); in case of more than one operator, it has been supposed that each maritime company attracts a share of demand, proportional to the actual attracted demand in the Straits of Messina;
- three different demand scenarios, with entering flows equal to $Q_1=88$ (measured entering flow), $Q_2=176$ (twice the measured entering flow) and $Q_3=440$ (five times the measured entering flow), in order to simulate both ordinary and extra-ordinary traffic conditions.

Combining these assumptions, the analyzed scenarios are summarized in figure 4, in which $\phi_1$, $\phi_2$ and $\phi_3$ are the frequencies of the three operators and A.C.S. is the Access Control System.
The results of the application of the proposed procedure on the case study of Tremestieri port terminal are represented in the following figures (fig. 5, 6, 7).

In particular, figure 5 represents:
- the number of served vehicles, i.e. the number of vehicles which can use maritime transport service during the reference period;
- the delayed vehicles, i.e. the number of vehicles which, although inside the system, cannot be loaded during the simulation period because of the maritime service capacity;
- the refused vehicles, i.e. the number of vehicles which are forced to queue up on the access links, because of storage area capacity.

As one can see, if the number of operators increases, the served number of vehicles decreases, because of the reduction in service frequency. On the other hand, the reduction of the storage areas implies that, if the number of operators increases, the delayed and, above all, the refused vehicles increase exponentially. It is worth underlining that the longer the same entering flow conditions last, the greater the number of delayed and refused vehicles is.

Figure 5. Served, delayed and refused vehicles.

In figure 6, the trends of the users’ total and average travel time on the network are reported, by varying the entering flow and the number of operators.

Figure 6. Total and average travel time.
Both of the two curves show that the most critical condition is the one in which all three maritime companies are operative; the comparison between scenarios 1 and 2 is less immediate. In particular, the figure on the left shows that the total travel time, when just one operator is present, is greater than the total travel time in the presence of two operators. This doesn’t mean that the single user spends a longer time in the first case (as one can see in the figure on the right) but that in the mono-operator case, the number of served vehicles is greater than in the two-operator case (fig. 5) and, therefore, also the total travel time (product of the number of served vehicles and single users’ travel time).

Another important indicator of the system functionality is the user waste of time, because of the interferences due to the interaction with other vehicles, to the traffic signs and to any other impedance factor. The trend of this indicator (fig. 7) is similar to the trend of the user travel time and, for this reason, the same consideration can be made.

Finally, figure 8 shows the trend of the average speed of the vehicles, by varying the number of operators and the entering flow. The figure shows, on the one hand, a decreasing trend of the speed by increasing the entering flow and, on the other hand, an increasing trend by reducing the number of involved operators.
In conclusion, after comparing the results, the best solution for the maximum functionality of the system is the presence of just one operator.

5. Conclusions

The main aim of this work is to focus on the importance of using an advanced approach based on linear programming in order to dynamically simulate the functionality of a Ro-Ro terminal and, therefore, to choose the most suitable condition of the supplied service, by considering both ordinary and extra-ordinary traffic conditions.

References


EXPERIMENTAL RESEARCH FOR THE EVALUATION OF DRIVER SPEED BEHAVIOUR ON TWO-LANE RURAL ROADS

Gianluca DELL’ACQUA and Domenico ABATE

Abstract. Numerous studies have been conducted, in different Countries and on various typologies of roads, to determine, on the base of the statistic analysis of instrumental direct measurements, the links among the operating speeds and some characteristic parameters of the roads. The objective of this experimental study is to get relationships of general validity for the prediction, in phase of project, of the speeds that will be indeed practiced on the infrastructure. The experimental investigation, object of the present paper, has been conducted employing traffic counters, able to record, for every vehicular passage in both senses: length, instant speed and direction of the vehicle. The plan of survey has been elaborated to satisfy different objectives of search (speed in free flow conditions, in entry and in gone out of the intersections, in rural to urban transition, etc.) and it has been applied to some belonging roads to the network of the Province of Salerno. The measures have been performed holding under observation every section for 2 or 3 hours. Then the data have been used for starting the formalization of some analytical relationships to predict the operating speeds.

1. Objectives of the study

The instrumental survey of the road functional characteristics is implied essential in the choice of the safety interventions. Moreover, the individualization of the criticises on a limited number of roads allows to draw considerations applicable to the improvement of the offered safety on all the other network segments; in fact, the elaboration of suitable methodologies of intervention comes down from the recognition of the factors producing risk situations, whose elimination

---

1 University of Naples “Federico II”, Department of Transportation Engineering, via Claudio 21, gianluca.dellacqua@unina.it
2 University of Naples “Federico II”, Department of Transportation Engineering, via Claudio 21, domenico.abate@unina.it
guarantees the attainment of performance levels more elevated in a larger context as long as homogeneous on the driver behavior.

During the years 2003 and 2004, about 60 infrastructures, belonging to the Provincial territory of Salerno but not necessarily managed from the Province, have been observed by the authors.

The relief data during these first investigations, crossed with those related to the road accidents, have been employed for constituting the initial nucleus of the informative base of the Road Safety Office of the Province of Salerno; the activities of monitoring we have revealed useful also profits for the standardization of the experimental procedures of data acquisition and for the way of the following investigation survey. The infrastructures have been observed for the two years 2003-04 and they constitute a representative sample of the whole rural viability in the Provincial territory of Salerno; they are roads with unique carriageway and two lane, situated primarily in rural contexts, morphologically level or hilly terrain. The information coming from the first monitoring survey have allowed to order the roads for homogeneous territorial environment and in function of the noticed traffic characteristics. Then the first objective of the research prediction model of the operating speeds has directed the choice of the roads to submit to the second stadium of the monitoring survey.

Figure 1. Extract of the plan of the SP30B road.

The most important criterion of the selection of the sample has been the necessity to recognize the determining factors of the so-called “memory effect” (Figure 1): in fact it is a consolidated acquisition of the modern studies the speed, that the driver adopts on a specific element of the layout (curve, tangent, etc.), does not depend only on the characteristics of the same element, but it is also influenced by the features of the horizontal alignment of the homogeneous layout context, in which the element is inserted. The measures of speed have been effected along the minor rural roads SP30A, SP30B, SP52, SP204, SP312, SP262 and the major rural roads SS19ter, SS166 and SS426, located in the Vallo di Diano and in the Piana del Sele. The traffic never exceeded 400÷500 vehicles/hour during data collection; the relief have always been conducted in daylight conditions, with no rain and dry paving. The measure sections have been located on geometric elements with constant curvature in free flow conditions as well as along transition sections from rural environment to the urban one.
2. Operating procedures adopted

For the purposes of these researches, we used three traffic counters “KV Laser” made by SODI Scientifica (Figure 2).

The operation feature of these devices, endowed with two photocells, is based on the issue and the receipt of a laser couple beams perpendicularly direct to the road axis. The laser beams are to issued low-power (class 1) and, therefore, harmless for the vehicles occupants; the vehicle speed is inferred by the temporal interval between passage of the vehicle from the first one to the second photocell. The measurer is lodged in fixed posting on easel or (where possible) in the cabin of a car parked on the margins of the roadway or inside a box, so that to protect it from the drivers sight not to condition its behaviour.

The instrument has been fed by a battery (accumulator to the lead), that guarantees the autonomy of operation of it, but the low consumption would also have allowed to use the endowments of the support vehicle. The tool is endowed with a hard memory, inside which the recorded data are accumulated; then all the information can be transferred with the devoted software, on a personal computer through a serial cable.

To every vehicular passage the devices have recorded:
- time (date, hour, minutes, seconds);
- vehicle speed (in km/h);
- length of the vehicle (in meters);
- direction of march (binary variable: “direction 0” and “direction 1”).

The available mass storage has allowed to manage an elevated number of transits for long temporal intervals.

In the transfer of the data some light anomalies of the system of relief have been managed, relative, for instance, to the case of passages outdistanced less than 0,5 seconds

Figure 2. SS426, SP30A, SP312: speed measures.
having direction of opposite march, that they contemporarily transit in front of the laser measurer; it is not been able, besides, to intervene in the case in which the axis of the couple of laser beams were projected on surfaces with low reflecting ability and in the cases in which the axles of the vehicle were not perpendicular to the axis of the roadway.

3. Data Analysis

The elaboration has been completed for the only relief performed on the minor rural roads SP30A, SP30B and SP312 located in the zone of the Piana del Sele, and, instead, on the minor rural road SP52 and the major road SS426 located in the region of the Vallo di Diano (Table 1); particularly, the first three infrastructures analyzed distinguish from the other roads that constitute the sample, since they have predominantly straight layouts with the interposition of short circular curves and long radius of curvature; the authors have been recorded vehicle speed more elevated than the design speeds and by far superior to the imposed legal speed limits.

The models drawn in the first stations of measurement for the relationship between the operating speeds (derived variable), curvature and $V_{env}$ (independent variable), are delivered in Table 2.

The surveyed data on the circular curves with long radii return, as from literature [1, 3, 5, 6], values with high dispersion of the operating speeds, also with the same curvature; therefore it is confirmed the geometry of these curves influences the behavior of the drivers less than the ones with minor radius.

The surveyed speeds almost result anywhere, also in proximity of curves with inferior radii, superior to the imposed legal speed limits, confirming the poor attitude of the drivers to receive the Italian standards (particularly, for the majority of the examined roads, values of the $V_{85}$ about 70 ÷ 90 Km/h have been recorded, comparing to a legal speed limit of 50 Km/h).

<table>
<thead>
<tr>
<th>PIANA DEL SELE</th>
<th>VALLO DI DIANO</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Road</strong></td>
<td><strong>Stations</strong></td>
</tr>
<tr>
<td>S.P. 30B</td>
<td>43</td>
</tr>
<tr>
<td>S.P. 30A</td>
<td>82</td>
</tr>
<tr>
<td>S.P. 204</td>
<td>10</td>
</tr>
<tr>
<td>S.P. 262</td>
<td>45</td>
</tr>
<tr>
<td>S.P. 312</td>
<td>80</td>
</tr>
</tbody>
</table>

Table 1 – Location of survey stations.
From the data examination the first prediction bonds of regression have sprung between the $V_{85}$ and the different geometric features of the horizontal layout of the sections; as it is already told before (the vertical alignment has not been taken into consideration, because in selection phase of the infrastructures, object of investigation, we consider only the one with longitudinal slope inferior to 6%; value below which no meaningful correlation is statistically admitted between operating speed and longitudinal slope).

For the circular curves the database confirm the most important variable of $V_{85}$ is the local curvature. On the in relief data different typologies of relationships have been developed: among all the linear and square one are resulted more representative vs. of the only variable C.C.R. (Curvature Change Rate) (Figure 3, 4 and 5).

![Figure 3. Diagram of the CCR calculated for the SP30b road.](image)

For the curves of longer radius ($R > 1000$ m) the correlation is much weaker, because of the notable dispersion of speed determinations with the same radius.

The ample variability of the indeed practiced speeds, also found in analogous experimentations of other researchers in the world, depends on the different driving behavior, especially in comparison to other factors (said “secondary”), that characterize the road environment (tortuosity of the section in which the curve is inserted, width lane, widening in curve, presence of obstacles, free views, paving conditions, etc.) and are more difficult to be systematically treated. Along circular curves of radius smaller than 1000 m the incidence of the curvature on the effected speed grows to damage in order to all other factors.

This important conclusion is testified on the same radius, with the sample distributions of the rather similar speeds are noticed.

Nevertheless also on this typology of curves, in the case in which in more flowing road sections they revert, the influence of all the other factors “secondary” grows up; insofar, on circular links with the same radius, meaningful differences are often found among the characteristic parameters of the distributions in relief speed.
The followings models propose, respectively consequential from the elaboration, both of the single roads that related to the first 350 stations of measure considered in function of the only plane features of the elements that the infrastructures are constituted.

<table>
<thead>
<tr>
<th>Road</th>
<th>Independent variable</th>
<th>Model</th>
<th>Regression parameters</th>
<th>R²</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Z.P. MB</td>
<td>CCR</td>
<td>V85=a0 + a1*CCR</td>
<td>98,68</td>
<td>-0,082</td>
<td>0,54</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>V85=a0 + a1*(1/R) + a2*(1/R²)</td>
<td>78,21</td>
<td>2139</td>
<td>-7346056</td>
</tr>
<tr>
<td>Z.P. 112</td>
<td>CCR</td>
<td>V85=a0 + a1*CCR</td>
<td>72,24</td>
<td>-0,020</td>
<td>0,58</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>V85=a0 + a1*(1/R) + a2*(1/R²)</td>
<td>75,86</td>
<td>-3151</td>
<td>71681,8</td>
</tr>
<tr>
<td>Z.P. S2</td>
<td>CCR</td>
<td>V85=a0 + a1*CCR</td>
<td>62,87</td>
<td>-0,015</td>
<td>0,58</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>V85=a0 + a1*(1/R) + a2*(1/R²)</td>
<td>47,16</td>
<td>-3953,2</td>
<td>27945,4</td>
</tr>
<tr>
<td>S.S. 426</td>
<td>CCR</td>
<td>V85=a0 + a1*CCR</td>
<td>84,4</td>
<td>-0,058</td>
<td>0,56</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>V85=a0 + a1*(1/R) + a2*(1/R²)</td>
<td>81,8</td>
<td>-2261,1</td>
<td>-16081,9</td>
</tr>
<tr>
<td>All roads</td>
<td></td>
<td>CCR</td>
<td>V85=a0 + a1*CCR</td>
<td>74,91</td>
<td>-0,023</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R</td>
<td>V85=a0 + a1*(1/R) + a2*(1/R²)</td>
<td>76,49</td>
<td>-2702,1</td>
</tr>
</tbody>
</table>

Table 2 – Regression models of the V85 for the single roads.

Figure 4 – Linear regression of the V85.
4. Conclusions

Experimental evidence has confirmed on the two lane rural roads, that represent the most part of the Italian road network, the speeds adopted by the drivers are not congruent with the road design standards and with the imposed legal speed limits. The high dispersion of the speed distributions is an indicator the drivers, in the choice of the driving behavior, are little conditioned from the layout.

The models developed for the estimate of the operating speed $V_{85}$ show very good prediction ability with the introduction of the environmental speed $V_{env}$. To the $V_{env}$ we have attributed, for every homogeneous section, the values of the $V_{85}$ measured on the tangents or on the greatest radius curves belonging to the same section.

To get tools of the operating speeds of general validity it has been necessary to release from the singleness of the specific survey roads, setting original prediction models of the $V_{env}$ according to a scheme analogous to that proposed by the German design standards in force. To such end, the choice of the survey homogeneous sections has also been directed by the necessity to get a sufficiently greater randomness of $V_{env}$ determinations.

The data base available has allowed, as estimated to the action of the planning of the monitoring survey, to develop a prediction model of the operating speeds that overcomes the distortions caused by the specific features of the narrow sample of the infrastructures object of the searches illustrated in the proceedings previously published [2, 6].

The models formalized, at the end of the regression analyses previously illustrated, are following here brought:

$$
V_{85} = -2073.7 + 31029 \cdot (1/R)^2 + 0.87 \cdot V_{env} \quad \rho^2 = 0.81 \quad (1)
$$

$$
V_{env} = 82.84 - 0.1033 \cdot CCR + 3.44 \cdot L \quad \rho^2 = 0.84 \quad (2)
$$
Where \( V_{85} \) is the operating speed on curve, \( R \) is the radius (30÷5000 m), \( V_{env} \) is the environmental speed, \( CCR \) is the curvature change rate and \( L \) is the lane width (2.5÷5 m).

The model with “two steps” proposed allows to predict the \( V_{85} \) and the \( V_{env} \) when the geometry (either of the specific element component the planimetric layout or of the homogeneous section in which it is inserted) is known.

The model, calibrated on an ample experimental casuistry, is representative of the real flow conditions on the territory object of investigation and, at the same time, of immediate application by the highway administrations.

Acknowledgments

The authors would like to thank the Assessorato of the “Transportation, Infrastructures and Mobility” of the Province of Salerno for promoting the research, particularly chief engineer Giovanni Coraggio and eng. Alessandro Annunziata of the Road Safety Office.

References


GOODS DELIVERY DISTRIBUTION IN ROME: AN APPROACH TO THE SYSTEM REORGANIZATION

Antonio Musso, Maria Vittoria Corazza

Abstract. The paper deals with the organization of the delivery system in Rome. After an outline of the “best practices” abroad on this topic and of the results of a cluster analysis on the relationships among implemented measures and some socio-economic features of the sites in which they are implemented, the paper reports the study run by DITS – University of Rome “La Sapienza” concerning the feasibility of a re-organization process for delivery operations in Rome’s central areas. The study focuses, first, on the design of a suitable number of loading/unloading lots, currently not sufficient, and next, once that such basic requirement is met, on the development of successful strategies to improve delivery operations, under the sustainability point of view.

Introduction

The poor supply of loading/unloading bays in urban areas is one of the most recurring mobility problems. Citizens and operators perceive only the most noticeable aspects of such malfunction and very seldom are able to fully understand the reasons lying beneath.

The causes of bad performance in delivery operations are manifold: the poor knowledge of the phenomenon, both in terms of statistics and regulations, the lack of dedicated governance measures, the modest participation of operators to decision making processes.

1 DITS – Dipartimento di Idraulica Trasporti e Strade, Facoltà di Ingegneria, Università degli Studi di Roma “La Sapienza”, Via Eudossiana 18, 00184 Roma, antonio.musso@uniroma1.it
In order to support this statement, it is worth reporting some facts on the phenomenon, at national level. In Italy, about 500,000,000 tons of goods are yearly moved via road, within a distance range minor than 50 km, by operators directly involved in retail activities; unfortunately, national statistics, which provide data on these facts, are based on surveys concerning delivery operations performed by heavy vehicles (> 3,5t) only. This means that most of deliveries in cities, usually operated directly by retailers or minor carriers, are not even surveyed. The lack of investigations on the delivery activities as an urban phenomenon partly explains why, under the regulatory point of view, to national directions and specifications very detailed on planning and management issues on delivery general activities, do not correspond a suitable commitment by local administrators. Deliveries are usually managed by regulatory tools as Urban Traffic Plans, which often are too private-traffic-dedicated to appropriately deal with the delivery problems. Another point to consider is the local operators’ general wariness; they are not ready for negotiations, so it is very difficult to have proper knowledge of their requirements, which are often based on “at-all-cost deliveries” rather than on “smart deliveries”. Furthermore, historical urban features and very dense land use greatly complicate delivery system networks in historical centers. Needless to say, the resulting approach used by most of local administrators is to provide short-term solutions to emergency situations, short-sighted answers which do not meet citizens’ requirements.

1. The lesson learned from best practices

Sustainability is often the main engine which prompt local administrators to start enforcing regulatory measures to govern delivery problems. Starting from the analysis of best results, three main trends of governance actions can be outlined: incentives/disincentives to some forms of deliveries, participation and support by external actions; in some cases policies based on a mix of the three trends are applied, but more often decision makers prefer to implement just one of them, in accordance to the availability of resources, to the common awareness levels, to the built environment. Best practices show, however, that in some cases, these three parameters can turn into bonds which may interfere one to the others, becoming even hindrances for the administrators’ choices.

A cluster analysis, based on the survey of 15 European cities and towns current policies on delivery organization, was run; thanks to this analysis it is possible to investigate relationships between implemented measures and respectively: number of inhabitants, urban density and size of urban areas. From this analysis it seems that soft approaches (i.e. those based on participation, on infrastructural improvements and fundings) occur mainly in cities with population between 500,000 and 1,000,000 inhabitants; approaches mostly
based on restrictiveness (time slots for operations, emissions control, weight limits, etc.) are usually implemented in urban areas with 1,500,000 inhabitants and more.

Focusing on some relevant, recurring measures as, for instance, the compulsory fleet renewal to meet sustainable emissions standards, the creation of transit points and of distribution centers and eventually the infrastructural interventions, the relationship between cost per capita and number of inhabitants was studied for middle size towns (less than 1,000,000 inh.), as well. In synthesis, from the analyzed sample, transit points seem to be in average the most expensive solution (12.5 Euro/inh. as mean value), whereas the renewal of the fleet seems to be the cheapest (always < 5 Euro/inh.).

Such results, even though based on a restricted sample, outline how difficult could be the selection of the appropriate solution/s to manage the delivery organization and, conversely, how mere infrastructural interventions (always < 10 Euro/inh.) appear to be the most feasible, in spite of their middle-range costs, because of the easiness of the decision and implementation processes.

2. The Rome case study

2.1. Contemporary situation

Everyday deliveries are an important issue in the mobility scheme of Rome. Freight entering via road the city amounts to about 90,000 tons/day; the historical centre represents 1.1% of the urban area but it attracts 33% of the total amount of delivered goods. About 25,000 freight vehicles enter the Freight Limited Traffic Zone (FLTZ, an area larger than city centre) per day [1]. The most striking aspect of this process is the lack of an appropriate number of parking lots, thus the Municipality’s first answer seemed, since ever, to be just to plan more loading/unloading areas. It was soon evident that this was not the most appropriate solution for the city centre, where narrow streets, operators’ non respect of the rules and citizens’ basic parking requirements make the situation worse, day by day.

2.2. The case study

Naturally, Rome municipality would like to apply lessons learned in other European cities to improve the logistic system, specifically in the Limited Traffic Zone, in order to optimize routes. Hence, a study was conducted to assess the real need of loading/unloading areas in the city centre as a pre-requisite to whatever other initiatives should be implemented to solve delivery problems at urban scale. Main results of such study, performed by DITS – University of Rome “La Sapienza”, focus on the detection of a suitable loading/unloading areas supply, as from the comparison between place performances and use demand; such comparison defines a kind of do-something scenario, i.e. what would happen if a progressive increase of lots for delivery would be applied in a short-term period.

The reasons that prompted to develop the above scenario relied on the unsuitability of a first estimate based on the creation of a “Business As Usual (BAU) scenario”, in which the
implementation of just local initiatives was assumed to occur. This forecast, merging information on existing data and on the Municipality practice implemented so far, led to assess an increase of approximately 10% on lots availability, raising the number of load/unload areas from 186 to 201 in four years as a result of natural growth in supply, which seemed unsatisfactory.

2.3. The methodology

It was thence necessary to develop a comparison between operators’ requirements and space supply, aimed at depicting a realistic do-something scenario. The study was developed starting from spots analyses and an example of how many lots would be needed in a representative street of the city centre is reported as follows. The number of lots required was estimated using a simple formula (1)

\[
\text{u.r.} = \frac{\Phi}{nK}
\]

(1)

u.r. – utilization ratio
\(\Phi\) – frequency of commercial vehicles arrivals (vehicles/hour)
\(K\) – frequency of service for a given loading/unloading area, esteemed as 4 vehicles/hour
\(n\) – number of loading areas

Main assumptions for the calculation of the number of loading/unloading areas were: these lots must be used by commercial vehicles only (i.e. no illegal use by private vehicles is possible); frequency of commercial vehicles arrivals can be either the surveyed maximum one or the average one; u.r. is fixed either 0.75 or 0.5. Data on arrivals, traffic flows and infrastructural supply were collected during a set of dedicated surveys, which took place during the winter 2005. Each kind of commercial activity was surveyed, thanks to interviews with retailers aimed at determining the number of deliveries they have and they send daily and hence the amount of vehicles each kind of retail manages everyday. The total amount of operated deliveries was summarized into a daily attractiveness coefficient which relates each kind of commercial facility to the daily average number of vehicles that serves it.

For an analyzed street (in Figure 1 Via dei Serpenti, an example of a typical commercial street in the city centre, is reported, where currently only three loading/unloading lots are available), the number of lots increases relevantly. In case of 0.5 u.r., 11 extra parking lots should be required for operations from 7.30 a.m. to 1.30 p.m. (providing 14 lots in total) and 7 new ones for operations from 1.30 p.m. to 6.30 p.m.; in case of 0.75 u.r., 6 extra parking lots should be required for operations from 7.30 a.m. to 1.30 p.m. (providing 9 lots in total) and 5 new ones for operations from 1.30 p.m. to 6.30 p.m. Extra surveys of commercial vehicles operating on Via dei Serpenti confirm such needs; moreover calculations based on 0.5 u.r. seem to be the most appropriate, given the long time needed for deliveries to commercial facilities located at the upper side of the street.

By expanding the obtained results to the downtown level, it was estimated that about 643 new lots are required, in accordance to the current development of the city and
consistent with the commercial high-density nature of the area; the consideration of residents’ parking requirements, in light of the fact that creation of new bays for deliveries often results in fewer parking lots for residents, was also crucial for the assessment of the number of loading/unloading areas. It was reasonable however to consider a smaller amount of lots (about 600), since other hindrance factors as driveways, garbage collection points, parking areas for administrative, religious, tourist purposes reduced the available space, as well.

Figure 1. Unloading/unloading lots supply according to 0.5 u.r. calculations

2.4. The creation of interventions scenarios

Once the individuation of the proper number of loading/unloading lots in central areas was achieved, the next step concerned the development of strategies to improve the overall delivery system and hence to evaluate the most successful strategy.

On this purpose, it was necessary to define the objectives for the development of the improvement strategies/scenarios, which were:
- Increase of the delivered load by 15 %
- Rationalization of the delivery process
- Promotion of eco-friendly vehicles for delivery operations.

These three main goals allowed developing three strategies/scenarios, different, in terms of increasing complexity of application (quantity and quality of the measure to implement) and in terms of “premium” benefits; i.e. the more complex scenario, the best results. The
selection of measures to implement within each scenario was run also according to the results of the analysis of best practices as shortly synthesized at paragraph 1. The three scenarios, described as follows, are synthesized in Table 1, as well:

- **Scenario A** was aimed at meeting operators requirements, in order to have an increase of the loading factors; measures to achieve the goal were related to regulations, infrastructural and management domains; expected results should reduce deliveries fragmentation and not efficient operations. This scenario was considered a soft one because it was just targeted to control and reduce accidental and not organized operations and to prevent shopkeepers from making delivery operations by themselves.

- **Scenario B** was aimed at amending the infrastructural parking situation; indeed, most of the scenario measures were targeted both to improve physically the parking supply and operatively the delivery processes, especially for what concerned time and way of performing such activities. Needless to say that the difference with the previous scenario relies on the highest cost of the measures requested by scenario B.

**Objectives:**
1) Increase of the delivered load by 15%  
2) Rationalization of the delivery process  
3) Promotion of eco-friendly vehicles for delivery operations

<table>
<thead>
<tr>
<th>Scenarios</th>
<th>A</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Achievable objective: 1</td>
<td>Achievable objectives: 1,2</td>
<td>Achievable objectives: 1,2,3</td>
<td></td>
</tr>
<tr>
<td><strong>Measures</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Check on vehicles with carried load less than 1.5 t</td>
<td>Routes optimization</td>
<td>Time permissions</td>
<td></td>
</tr>
<tr>
<td>Routes optimization</td>
<td>Vehicles Routing</td>
<td>Check on vehicles with carried load less than 1.5 t</td>
<td></td>
</tr>
<tr>
<td>Vehicles Routing</td>
<td>Creation of a TP - transit point</td>
<td>Creation of loading/unloadin areas</td>
<td></td>
</tr>
<tr>
<td>Creation of a TP - transit point</td>
<td>Delivery to TP during off-peak hours</td>
<td>Creation of lanes multifunction</td>
<td></td>
</tr>
<tr>
<td>Delivery to TP during off-peak hours</td>
<td>Routes optimization</td>
<td>Pricing for not eco-friendly vehicles</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Vehicles Routing</td>
<td>Creation of UDC – Urban Distribution Centers</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Creation of a TP - transit point</td>
<td>Loading/unloading area booking</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Delivery to TP during off-peak hours</td>
<td>Routes optimization</td>
<td></td>
</tr>
<tr>
<td>Expected outcomes</td>
<td>Reduction of the number of delivery vehicles</td>
<td>Reduction of the delivery time by 15%</td>
<td>Reduction of the number of delivery vehicles</td>
</tr>
<tr>
<td></td>
<td>Reduction of the number of unfulfilled deliveries</td>
<td>Reduced congestions on some operation areas</td>
<td>Reduction of the number of unfulfilled deliveries</td>
</tr>
<tr>
<td></td>
<td>Improvement on air quality</td>
<td>Reduction of the delivery time by 15%</td>
<td>Improvement on air quality</td>
</tr>
<tr>
<td><strong>Prerequisites</strong></td>
<td>Specific regulations</td>
<td>Enforcement</td>
<td></td>
</tr>
<tr>
<td><strong>Accompanying measure</strong></td>
<td>On-board POS</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table 1. The Scenarios and the related measures**
Scenario C, the most complex, was virtually dedicated to meet environmental requirements: this means that to put into practice the scenario’s measures, political will and operators acceptance had to be pursued first, being the set of measures very restrictive (due especially to the pricing for not eco-friendly vehicles). This was also a very long-term scenario, so its feasibility must be based also on the possibility to achieve the expected environmental benefits in a very long time.

Measures of the three scenarios were meant to meet the following general requirements: sustainability (meant as control of pollution due to commercial vehicles emissions), parking, safety and security, fight against illegal behaviors, easiness of operations, time savings, reduction of enforcement and control resources, as reported in Table 2. Such requirements came from the needs as achieved by the interviews with local delivery operators and retailers.

To assess the suitability of each proposed scenario, a multicriteria analysis was carried out, providing a different weight to each measure, according to its influence level (high = weight 3, medium = weight 2, low = weight 1). It was possible to evaluate in this way not only how much successful a given scenario could be, but which were the most effective measures in relation to the requirements expressed by the interviewees. Accordingly, most needed and effective-assumed measures seem to be the time-related ones: time slots to enter the city centers and hence improved travel times. Both actions, however, need a strong regulatory support and strict enforcement, the latter enhanced even by an upgrading of the telematic access control system, already operative in the city centre.

Switching to a more general level of assessment, from the comparison among the three scenarios it is possible to stress how Scenario B seems to be the most effective in terms of
overall favourable achievements, but if priority is given to sustainability issues both Scenarios A and C seem to be preferred, as synthesized in Table 3.

<table>
<thead>
<tr>
<th>Requirements</th>
<th>Sustainability</th>
<th>Policing</th>
<th>Safety and security</th>
<th>Illegal behaviours</th>
<th>Operations</th>
<th>Time</th>
<th>Cost</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scenario 1</td>
<td>43/45</td>
<td>19/40</td>
<td>18/5</td>
<td>12/30</td>
<td>27/45</td>
<td>20/60</td>
<td>3/15</td>
<td>134/500</td>
</tr>
<tr>
<td>Scenario 2</td>
<td>44/72</td>
<td>55/96</td>
<td>57/2</td>
<td>27/43</td>
<td>54/72</td>
<td>73/96</td>
<td>15/24</td>
<td>272/488</td>
</tr>
<tr>
<td>Scenario 3</td>
<td>55/72</td>
<td>55/96</td>
<td>47/2</td>
<td>33/43</td>
<td>42/72</td>
<td>51/96</td>
<td>11/24</td>
<td>253/488</td>
</tr>
</tbody>
</table>

Errors are ordered to the highest achievable value; the each class of requirements, the social sustainability; the increase, in scenario 3, sustainability achieve 45 as a score, whereas 45 is the achievable maximum, which is obtained providing with a weight 10 each item belonging to the sustainability class.

---

Table 3. The Scenarios relevance

**Conclusions**

An important issue emerging from this study is the mandate to meet parking requirements. Having provided this basic requirement, every other intervention is possible. In Rome, given the poor supply of loading/unloading areas for commercial activities, any extra action in terms of fleet renewal, e-commerce development, logistic platforms location, etc. must be postponed until the loading supply question is addressed and solved. Another point to consider is the participation process: until communication barriers between operators and administrators will be overcome, not shared, suitable solutions will not be possible.

Political will is required, both in terms of measures to implement and in terms of acceptance of disincentives, but political measures are not sufficient when the built environment has premium value. Typically in this case, the provision of loading/unloading areas is poor and conflicts between residents and operators are likely to arise. In this case, only a relative optimum can be reached for each solution.

**References**

TRAVEL TIME ESTIMATION
USING TOLL COLLECTION DATA

Francesc SORIGUERA¹, Leif THORSON², Francesc ROBUSTÉ³

Abstract. Transportation System Management (TSM) permits optimizing the current available road network. Travel time and its reliability are key factors in road management systems, as they are good indicators of the level of service in any road link, and perhaps the most important parameter for measuring congestion. This paper presents a new approach for direct travel time measurement using existing toll infrastructure. An efficient and simple algorithm has been developed and implemented in the toll highways in Catalonia, around Barcelona, Spain.

1. Introduction

Travel time and travel time reliability are key factors in road management systems, as they are the best indicators of the level of service in a road link, and perhaps the most important parameter for measuring congestion (Palen, 1997). Travel time estimation is necessary to assess the operational management and planning of a road network. Moreover, travel time information is the best and most appreciated traffic information for road users.

Most of the European countries (Spain, France, Denmark, Italy, Finland, United Kingdom, Sweden, the Netherlands, Norway and Germany) consider travel time as an emerging issue in Europe and realize of the growing demand for real time travel time information. Most of the Trans-European Road Network (TERN) in these countries is now covered by a euro-regional project in travel time.

Because of the growing interest in measuring travel time, there have been several studies attempting to determine link travel times in a road segment. So, Dailey (1993) and

¹ CENIT, Jordi Girona 29, 2-A 08034 Barcelona (Spain), francesc.soriguera@upc.edu
² CENIT, Jordi Girona 29, 2-A 08034 Barcelona (Spain), leif.thorson@upc.edu
³ School of Civil Engineering, Technical University of Catalonia, Jordi Girona 1-3, B1 08034 Barcelona (Spain), f.robuste@upc.edu
Petty et al. (1997) focus on the travel time estimation from loop detector data. These algorithms have been proven to be quite effective in free flow conditions, but they don’t perform well under heavy congested situations. For these last situations, Westerman and Immers (1992) propose a cumulative flow balance algorithm at successive detector sites for estimating travel time. The counter detector drift is the problem in this case. Other studies analyze the direct measure of travel time, using AVI systems (Levine and McCasland 1994, Christiansen and Hauer 1996, Seki 1995) or license plate matching (Cui and Huang, 1997), among others.

This paper presents a new approach in travel time measurement using toll infrastructure. A simple algorithm for estimating travel times in toll highways is developed. The algorithm uses the travel time data included in the toll tickets for different itineraries in order to estimate the section travel time for consecutive entry and exit points. The results of a pilot test in the AP-7 highway in Spain are also outlined in the paper.

2. Contribution of the proposed algorithm

The data on travel times can be obtained by measuring the time taken for vehicles to travel between two points on the network. In toll highways, the data needed for the fee collection system, can be also used for travel time measurement.

The toll tickets allow for measuring travel time for all origin – destination relations in the highway. However, travel time data is obtained once the vehicle has left the highway. This involves a great delay in the information in long trips. Moreover, travel time for long trips can be increased by factors that are unrelated to traffic conditions, for example if the driver stops for a break. In order to reduce the influence of such events and the information delay on the calculated average travel time, it is needed to estimate the section travel time between consecutive entry and exit points (see Fig. 1). This will also provide valid information for all drivers in the highway (not only for vehicles with the same origin – destination itineraries), and could also enable incident detection applications.

![Fig. 1. Highway travel time definition](image-url)

One possible solution to estimate the section travel time could be to include only measurements between consecutive entry and exit points. This solution, used in the Italian
“AutoTraf” system (Hopkin et al. 2001), may reduce in excess the amount of data in certain sections of the network, where the volume of traffic entering and leaving the highway at consecutive junctions is low, but there is a large volume of passing traffic. Moreover, this solution does not account for the “exit time” (i.e. the time required to leave the highway). The exit time includes the time required to travel along the exit link plus the time required to pay the fee in the toll plaza. Then, if the time to travel along a particular itinerary, composed of various single sections, is calculated by adding the single section travel times, the resulting itinerary travel time will be excessive, because it includes as many exit times as single sections compose the itinerary.

The algorithm presented in this paper solves the problem of long journeys by estimating single section travel times and allows for dividing the section travel time and the exit time. The estimated exit time is a very useful measure for highway operators, as it is an indicator of the level of service in the toll plazas. Moreover, the algorithm uses data from most of the itineraries in the highway increasing the accuracy and reliability of results.

3. Estimating section travel times from toll ticket data

For each particular vehicle “k” been driven along a highway with a closed tolling system (see diagram in Fig. 2), the travel time spent in its itinerary between “i” (origin) and “j” (destination) expressed as “ti,j,k” can be obtained by matching the entry and exit information recorded in its toll ticket.

The average travel time for the itinerary in a particular time period “e” (e.g. 1 hour) can be obtained by averaging the travel times of all vehicles that have entered the highway within the time period and have traveled along the same itinerary “i,j”.

\[
\bar{t}_{i,j} = \frac{1}{n_{e_i}} \sum_{k=1}^{n} t_{i,j,k}^{(e_i)} \quad \forall i = 0,\ldots,m-1 \quad \forall j = 1,\ldots,m \quad \forall e = 0,\ldots,23 \quad (1)
\]

Where: \((e_i)\) is an hourly time period in the vehicle entrances in the toll plaza “i” of the highway.
$t^{(e_i)}_{i,j,k}$ is the travel time for the itinerary “$i,j$” for a particular vehicle “$k$” that has entered the highway within the time period “$e_i$”.

$t^{(e_i)}_{i,j}$ is the average travel time for the itinerary “$i,j$” in a particular hourly time period “$e_i$”.

$n_{e_i}$ is the number of vehicles that have entered the highway within the time period “$e_i$”.

From these calculations, the average travel time for all the itineraries in a particular time period, “$t^{(e_i)}_{i,j}$”, is obtained. The next step is to calculate the single section travel time (i.e. travel time between consecutive entry and exit points) and the exit time (i.e. the time required to travel along the exit link plus the time required to pay the fee in the toll plaza).

In general, the average travel time “$t_{i,j}$” can be divided in two parts: the section travel time “$t_{s(i,j)}$” and the exit time “$t_{e(x(j))}$” (see Fig. 1).

\[ t_{i,j} = t_{s(i,j)} + t_{e(x(j))} \] (2)

If we consider the highway stretch between entrance 0 and exit 1:

\[ t_{0,1} = t_{s(0,1)} + t_{e(1)} \] (3)

Then it can be seen that subtracting different travel times of selected itineraries, the section travel times and the exit times can be obtained (see Fig. 3). Then for the (0,1) itinerary:

\[ t_{s(0,1)} = t_{0,2} - t_{1,2} = t_{0,3} - t_{1,3} = ... \] (4)

\[ t_{e(1)} = t_{01} - t_{e(0,1)} \] (5)

Fig. 3. Section (0,1) travel time estimation

Note that for sections with entrance different from the initial toll plaza, equation (3) should be rewritten as:

\[ t_{i,j} = t_{e(n(i))} + t_{s(i,j)} + t_{e(x(j))} \quad \forall i = 1,...,m-1 \quad \forall j = 2,...,m \] (6)

Where: $t_{e(n(i))}$ is the “entrance time” (i.e. the time required to travel along the entrance link)
In the present paper it is assumed that the entrance time "ten(i)" is smaller enough in relation to the section the section travel time "ts(i,j)" and to the exit time "tex(j)" to be rejected. So, in a general expression, the single section travel times and exit times can be calculated for each stretch as:

\[ t_{s(i,i+1)} = \frac{\sum_{j=i+1}^{m} (t_{i,j} - t_{i+1,j})}{m + (i + 1)} \quad \forall i = 0, \ldots, m - 2 \quad (7) \]

Where “m” is the last toll plaza in the highway.

To obtain the exit time we only need to subtract this single section travel time to the total itinerary travel time in adjacent entrance/exit points. Then:

\[ t_{ex(i+1)} = t_{i, i+1} - t_{s(i,i+1)} \quad (8) \]

4. Application to the AP-7 highway in Spain

The AP-7 highway runs along the Mediterranean cost corridor in Spain, from the French border to the Gibraltar strait. The pilot test was restricted to the north east stretch of the highway from “La Roca del Vallès” toll plaza to the French border at “La Jonquera”, of approximately 120 km long.

The pilot test was performed with the data of 10th of July of 2005, which was a particular conflictive day in the highway. The results of the application of the algorithm allow for obtaining the single section travel times for each of the 13 sections of the stretch and for each of the 24 hours of the day. Moreover the exit times for each toll plaza are also calculated in an hourly time period basis. The analysis of the obtained results determines that the proposed algorithm is a suitable new approach to estimate the single section travel times in toll highways.

References


Petty, K., et al. (1997) Accurate estimation of travel times from single loop detectors, 76th annual TRB meeting, Transportation Research Board.


A PROBIT MODEL FOR DEPARTURE TIME CHOICE

Mike MAHER¹
Andrea ROSA¹

Abstract. A random utility model of departure time choice should account realistically for the correlation between costs of travelling at alternative times and should not be unduly sensitive to time discretisation. With these objectives we put forward a model with autoregressive structure of the stochastic term which results in a multinomial probit model. Its formulation and its analytical solution are discussed as well as the resulting choice patterns.

1. Introduction

Departure time choice (DTC) has received an increasing amount of interest in the literature thanks to the recognition of the importance of time shifts as a response to transport system variations and the need to understand and model them.

Much of the recent literature treating DTC within the random utility framework investigates the most appropriate specification of the systematic utility for the model in a practical case. The present paper is concerned, instead, with the effects of making different assumptions about the random term of the utility when the systematic utility profile is known. In particular, this paper examines the implications of choosing the multinomial probit model (MNP) or a logit-based formulation, characterising the differences in their results, and their sensitivity to the level of discretisation of the time period being modelled.

2. A brief survey of the literature

Many DTC models appeared in the literature and based on random utility theory are instances of the multinomial logit (MNL) model: for example, those presented in [6], [1] and [10]. However convenient from the analytical and computational point of view, the

¹ Both authors at the School of the Built Environment and the Transport Research Institute, Napier University, 10 Colinton Road, Edinburgh, EH10 5DT, United Kingdom; emails, respectively, m.maher@napier.ac.uk and a.rosa@napier.ac.uk
MNL model has long been considered inadequate for modelling DTC, as it does not account for the correlation between the costs of travelling at different times. This point had already been made, for instance, by McFadden et al. in [6], who noted that accounting for correlation is especially necessary as the time intervals into which the modelled period is discretised become smaller.

Several studies have reported on work to include correlation amongst the costs of travelling at different times. For instance, [11] introduced the OGEV model (Ordered Generalised Extreme Value model), which has a closed-form expression for the choice function. Mahmassani and co-researchers (see e.g. [5]) employed the multinomial probit model within the bounded rationality framework to account for correlation between the decision of changing or keeping the departure time over a series of days. Their framework was also extended to model jointly departure time and route switching. Recently, [13] considered modelling departure time and route choice together. It looked at an idealised situation with a limited route choice set and included investigations on the MNL, the nested logit model, the mixed logit model, the MNP model and a model combining the OGEV model for departure time choice and the nested logit for route choice.

3. Models for departure time choice

In this paper departure time choice is modelled within the random utility framework: a tripmaker has a (discrete) set of choices for departure time $t$, with perceived total travel cost for choice $t$ being $C(t) = c(t) + e(t)$, where $c(t)$ is the mean and $e(t)$ the random perception error. The tripmaker chooses the departure time that minimises his or her perceived cost of travelling.

A satisfactory DTC model would be one in which the correlation between the perception errors of the costs at a time $\tau$ apart form each other, $e(t)$ and $e(t + \tau)$, tends to 1 when $\tau$ tends to 0, and decreases monotonically as $\tau$ increases. Another desirable property for a choice model to be used for DTC is that results are not unduly sensitive to the discretisation of the time period being modelled and remain consistent as the discretisation of the time is further refined, so that the model defined over a discrete set of departure times becomes, in the limit, a model defined over continuous time.

One way to obtain such a specification of the model, and one that is considered here, is to specify $e(t)$ as having an autoregressive AR(1) structure. If $t-1$ and $t$ are two consecutive departure time alternatives in the chosen discretisation of the time, the stochastic part of the cost is given by:

$$e(t) = a \cdot e(t - 1) + e(t)$$

(1)

where the $e(t)$ are defined as independently and identically distributed Normal random variables $N(0, s^2)$. If the variance of $e(t)$ is defined as equal to $\sigma^2$ for each time $t$, the relationship between $s^2$ and $\sigma^2$ is:

$$\sigma^2 (1 - a^2) = s^2$$

(2)

Moreover, considering (1), the correlation between the errors at any two time periods $t$ and $t + k$, that are $k$ time periods apart, results in:

$$corr(e(t), e(t + k)) = a^k$$

(3)
To achieve a correlation structure that has the properties outlined above, the term \( a \) is defined as a function of the rate of decrease of the correlation over time \( \lambda \):

\[
a = (1 - \lambda \Delta t)
\]

(4)

where \( \Delta t \) corresponds to the duration of each time interval considered, so that:

\[
\text{corr}(e(t), e(t + k)) = (1 - \lambda \Delta t)^k
\]

(5)

In the limit, as \( \Delta t \) tends to zero, the model is defined over continuous time, and the correlation of the perceived costs of travelling at two time periods that are a time \( \tau \) apart can be written, starting from (5), as:

\[
limit_{\Delta t \to 0} (1 - \lambda \Delta t)^{\frac{\tau}{\Delta t}} = \exp(-\lambda \tau)
\]

(6)

Thus, given the correlation between the costs of travelling at two times that are a time \( \tau \) apart, it is possible to calculate \( \lambda \), the rate of decrease of the correlation over time, and obtain, from (4) and (5), the correlation for the costs of the departure time alternatives defined over discrete intervals.

The AR(1) process defined above results in costs that are MVN distributed and thus in a multinomial probit choice model.

We contrast the model above with the simpler multinomial logit, which is frequently used in the DTC literature, notwithstanding its limitations remarked in section 2. In the MNL model the perception errors are identically and independently Gumbel distributed, and the probability of departure at time \( t \) is given by the formula:

\[
p(t) = \frac{\exp(-\theta c(t))}{\sum_{r} \exp(-\theta c(r))}
\]

(7)

where \( R \) is the set of available choices and \( \theta \) is a scale parameter related to the variance of the Gumbel distributions.

4. Analytical solutions of the multinomial probit model

While much of the recent literature on applied choice modelling, including traffic assignment, reports on MNP models solved by simulation, in this work the MNP is solved analytically with the numerical integration method and some of the approximations reviewed in [8], partly anticipated in [9]. Those works pointed out that there are more numerical methods for solving the MNP than previously discussed in the transportation literature and that several of them are of practical interest. The investigations for [8] and [9] included an assessment of the precision of different approximation methods and suggested that, amongst the approximations examined, the methods developed by Tang and Melchers in [12] and by Mendell and Elston in [7], the latter also discussed and modified in [4], give the best trade-off between results’ accuracy and computational effort and should be further considered in practical applications. The accuracy of the approximations had been evaluated by applying them to route choice test problems and comparing their results with those given by the numerical integration method of Genz (see [2] and [3]), in which the precision of the results can be set as an input. In this paper the approximations of
Mendell and Elston and of Tang and Melchers are used to solve the MNP as well as the numerical integration method of Genz. The latter, unlike numerical integration methods previously considered in the transportation literature, provides a feasible approach to MNP problems of relatively large dimension (for instance, in the work reported here it is used with up to 18 choice options).

5. Numerical experiments

To compare the effect of using different choice models on the same systematic cost profile, numerical tests have been carried out on a number of artificial departure time choice scenarios, for reference defined on a period conventionally taken of 3 hours, divided in 6, 9, 12, 18 periods of equal duration. Scenarios tested include constant systematic costs and systematic costs increasing or decreasing in the central part of the modelled period. The stochastic part of the costs has been set according to the AR(1) process defined above considering different levels of cost variance (2, 5, 10) (from which the logit parameter has been calculated) and different levels of correlation (0, 0.20, 0.40) between time points that are 1 hour apart. The correlation decay $\lambda$ has then been calculated from (6) and the correlation between pairs of options with (5).

A first result from the numerical experiments is the good accuracy of the MNP approximation methods employed. The numerical integration method of Genz can be taken as precise as its accuracy is set by the user ($10^{-4}$ in this case). Fig. 1 allows appreciating the good accuracy of the Mendell-Elston and Tang-Melchers approximations to the MVN integral. For both approximations inaccuracies are of larger importance for small actual choice probabilities (less than 0.05-0.10). For instance, in the experiments carried out, for the Mendell-Elston approximation the maximum percentage inaccuracies are in the range [−5%;+6%] and are due to very small actual probabilities; the maximum percentage inaccuracies for actual probabilities larger than 0.10 are within the range [−2%;+2%], but typical percentage inaccuracies are smaller.

The tests have also allowed us to characterise differences in the choice profiles from the probit and the logit models as well as obtain indications on their extent and the elements affecting them. For instance, in cases with a flat or approximately concave cost profile with correlated costs the profile of the choice probabilities is convex (or approximately convex) but it is flatter for the MNL than for the MNP (see fig. 1). The test results show how correlation, variance and importance of the deterministic part of the costs affect the differences between the choice profiles.

To evaluate the sensitivity of the models to the discretisation of the problem, the results from coarser time divisions have been compared with those obtained aggregating the results from finer ones. The aggregation has been carried out simply by summing the results for the time intervals corresponding to each period of the coarser discretisation.

The MNL model shows very good aggregation properties: the sum of proportions for finer discretisations practically coincides with the choice proportions for the coarser ones. The MNP results show a good agreement of choice profiles for different discretisations of...
the problem (see e.g. figure 2). However, the differences between the results from different
time divisions are clear, even though of small magnitude. This suggests that, although
results remain consistent, there is a loss of accuracy for coarser time discretisations.

![Figure 1](image1.png)

**Figure 1.** Logit and MNP choice probabilities (NI: num. integr.; TM: Tang Melchers;
ME: Mendell-Elston) for a concave cost profile (variance 5 and correlation level 0.20).
Modelled time divided in 18 intervals. Lines between markers are only indicative.

![Figure 2](image2.png)

**Figure 2.** MNP choice probabilities (obtained by numerical integration) for a concave
cost profile, cost variance equal to 2 and correlation level 0.20. The graph reports the
original results for a division in 6 time intervals and those obtained aggregating
results from finer time discretisations. Lines between markers are only indicative.

### 6. Further work

Planned further work includes the investigation of the AR(1) MNP model discussed above
on more cost profiles and its approximation and comparison with a GEV model able to
capture the similarity between pairs of alternatives described by the AR(1) structure.
An interesting extension of the models described would be to look at time intervals with duration tending to zero (i.e. the choice alternatives become infinitesimal time intervals). The practical difficulty is the characterisation of a method to solve the resulting continuous MNP model. By contrast, a continuous MNL model has already appeared in the literature.

References


NETWORK OF MAGNETIC DETECTORS FOR ENHANCING AIRPORT SAFETY

Kosmas DIMITROPOULOS, Nikos GRAMMALIDIS, Ioannis GRAGOPOULOS, Niovi PAVLIDOU, Thomas HEUER, Haibin GAO, Constanze STOCKHAMMER and Uwe HARTMANN

Abstract. Airport surface management is increasingly recognized as a critical process with respect to flight safety and air transport system capacity. The EC-funded ISMAEL project, presented in this paper, aims to determine whether recent advances in magnetic sensing techniques can provide a better means of surface movement surveillance at airports, either as a cost-effective alternative of ground radar or as an additional sensor in a multi-sensor A-SMGCS application.

1. Introduction

A steady increase of air traffic over the last decades has led to the situation where many major airports are working at their capacity limits. Hence, efficient solutions are required to overcome the resulting delays and to maintain safety of surface traffic at airports. Various systems have been developed in order to improve traffic capacity and increase safety, each with its own disadvantages and limitations:

- Surface Movement Radar (SMR) systems are subjected to masking and distortion in the vicinity of airport buildings, terrain features or airport installations.
- Optical detection systems are severely degraded under adverse visibility conditions (e.g. fog). However these are precisely the conditions in which optimum surveillance is required for safety.
- Advanced Surface Movement Guidance and Control Systems (A-SMGCS) and Airport Surface Detection Equipment (ASDE) suffer from reflections and multi-path phenomena, and gaps in surveillance coverage.

1 Informatics and Telematics Institute, 1st Km Thermi-Panorama, Thessaloniki Greece, dimitrop@iti.gr, ngramm@iti.gr, grag@iti.gr and niovi@eng.auth.gr
2 Institute of Experimental Physics, Saarland University, D-66041 Saarbruecken, Germany, t.heuer@mx.uni-saarland.de, h.gao@mx.uni-saarland.de, u.hartmann@mx.uni-saarland.de
3 HiTec Marketing, Lothringerstrasse 14/6, 1030 Vienna, Austria, cs@hitec.at
• The newly-introduced multi-lateration systems work only with transponder-equipped aircraft (i.e. co-operative targets). They cannot detect non-cooperative vehicles or aircraft that are not equipped with such a transponder, or whose transponder has failed, is malfunctioning or switched off. This is a potential safety hazard if multi-lateration is used as the sole means of maintaining separation and control.

• The same applies to GPS-based systems that cannot provide the necessary reliability in terms of availability, continuity and integrity.

Therefore, there is a strong demand for a new sensing technology providing for cost-effective detection of non-cooperative targets under any weather condition. The EC-funded research project ISMAEL [2] aims to determine whether recent advances in magnetic sensors can provide a better means of surface movement surveillance at airports. In this context, a dual approach is pursued, either to serve as a cost-effective alternative to SMR for smaller airports or as an additional point sensor in a multi-sensor A-SMGCS at major international airports. The aircraft position within the taxiway and runway system will be detected by a network of magnetic detectors and will be transferred to Air Traffic Control (ATC) in ASTERIX Cat.10 format, which is an international standard for surveillance data. Extensive testing under operating conditions at the airports of Saarbruecken, Frankfurt (Germany) and Thessaloniki (Greece) and assessment of the technical and socio-economic viability of the system is planned. This will provide the basis for the development of a value added air transport management solution with high economic and operational impact.

2. System Architecture

The system is based on the detection of ferromagnetic objects (e.g. vehicle motors, aircraft components) through their interaction with the earth's magnetic field. The earth's field acts as a biasing magnet, resulting in a magnetic signature (fingerprint) from ferromagnetic objects [3, 4]. The local change of the earth's magnetic field is extremely small – less than 1 µT - but the new detector proposed is able to detect it reliably. This property can be used to detect and locate the objects, either using a single point detector or an array of detectors, small enough to be installed within existing Aeronautical Ground Lighting (AGL) units and cheap enough to be a viable alternative to other technologies.

Based on the experience gained in the early stages of development, the current design of the detector uses a processor with Digital Signal Processing (DSP) functionality with three identical x-y-z channels (magnetoresistive (MR) sensors) for the three-dimensional detection of magnetic field changes. The three-axes-signal is sampled and mathematically filtered inside the controller. Filtering techniques are used to filter out ambient noise (e.g. power line frequency). The detector’s firmware reacts to changes of the sum of the (unsigned) magnitudes of the three axes and currently transmits on demand either the binary states “Field disturbed” and “Field not disturbed” or the three-axes-signal via an RS485 network. Core functions of the detector including sensitivity and noise level have been tested under well-controlled laboratory conditions. In earlier versions, the output shift of the detector caused by the temperature coefficient of the resistance of the MR sensor element and by the change of its sensitivity to magnetic fields was found to have an adverse effect on detector performance. The most promising approach to this problem is a local
compensation of the magnetic field at the sensor’s position. This compensation was realised via a control loop. In the respective setup a compensation coil is deployed, which can produce a magnetic field reverse to the external magnetic field by adjusting the current through the coil. Thus, the magnetic sensing element always works under zero field conditions, eliminating most of the temperature behaviour found in the MR elements. The external magnetic field can be determined from the current needed to create the compensation field in the coil.

As shown in Figure 1, all available ground information obtained from magnetic sensors (i.e. signals from magnetic sensors) is sent to the Sensor Data Fusion (SDF) server, which processes this data in order to extract observations (plots) and target tracks. Observations contain information about the position of the targets, date and time of detection and the size of the targets, while tracking estimates the position and the motion of targets moving on the airport surface.

Figure 1. The architecture of ISMAEL system

Specifically, the SDF server collects magnetic signals (x, y and z components) from all detectors, calculates the overall change of the earth's magnetic field for each detector and compares the field change with a predefined threshold to detect ferromagnetic objects. For determining observations the detectors are initially assigned to predefined sequences of topologically consecutive sensors (“chains”). Within each chain, each sequence of activated consecutive sensors produces exactly one observation. The output of this procedure is a vector (observation vector) containing the position (in ground coordinates) and size (in metres) of observations corresponding to the specific polling cycle. After the estimation of observations, the next step is the generation and update of target tracks. The ISMAEL system requires tracking of multiple targets moving on the airport surface. Thus, a Multiple Hypothesis Tracking (MHT) algorithm is employed [5]. MHT is a deferred decision logic, in which alternative data association hypotheses are formed whenever there are observation-to-track conflict situations. Then, rather than combining these hypotheses, the hypotheses are propagated in anticipation that subsequent data will resolve the uncertainty [6].

The final data (plots or tracks) are then coded in ASTERIX format (a standard developed by EUROCONTROL for the exchange of radar data and extended to any kind of
surveillance data) and thus can be sent to any A-SMGCS system via a UDP/IP communication. The SDF server also provides a real-time display window for the surveillance of traffic on the airport surface. The display shows the target tracks (visualised as crosses) moving on the airport’s map, the state of each magnetic sensor (green in case of activation and red if the detector is not affected) as well as the position and range of each sensor. As a result the system can also serve as stand-alone solution at small or medium-sized airports.

3. Applications

After an extensive user requirements analysis, three applications - Airport Surveillance, Runway Incursion Prevention and Gate Management - were identified as the most promising ones to be targeted by the ISMAEL project in first instance.

In the case of airport surveillance, the controller must have a clear picture of the current traffic situation under any weather condition. The ISMAEL system can be used either as a low-cost alternative to radar at smaller airports or as a “gap-filler” for areas where existing surveillance techniques are not effective.

As regards runway incursion prevention, this application relies on the ability of the system to detect whatever aircraft or vehicle passing a specific location and is a critical task in assuring airport safety. Installing magnetic detectors at each runway entrance intruding vehicles can be detected. Thus, severe accidents can be prevented, like the one in Milan in 2001, when two aircraft collided during take off due to limited visibility under foggy conditions and an inoperative ground radar [7]. In this context, a simple detector system based on the ISMAEL technology could have been used to create a redundant and independent warning system.

Finally, gate management means knowing precisely when a particular gate becomes available, so that the gate can be assigned to a new arrival as quickly as possible and occupation charges can be calculated more accurately. Since ISMAEL’s magnetic field sensors are also able to detect static vehicles, the occupancy status of all equipped parking positions can be identified in real time by the ISMAEL system. In this case the architecture of the ISMAEL system is simpler, since the tracking module of the SDF server is not required.

To investigate the three identified airport applications, detector locations have been defined at both test sites - Frankfurt airport and Thessaloniki airport. Detector positions and numbers have been optimised based on detector features as well as airport application requirements. Communication aspects between detectors and servers have been defined and first system level tests have demonstrated the successful communication among these components. Currently, first tests are being performed with the established detector array at Thessaloniki airport to validate the expected performance. Additional tests with a different detector configuration will also be conducted at Frankfurt airport providing more information about the functionality of the magnetic detectors in a major airport environment. At Frankfurt airport, gate B46 has been selected for evaluating the aircraft parking application. Taxiway S will be used for evaluating runway incursion protection and airport surveillance.
The first prototype system for airport surveillance has been installed at Thessaloniki airport in Greece, where eight magnetic detectors have been installed along the centreline of the main taxiway (Taxiway A) before the stop line of Runway 16. The distance between the magnetic detectors is 30 metres, except for the last detector (under the stop line) that has been installed 20 metres away from the previous one for better monitoring of the stop line. Monitoring the stop line is considered as being of great importance, since it is here that aircraft enter the runway (prevention of runway incursion). The central SDF server computer has been installed at the tower and connected with the magnetic detectors network.

4. Experimental Results

During the initial phase of the project, five early stage detector prototypes were installed at a taxiway of Saarbruecken airport. Initial tests already revealed promising results with respect to the ability of the system to detect both aircraft and ground vehicles. In the case of aircraft, tests revealed that the most detectable parts are those components that are made of ferromagnetic materials, e.g. parts of turbine engines and landing gear [2]. So far, 150 magnetic profiles of 61 individual aircraft and 21 types and relevant subtypes of aircraft have been recorded. A database containing the signals from all tested aircraft is under construction. These data will be subjected to further analysis. Figure 2 shows the magnetic profile of a DHC-8-314 passing the detector.

![Figure 2. The magnetic profile of a DHC-8-314](image)

Further tests at Thessaloniki airport confirmed the promising results of the initial tests in Saarbruecken. It could be demonstrated that cars produce stronger signals than aircraft, due to the larger amount of ferromagnetic materials used and the much lesser distance to the detectors. This significant observation sets the base for an automatic classification of
targets on signal level. Figure 3 (a) illustrates a sequence of activations (eight magnetic detectors) caused by a vehicle moving along taxiway A as shown in the display of the SDF server in Figure 3 (b).

![Graph](image1)

(a)

![Image](image2)

(b)

Figure 3. (a) Magnetic signals from all sensors as the vehicle is passing over them, (b) the SDF display showing the vehicle moving along the taxiway. The light grey detector (detector 4) is activated by the moving vehicle, while the cross indicates its track.

5. Conclusions

The major benefits of the presented solution are derived from the fact that it can be applied as both an efficient low-cost complementary position surveillance technology to be included in existing A-SMGCS and as a stand-alone solution monitoring key points at
airports where full A-SMGCS is not feasible. Furthermore, it is easy to implement due to the small size of the detector allowing for its installation at almost any location, such as in existing ground lighting systems. Unaffected by weather conditions, interferences and shadowing effects, the system provides reliable position, velocity and direction information. In addition, a broad classification of passing vehicles is possible. Based on a passive detection principle, the system does not interfere with other systems like aircraft radios, and it does not depend upon transponders or other equipment within the vehicles. In general the solution is characterised by low energy consumption and a modular architecture that allows for easy system upgrades and extensions accounting for the heterogeneity of different airports.

References


A GIS APPROACH TO OPTIMAL LOCALIZATION OF RESERVED PARKING SPACES FOR URBAN FREIGHT DISTRIBUTION

Giuseppe SALVO¹, Gianfranco AMATO¹, Pietro ZITO¹

Abstract. In the last years, urban freight distribution is increasing its influence on road congestion. The just-in-time supply has raised the number of vehicles used for freight distribution in urban areas. These vehicles cause a significant contribution to congestion and environmental nuisances, such as emission and noise, that impact adversely on the quality of life in urban centres. Moreover, inefficient location and lack of reserved parking space near delivery points, force carriers to the illegal parking. A possible way to improve freight transport efficiency is then to allow freight carriers to have access to the reserved unload/load zones, thus ensuring their availability at the requested time and improving the reliability of deliveries. The aim of this paper is to define a methodology using GIS tool for the optimal localization of reserved parking space for urban freight distribution vehicles in congested areas.

1. Introduction

The growing flows of freight have been a fundamental component of contemporary changes in economic systems at the global, regional and local scales. Structural changes mainly involve manufacturing systems with their geography of production, whereas operational changes mainly concern freight transportation with its geography of distribution. The just-in-time supply and the improved telematics technologies have raised the vehicle’s number used for freight distribution especially in urban areas. Moreover, freight transportation is also a disturbing activity in urban centres.

Freight distribution is carried by vehicles that move on the same streets and arteries used by the private and public transport vehicles. These vehicles make a significant

¹ Department of Transportation Engineering, University of Palermo, Viale delle Scienze al Parco d’Orleans, Edificio 8 – 90128 Palermo, tel. (+39)0916459931, E-mail: {salvo, amato, zito}@ditra.unipa.it
contribution to congestion and environmental nuisances, such as emissions, noise, and so on, that impact adversely on the quality of life in urban centres.

Some measures optimizing urban freight distribution could have some positive effects on congestion as well. In fact it is possible:

- To localize Urban Distribution Centres (UDC), aiming to improve load factors in delivery vehicles, since the city terminal can be accessed by larger trucks, and the goods are then transferred to smaller vans for their final delivery, which now have to cover smaller distances (Thoma, 1995);
- To localize parking lots, especially near congested commercial areas of the city, usually used as city terminals, with easy access from ring roads, space for vans and trucks, and possibly transhipment devices and warehousing capabilities (Tellus-Berlin, 2004);

in order to decrease vehicle’s number in road network and to improve delivery operation with reserved parking spaces. In fact, lack of parking spaces forces double-parking of freight vehicles, causing congestion problems especially on narrow streets. The optimal location and an advanced management of these reserved slots for freight parking could improve freight distribution and traffic condition on urban areas.

2. The model used to localize reserved parking spaces for freight distribution

The optimal location of reserved parking spaces should examine freight demand (density and typology of commercial activities) and several economic principles (Drezner, 1995). The optimal location problem could be solved to take advantage of FLP (Facility Location Problems) as follow: given existing facilities location and the traffic flow between these facilities, to determine the location of a new facility in order to minimize the total cost of transportation (distance multiplied to amount of transported goods).

In restricted zones where there are limited region of solution, like urban areas, an optimal model works fine using Euclidean distance. The Euclidean distance \( d(A, P) \), between an existing facility \( A \) with coordinates \((x, y)\) and the new facility \( P (a,b) \), is defined as follow:

\[
d(A, P) = [(x - a)^2 + (y - b)^2]^{1/2}
\]  

(1)

When there are many facilities it is possible to distinguish:

- \( N \) is the number of different commercial activities (clothes, food, etc.) in a specific zone
- \( M \) is the total number of commercial activities for each typology
- \( A_i(x_{ij}, y_{ij}) \) is the position with coordinates \( x_{ij} \) e \( y_{ij} \) for \( i = 1, \ldots, N, j = 1, \ldots, M \);
- \( f_{ij} \) is the freight flow required by any commercial activity \( A_i(x_{ij}, y_{ij}) \) \( \forall i = 1, \ldots, N, j = 1, \ldots, M \);
- \( r_{ij} \) is the transport, load and unloaded costs for freight unit from the new facility to any commercial activity;
- \( w_i \) is a factor to take in account the average time frequency for each typology of commercial activity \( \forall i = 1, \ldots, N \).
It is possible to determine the optimal location minimizing the total cost of the transport as follow:

\[
\min \sum_{i=1}^N \sum_{j=1}^M w_i r_{ij} \left[ (x_i - a)^2 + (y_i - b)^2 \right]^{1/2}
\]  

(2)

The minimization problem described by the eq. [2] it is a problem not bound, that can be solved through the simple calculation of the derived partial (Winston, 1987). Considering the derived partial in comparison to \(a\) and \(b\), equalizing to zero and solving in comparison to \(a\) and \(b\), we obtain:

\[
a = \left( \sum_{i=1}^N \sum_{j=1}^M w_i f_{ij} x_{ij} / d_{ij} \right) \left( \sum_{i=1}^N \sum_{j=1}^M w_i f_{ij} r_{ij} / d_{ij} \right)^{-1}
\]  

(3)

\[
b = \left( \sum_{i=1}^N \sum_{j=1}^M w_i f_{ij} y_{ij} / d_{ij} \right) \left( \sum_{i=1}^N \sum_{j=1}^M w_i f_{ij} r_{ij} / d_{ij} \right)^{-1}
\]  

(4)

To note that the factor \(d_{ij}\) is the Euclidean distance among the position of the commercial activity \(A_{ij}\) and the parking space \(P\) (eq. 1). This is a classic gravitational model.

Because the parking location is unknown, it is possible to use an iterative processes converging to an optimal solution, as the following steps:

- **Step one:** to determine \((a, b)\):

\[
a = \left( \sum_{i=1}^N \sum_{j=1}^M w_i f_{ij} x_{ij} / d_{ij} \right) \left( \sum_{i=1}^N \sum_{j=1}^M w_i f_{ij} r_{ij} / d_{ij} \right)^{-1}
\]  

(5)

\[
b = \left( \sum_{i=1}^N \sum_{j=1}^M w_i f_{ij} y_{ij} / d_{ij} \right) \left( \sum_{i=1}^N \sum_{j=1}^M w_i f_{ij} r_{ij} / d_{ij} \right)^{-1}
\]  

(6)

- **Step two:** to calculate the distance \(d_{ij}\) using the eq. [1], knowing the previous \((a, b)\) values;
- **Step three:** using eq. [3] and [4] to obtain new \((a, b)\) values;
- **Step four:** repeat steps two to four until the difference of \((a, b)\) in step three and two is arbitrarily changed more than a small quantity.

This model uses GIS data, since it is easier to achieve location data of commercial activity, classified for typology and freight flows. Conversely to the normal Facility Location Problems, the proposed model takes into account, in its mathematical formulation, a factor related to the temporal frequency of average loaded/unloaded flows of the commodity classified for any typology of commercial activity.
3. A GIS support to localization model

Geographic Information System is able to process different kinds of database and also allows to manage territorial information. Moreover, the structure of database allows the integration in complex mathematical model providing to effective decision support systems. A GIS application is proposed to determine the optimal location of parking space taking into account the cyclical goods demand of commercial activities in the study area.

![Figure 1 – Study area](image1)

![Figure 2 - Optimal location of reserved spaces](image2)
In particular, each record holds the location (the coordinates X and Y to determine the
distance from the reserved parking space), typology, supply of goods in terms of frequency
and quantity of the commercial activity.

The application was applied to a study area with various commercial activities of
Palermo.
A survey was carried out to determine the goods demand and the transport supply by
knowledge of geometric and traffic characteristics of the road network.

The GIS model has provided the coordinates of reserved parking place considering a
circle area with radius of 40 m as maximum distance that the operator is available to run up
to the commercial activity.

4. Conclusions

The purpose is to define a methodology using GIS tool for the optimal localization of
reserved parking space for urban freight distribution vehicles in congested areas.

The maximum distance that the operator is available to run up to the commercial
activity is about 40m.

The GIS to support the optimal location of the reserved parking space of goods
distribution allows:

− to reduce the number of vehicles for illegal parking, fluidifying traffic conditions;
− to improve the freight distribution and the load/unload operations.

References

[1] Jesús Muñuzur et al. (2005), Solutions applicable by local administrations for urban
logistics improvement, Cities, Vol. 22, No 1: 15-28
delle merci, 8ª Conferenza Utenti ESRI, Roma
milanese, Franco Angelo, Milano
Mobilità e Traffico Urbano Anno XII N. 2 – Febbraio 2001, Milano
e Traffico Urbano Anno XII N. 8 - Settembre 2001, Milano
urbano delle merci, Working paper del Politecnico di Milano Dipartimento di
Architettura e Pianificazione, Milano


START-UP LOST TIME AT URBAN TRAFFIC LIGHT JUNCTIONS

Sergio D’ELIA¹ Domenico Walter E. MONGELLI²

Abstract. The present work aims at an analytical evaluation of vehicular headway and therefore of the time loss at setting off and of any non-applicability of the formulations still in use today. The results obtained show a behaviour and numerical values different from the existing formulations used today, also in Italy, when calculating traffic light cycles; in fact, shorter headway and different trends among the leading vehicles when crossing the stop line were recorded.

1. Introduction

In general terms vehicle flow in the urban context is characterized by continual variations of motion with alternating acceleration and deceleration, with a frequency depending on both the entity of the circulating flow and on the fragmentation of the urban road network. In practice the “stop and go” phenomenon is found with regularity at traffic light junctions that produce rhythmic interruptions to the flow and the formation of traffic queues waiting for the green light; when this light shows the queuing vehicles set off one after the other according to the perception and reaction time of each driver, leading therefore to the inevitable formation of “start-up lost time”. This phenomenon is important because it occurs systematically at each of the entry points at traffic light junctions and is cyclical all day long, occurring at each green light phase and being repeated hundreds of times throughout the day. This situation involves at least the first six vehicles for each branch, that is, almost 30% of the vehicles that manage to pass at each cycle. Given the above-mentioned situation the importance of a specific study on start-up lost time is clear, since it notably affects vehicle flow, the number of vehicles crossing, the level of service of each of the branches making up the traffic light junction and the degree of pollution at the node. The importance of a study on start-up lost time is revealed also by the influence that this phenomenon has on vehicle flow and on the level of service of the entry; these time losses moreover, as is known, condition the calculation of the gap to the amber light.

¹ Full Professor University of Calabria – Di.Pi.Ter.
² PhD Student University of Calabria – Di.Pi.Ter.
2. State of the art

The phenomenon was studied for the first time in the United States over 30 years or so ago, with experimental surveys at traffic light junctions in various operating conditions; the results, illustrated by the Highway Capacity Manual, led to the establishment of a well-defined behavioural law and to specific numerical values of the mean vehicular headway and therefore of the start-up lost time at a generic entry of a traffic light junction. According to what was reported in the first edition of the H.C.M., the phenomenon of time loss especially regards the first six queuing vehicles waiting at the stop line, with variable and quite considerable delays decreasing progressively with respect to the headway of the successive vehicles. The rhythm of vehicles after the sixth was believed to be constant and equal to 2.1 seconds. In the last edition of the manual the phenomenon is described analogously, with identical behavioural laws, but the numerical values of fixed vehicular headway have been eliminated, pointing out user behaviour only graphically. The trend proposed by the H.C.M. is also taken into account in the text “Traffic Flow Theory and Control”

Other Authors have showed interest in the phenomenon and, in the absence of experimental surveys, have more simply suggested a constant numerical value independent of the operating conditions and geometry of the junction in their calculations for the determination of the traffic light cycle. For example, in the notebooks of the AIITT and specifically, in Notebook no. 1 entitled: “Basic Techniques of Traffic Light Regulation” by Ignazio Morici establish a fixed start-up lost time at 3s.

3. Schematization of the phenomenon

The vehicular headway phenomenon can therefore be schematized by means of a behavioural model, as illustrated by the H.C.M. The fundamental variables can therefore be defined that are represented by headway “$h(n)$” that a queuing vehicle must have to cross the stop line of a green light junction in the presence of a theoretical queue of vehicles moving through a junction and equally represented by “$\Delta t_i$” the start-up lost time of vehicle $i$ “start-up lost time”. This phenomenon is highlighted by the fact that there is a change from a null velocity of the queuing vehicles to a velocity $V$, carried out by the successive vehicles in that traffic situation. The influence of the “start-up lost time” decreases progressively to a condition of “$\Delta t_i = 0$” where the function “vehicular headway” is stabilised and settles to mean values equal to “$h$”. The headway trend can be schematized in two successive phases: in the first, the headway of single vehicles is influenced by time loss “$\Delta t_i$” and the trend is represented by an almost parabolic function with vertex $V=(n,h)$ and with increasing velocity (figs.1; 2);
The surveys carried out and the behavioural model of the start-up lost times

The behaviour model that best describes the trend of the start-up lost times is certainly influenced by a high number of variables, often closely correlated among themselves or having little importance in the definition of the model.

The variables can be: a) geometrical variables (L width of the entry, P gradient of the entry, Rc radius of the bend), b) physical variables (n queuing vehicles, V velocity of crossing, tpr perception and reaction time, Pt percentage of heavy vehicles), c) behavioural variables (including day of the week, atmospheric conditions, h time of day). Given the
high number of factors to examine, it is extremely difficult to define a model that takes into account each of the variables in play without having to employ a very high number of experiments and results analyses. Therefore, targeted survey campaigns were conducted, at various traffic light junctions to take into account each of the above-listed variables and to look for systematic behaviour. The experiments were carried out in some Italian towns according to the usual criteria and the modalities of the theory and technique of circulation. A representative sample was obtained of vehicular headway and they are shown as an example, in (figs. 3, 4) For a more complete representative framework the mean value of the instantaneous velocity of the single queuing vehicles was also recorded, also measured at the instant of crossing the stop line.

Figs. 3, 4 Survey at Rende Entry 1 8.30-9.30 a.m., 6.00-7.00 p.m.

The samples observed are representative of the urban context of this study. The trends researched were subsequently compared to the observed trends. The accuracy of the theories formulated was verified by a statistical test and a significantly good fit was found.

The trend found, as in the American experience, presents two phases: a non-linear and a constant one. The notable difference of the first phase is that while the trend proposed by the H.C.M. is of the parabolic type \( y = an^2 + bn + c \), the observed one, instead, has a fifth grade trend of the type: \( y = a_5n^5 + a_4n^4 + a_3n^3 + a_2n^2 + a_1n + a_0 \).

The second phase is almost constant (fig. 6), undergoing a slight influence of the number of queuing vehicles, in fact, when the number of vehicles exceeds the fourteenth there is a negative correlation (fig. 5). It can be noted by studying all the elaborated graphics, that the first queuing vehicle crosses the stop line with a time loss lower than that found for the second vehicle, a sign that the driver pays much more attention to the colour of the traffic light and starts his vehicle up more quickly than the second. Probably the second driver does not look directly at the light but at the preceding vehicle and starts moving only after the first has moved off, therefore making him slower than the car ahead. It can also be noted that time losses are only relative to the first 5 queuing vehicles and already not to the sixth as in the American experience; in fact, the sixth vehicle has minimal headway. Moreover, it can be noted that the successive vehicles improve their performance and they arrive over the stop line with progressively decreasing headway, even though with minimum deviations. The equations for the two phases are the following:

**Phase 1:** Fifth grade non-linear trend:

\[
h(n) = 0.0051n^5 - 0.1022n^4 + 0.7901n^3 - 2.8956n^2 + 4.6246n + 0.289
\]

**Phase 2:** Constant/linear trend:
5. Test of the validity of the models

To verify the validity of the results of the study obtained in Calabria, a further experimental campaign was conducted at junctions in the city of Taranto and the results were successively compared. The behaviour of Apulian drivers proved to be identical to their Calabrian counterparts, again showing a progressively greater headway for the second vehicle passing the stop line, with values decreasing to the fifth vehicle and almost constant values for the successive vehicles. In particular for the latter, when there are many queuing vehicles (more than 14) there is less aggressive behaviour and constant headway. Instead, when there are less than 14 queuing vehicles there is a slight improvement in performance and slightly decreasing headway. By means of a statistical comparison it was found that despite the territorial diversity, but in the same urban context, the values of the headway curve are comparable with a significantly good fit, in fact, in the definition of the $\chi^2$ test values close to zero were reached, confirming the theories. For brevity the specific graphs are not included here, but it was found that the behaviour was completely analogous to that previously reported (fig. 5, 6).

6. Conclusions

The results achieved through specific experiments show that there is a more aggressive and attentive driver behaviour in Italy with lower time losses compared to those found in the United States. The numerical values of the headway are in fact much lower both in the first phase of acceleration of the first five vehicles, and in the second phase of passage of the successive vehicles that still have modest accelerations but constantly keep a very reduced headway.

Specifically, if the headway defined by the H.C.M. is compared to those found today in the two surveyed towns, there is a percentage variation of 15.3%.

This percentage is therefore to be considered as time saved that is really available to the vehicles that have to cross a branch of the junction during the green light phase. This advantage is concretized, taking into account the results obtained, in the calculation of the
cycle duration, in the two following aliquots. A first advantage in the calculation of the
effective green light time since a lower time loss should be subtracted from the real time. A
second advantage is also gained in counting of the vehicles that effectively cross in the
green light phase, since the real headway between the vehicles is about 1.7 – 1.8 vehicles
and not 2.1 as in the previous formulations. This result is interesting, because it shows that,
at parity of green light duration, in the Italian situation and in the current operating
conditions, there is a greater vehicle flow, with greater saturation flows at various levels of
service. From the surveys there also emerges a further interesting observation that there are
no behavioural differences between daytime and evening situations; in fact the vehicular
headway and the velocity is identical in both cases.

This research will be extended to other traffic light junctions in an attempt to find
whether there is different behaviour in the presence of specialist manoeuvres, or in
situations of less than normal lane width.

Bibliographical references

[1] Cantarella G. E., Introduzione alla tecnica dei trasporti e del traffico con elementi di
intersections, URBAN TRANSPORT 96, 2-4 October 1996, Barcelona, pp. 573-582,
THE INFLUENCE OF ADVANCED TRAVELLER INFORMATION SYSTEMS (ATIS) ON PUBLIC TRANSPORT DEMAND

Gianfranco AMATO(1), Salvatore AMOROSO(1), Pietro ZITO(1)

Abstract. Advanced Traveller Information Systems (ATIS) include a broad range of advanced computer and communication technologies. These systems are designed to provide transit riders pre-trip and real-time information, to make better informed decisions regarding their mode of travel, planned routes, and travel times. ATISs include in-vehicle annunciators and displays, terminal or wayside based information centres, information by phone or mobile, and internet. A Stated Preference survey has been carried out in order to know the preferences of public transport’s customers related to different ATISs.

1. Introduction

In the last years, the application of telematics technologies in transportation field has developed many innovative services and management techniques. Now transportation companies are able to control in real-time the service provided and to take the right decision when any unpredictable event happen. Knowing exactly and immediately the position of each vehicle, the relative route and the current traffic condition it is possible to predict the average arrival time on bus stop. This information is very useful for the customers because they might be informed in real-time when their bus is expected to arrive at interested bus stop. There are several systems and different typologies of information able to reduce the irregularity of public transport service, especially in urban area, where the public transport service is influenced by private mobility. Actually, public transport companies are not able to afford the necessary cost to implement Advanced Traveller Information Systems without to know the effects on customers. Moreover it is hard to understand which ATIS and type of information is preferred by customers.

Under the DRIVE project was studied the effects of pre-trip information on travel behavior using a stated preference approach in Birmingham and Athens. The study revealed firstly, a requirement for multi-modal pre-trip travel information even if the sample studied were regular car users, and that the quantity and type of pre-trip information requested by

---

1 Department of Transportation Engineering, University of Palermo, Viale delle Scienze al Parco d’Orleans, Edificio 8 – 90128 Palermo, E-mail: {amato,amoroso, zito}@ditra.unipa.it
travelers depends on a range of personal, journey related, contextual and national factors. Moreover, it emphasized the importance to travelers of the timeliness and relevance of the provided information especially when relevant networks incidents happens [11].

Abdel-Aty M.A. et al. [2] studied the effects of ATIS on commuter propensity to use transit with a computer-aided telephone interview conducted in Sacramento and San Jose, California. The results indicated that about 38% of the respondents who currently do not use transit might consider public transport if the appropriate information is available. Moreover, using an ordered probit model produced results that show the significant effect of several commute and socioeconomic characteristics on the propensity to use public transport.

Nijkamp et al. [8] conducted a survey before and after the application of ATIS in the city of Birmingham and Southampton (QUARTET and STOPWATCH project respectively). Because of early development of QUARTET project their result was considered inconsistent, whereas in the city of Southampton the survey revealed raise of using public transport, especially in study and leisure trips, and mobility optimization of people in choosing the mode and route able to reduce travel time.

A methodology was developed [7], aiming to understand the effect of real-time information on bus stop, under three different methods to forecast bus stop arrival time: 1) static information, 2) real time information update using historical data, 3) real time information using data coming from Automatic Vehicle Location (AVL) system. Measuring the difference between predicted and effective waiting time when each people approach bus stop the third methods revealed more reliability.

To address the potential of ATIS was developed an innovative stated preference design aiming to investigate whether real time information would increase the acceptance of public transport [1]. Using ordered probit modeling technique was showed that commute by public transport, income, education were among the factors that contribute to the likelihood of using public transport given information was provided, and in general was revealed a promising potential of ATIS in increasing the acceptance of bus as a commute mode.

The impacts of socioeconomic benefits and technical performance of telematics application in the city of Helsinki was studied [6]. The system provided several public transport telematics such as real-time passenger information, bus and tram priorities at traffic signals and schedule monitoring. A before and after field studies, an interview and survey, a simulation, and socioeconomic evaluation indicated 40% reduction of delay at signals, improving on regularity and punctuality of public transport, reductions of 1% to 5% in fuel consumption and exhaust emissions. Moreover the information systems were regarded very positively, and, in particular the information displays on stops were considered necessary.

2. ATIS scenarios

Actually there are several ATISs providing different typologies of information on different stages of trip: at home/office (pre-trip), at bus stop, on public transport vehicle (on-trip). In the past, for high costs of ATISs and unknowing of their real effects, the main public transport companies were reluctant to adopt the information systems. Moreover the transport service in urban areas suffers from cyclical variation of traffic conditions, thus it
is very difficult to provide reliable service information especially on not reserved routes. In the last years, the evolution of informatics, telecommunication and the decrease in prices of these technologies induced public transport to invest in AVL and ATIS trying to provide more competitive and efficient service in order to attract more customers. These processes have also several social benefits: less congested cities, more easier and quick transport, and finally more sustainable life.

In this paper, a stated preference survey was built in order to know which combination of ATIS, typology of information and willingness to pay are preferred by Palermo’s citizens. At investigation time no real time information was provide by Local Public Transport Company (AMAT). The survey was conducted in the late 2005 using mail-back self-completion questionnaire.

Because of the number of different ATIS and different typologies of information, the first step of building the questionnaire was to identify the system, the typology and the costs of information to analyze. Considering to analyze pre-trip and on-trip information, the importance of real-time information, and then willingness to pay by the customers, we identified the following levels of any attribute:

<table>
<thead>
<tr>
<th>Information systems</th>
<th>Information typology</th>
<th>Costs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Display at bus stops (on trip real time information)</td>
<td>Waiting time</td>
<td>Free of charge</td>
</tr>
<tr>
<td>Web-site (pre-trip real time information)</td>
<td>Travel time</td>
<td>30 cent/€ for single enquire</td>
</tr>
<tr>
<td>Short Message Systems (on-trip and pre-trip real time information)</td>
<td>Route choice</td>
<td>20% more expensive tickets</td>
</tr>
<tr>
<td>Call center (on-trip and pre-trip real time information)</td>
<td>General information</td>
<td>5€ of monthly subscription</td>
</tr>
</tbody>
</table>

The full factorial design provides 64 different scenarios (three attributes four levels each), thus assuming the irrelevance of interactions between attributes we reduced, accordingly to Kocur et. al. [4] catalogue, in a 16 different scenarios (fractional factorial design). We divided these 16 scenarios in two groups in order to avoid fatigue effect in responding as the literature suggests. We asked firstly to select the most four preferred scenarios, then to order these selected scenarios from most to less preferred, and finally to evaluate the influence of each scenario, in an ordinal scale, in deciding to use public transport. Some other information was collected: frequency of use of bus and private vehicle, evaluation of the importance of some factors in deciding whether or not travel using private and public transport, some transport habits (frequency, reason, purpose and maximum distance traveled of following transport modes: walking, bicycle, motorcycle, car as passenger and as driver, taxi, bus, tram underground), knowledge and use of real time information, and some socioeconomic information.

3. Ordered probit model

When the dependent variable has a natural order and assumes more values, by literature the ordered probit model should be more appropriate (McKelvey and Zavoina, 1975).

The dependent variable is expressed by:

\[ y_i^* = \beta' x_i + \varepsilon \]
Let $y^*$ be dependent variable coded by 0, 1, 2, ..., $J$, $\beta$ the vector of coefficients, $x_i$ the vector of the attributes, and $\varepsilon$ is the random term, normally distributed $N(0,1)$. Whether the observation of the dependent variable is as follows:

\[
\begin{align*}
  y &= 0 \quad \text{if } y^* \leq \mu_0 \\
  y &= 1 \quad \text{if } \mu_0 \leq y^* \leq \mu_1 \\
  y &= 2 \quad \text{if } \mu_1 \leq y^* \leq \mu_2 \\
  \text{.....} \\
  y &= J \quad \text{if } y^* \geq \mu_{J-1}
\end{align*}
\]

where $\mu_k$ are threshold values that with together $\beta$ parameters have to be estimated by calibration process. The ordered probit model provides the thresholds which indicate the levels of users’ preference related to ATISs, without to make assumptions about the differences between the categories of the dependent variable. The threshold values have to be ordered from the lowest to the highest. The probabilities are provided by normal distribution:

\[
\begin{align*}
  \text{Prob}[y = 0] &= \Phi(-\beta x) \\
  \text{Prob}[y = 1] &= \Phi(\mu_1 - \beta x) - \Phi(-\beta x) \\
  \text{Prob}[y = 2] &= \Phi(\mu_2 - \beta x) - \Phi(\mu_1 - \beta x) \\
  \text{.....} \\
  \text{Prob}[y = J] &= 1 - \Phi(\mu_{J-1} - \beta x)
\end{align*}
\]

Therefore, the probabilities that $y_i$ falls into the $j^{th}$ category is given by:

\[
\text{Prob}[y_i = j] = \Phi(\mu_j - \beta x_i) - \Phi(\mu_{j-1} - \beta x_i)
\]

The ordered probit model includes two sets of parameters. The constant and other threshold values indicate the range of the normal distribution related to specific values of explanatory variables. The $\beta$ parameters represent the effect of changes in any explanatory variable on a given scale. These parameters let us understand the relative importance of any variable in the choice process of ATISs. It should be noted that the signs of coefficients in the ordered probit model are particularly important since they provide different effects on the probabilities of the ordered categories.

### 4. Survey results

We analyzed data using the ordered probit model by Limdep® software (see the following figure). We found that cost and info are attributes more important, whereas system is less important, highlighting a good significance of $P(|Z|)$ values.
Moreover to confirm the relative importance of any attribute for each level it was used the Conjoint module of SPSS® statistical analysis software. The results are summarized on the following table:

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Standard Error</th>
<th>[b/St.Er.]</th>
<th>[F]</th>
<th>[Z]</th>
<th>[Z]</th>
<th>Mean of X</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>.209453298</td>
<td>.36553165</td>
<td>.769</td>
<td>.4421</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SYSTEM</td>
<td>.5962786597E+01</td>
<td>.27349998E-01</td>
<td>2.130</td>
<td>.0313</td>
<td>1.500000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>INFO</td>
<td>.2688402168</td>
<td>.32614291E-01</td>
<td>8.259</td>
<td>.0000</td>
<td>1.155045</td>
<td></td>
<td></td>
</tr>
<tr>
<td>COST</td>
<td>.6333367316</td>
<td>.43990888E-01</td>
<td>14.397</td>
<td>.0000</td>
<td>1.2344828</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Moreover to confirm the relative importance of any attribute for each level it was used the Conjoint module of SPSS® statistical analysis software. The results are summarized on the following table:

<table>
<thead>
<tr>
<th>Importance</th>
<th>Utility</th>
<th>Factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>38.97</td>
<td>.0611</td>
<td>INFORMATION</td>
</tr>
<tr>
<td>25.45</td>
<td>.3265</td>
<td>SYSTEM</td>
</tr>
<tr>
<td>35.58</td>
<td>.8730</td>
<td>COST</td>
</tr>
</tbody>
</table>

Pearson’s R = .968                      Significance = .0000
Kendall’s Tau = .644                    Significance = .0003

SUBFILE SUMMARY

<table>
<thead>
<tr>
<th>Averaged</th>
<th>Importance</th>
<th>Utility</th>
<th>Factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>25.45</td>
<td>.3265</td>
<td>SYSTEM</td>
<td></td>
</tr>
<tr>
<td>38.97</td>
<td>.0611</td>
<td>INFORMATION</td>
<td></td>
</tr>
<tr>
<td>35.58</td>
<td>.8730</td>
<td>COST</td>
<td></td>
</tr>
</tbody>
</table>

It is possible to note that the respondents attribute most importance to the type of information (38.97%) followed by the cost (35.58%), and are less interested in which system provides the information (25.45%). The software provides also the part-worth for each factor level. These part-worth scores indicate the influence of each factor level on the respondent's preference for a particular combination. They are computed by the procedure through a set of regressions of the rankings on the profiles. Since they are all expressed in a common unit, the part-worth scores can be added together to give the total utility of a
combination. The part-worth scores point out that the “ideal” scenario of information could be represented by bus stop display, providing waiting time without any additional cost. The higher value of Pearson’s R and Kendall’s τ indicate that the model fit well the data.

The rating of the influence that the chosen scenarios have in deciding to use transport pointed out that about the 15% of the respondents in presence of bus stop display providing free of charge expected arrival time of the bus are willing to increase their use of public transport.

5. Conclusions

A stated preference survey was conducted in Palermo. The study’s main objectives were to investigate which combination of ATIS, real-time information provided, and costs would have influence in deciding to use public transport by the citizens. For this purpose an ordered probit model was calibrated and a conjoint analysis was carried out. Our analysis showed that cost and info were the most important variables. Furthermore the signs of ordered probit model coefficients confirm that the employment of ATISs increase the probability of using transit and its perceived quality.

The combination of display at bus stop providing expected real time information free of charge was preferred by the respondents and in particular 15% of them are willing to consider increasing their use of public transport if the previous scenario would be available in Palermo.

References


THE IMPACT OF WEATHER ON URBAN TRAVEL DEMAND IN THE NETHERLANDS

Eric VAN BERKUM(1,2), Wendy WEIJERMARS(1) and Anton HAGENS(1) *

Abstract. In this paper the impact of weather, in particular rain, on urban travel demand is presented. The research was performed in the Netherlands where in an urban environment next to the car and public transport, the bicycle is a major transport mode. Three different research approaches were followed, i.e. the Dutch National Mobility Survey (OVG) and traffic counts from an urban traffic control and information system were matched with weather data, and further a survey was held, where more specifically the impact of weather on travel behavior was recorded.

1. Introduction

In general, bad weather leads to a reduction of traffic flows on highways. In Australia Keay and Simmonds [6] found a reduction of 1.3% on rainy days compared to dry days. They found that in general flow decreases by 0.08% per day for each mm precipitation. Al Hassan and Barker [4] found similar, yet slightly higher effects for Scotland as did Edwards [2] for Wales, Chung [1] for the Tokio region and Hogema [5] for the Netherlands. In fact the reducing effect of rain seems most eminent in the weekend or on holidays (a reduction of 3% respectively 9%). Also fog seems to have a larger effect than rain.

Extreme conditions such as snow may have a significant larger impact. In some instances a reduction of 50% was found (Knapp and Smithson, [7] and Knapp et al, [8], Goodwin [3]). Keay and Simmonds [6] found further that the impact is largest for a combination of wet and cold, yet this was in Australia.

The impact of bad weather on travel demand in an urban environment is expected to be different from the highway system. In general trips are shorter, more diverse, and more modes are being used. In the Netherlands the modal share of the bicycle for all trips is 29% where in the UK this is 2% [9]. It is expected that bad weather may result in a modal shift from bicycle to car and public transport. Also bad weather may

*(1) Center for Transport Studies, University of Twente, PO Box 217, 7500 AE, Enschede, the Netherlands, e.c.vanberkum@utwente.nl; (2) Goudappel Coffeng, PO Box 161, 7400 AD, Deventer, the Netherlands
affect trip generation. In Sweden it was found that after illness extreme bad weather is the second important reason not to undertake an intended trip. Also shopping trips in many instances were found to be highly impulsive. Therefor it is expected that in an urban environment bad weather will yield an increase in car traffic, as opposed to the highway situation, where a decrease was recorded.

2. Approach

To research the impact of bad weather on urban travel demand three different approaches were taken.

**OVG** The *Onderzoek VerplaatsingsGedrag* (OVG) is an annual national survey from the Dutch Central Bureau of Statistics. For each year up to 100,000 respondent report their daily travel behavior. For the years 1995 to 2001 the data from these surveys are matched with national weather data, and further analyzed.

**Urban traffic volumes** In the city of Almelo at signalized intersections traffic flows are recorded on a 5 minute basis using induction loops. These data are matched with local weather data, and further analyzed.

**Survey** To get a deeper understanding of the impact of bad weather on travel behavior in an urban environment, a specific internet survey was held. Self-conducted inquiry is also part of the study-data.

2.1. Weather data

The Royal Dutch Meteorological Institute (KNMI) has 10 main stations scattered over the Netherlands, of which 9 were used. For each day data is recorded on temperature (mean, minimum, maximum), hours of sunshine, degree of cloud cover, visibility, humidity, precipitation (amount, duration), wind (mean speed, max hourly speed, max wind gust, prevailing direction), and air pressure. Since it is expected that season and precipitation are the most important factors that influence travel behavior, the impact of these factors is investigated.

For the analysis of the travel survey and traffic data, days were classified into wet days and dry days. Rain can occur in several different varieties, e.g. a short and severe cloud-burst may produce the same amount of precipitation as an all-day lasting drizzle. Both the duration (hour) and the intensity of the rain (mm/hour) might influence travel behavior. Therefore both characteristics are taken into account. For quantification of these criteria several items must be taken into account. The level of discernment must be high enough, i.e. a relatively large group of 'wet days' is required which can be realized by setting low values for the criteria. On the other hand, to make sure that only days with serious precipitation are labelled 'wet' the requirements for the criteria must not be too low. A compromise is made by labelling days with a rain duration of over 2 hours that have an average rain intensity of over 0.5 mm/hour.
as wet days. Precipitations in other forms than rain occur only on a rare basis in The Netherlands and are therefore not explicitly taken into account. For snow, if any, the same criteria are used as for rain to classify the amount of precipitation. By using this classification for precipitation, three groups of days can be distinguished, i.e. dry days (with no precipitation at all), wet days (over 2 hours of precipitation at an average of over 0.5 mm/hour) and other days (days with precipitation, but the criteria for 'wet' are not met). As a consequence of this classification, 50% of the days are labelled dry and about 25% of the days are labelled wet. The classification of days is done on a province basis, i.e. for each province in the Netherlands, the most appropriate weather station is selected.

For the second part of the research also weather data on a six hour level is used (http://www.WeerOnline.nl). These data are collected at a weather station that is located at a distance of approximately 20 kilometers from Almelo. The two data sources are combined to obtain an estimate of rain intensity for 6 hour periods. For the classification into dry and wet periods, the same criteria as for daily data were used. Approximately 75% of the periods were labelled dry, whereas about 5% was labelled wet.

2.2. OVG

The Onderzoek Verplaatsingsgedrag (OVG) is a national survey from the Dutch Central Bureau of Statistics (www.cbs.nl). It has been held annually since 1978. For this project the years 1995 and 2001 were used. The number of respondents is about 1% of the Dutch population. Of each respondent personal data are recorded, such as age, gender, education, income, car ownership, profession, residence etc. Further, for a specific day all trips, with departure time, arrival time, origin, destination, mode, purpose, trip length etc. are recorded. It is not recorded whether a specific trip is an urban trip or not. As an approximation only trips with a length under 7,5 kilometer were used in the analysis. The OVG dataset with trips under 7,5 kilometer is matched with weather data from the KNMI on the basis of the trip origin location, in particular the province where the origin is located. For dry, wet and other days the average number of trips and the modal split is computed. The same is done for winter (December, January and February) and summer (i.e. June, July and August) periods. Since the average numbers of trips are based on a relatively small number of people (47 persons per day per weather station), the number of trips cannot be assumed to be normally distributed. Therefore, non-parametric tests (Wilcoxon and Mann-Whitney) are used to test whether the differences between dry and wet days are significant.

2.3. Urban Traffic Volumes

In the second part of the research, traffic count data from the city of Almelo is used to analyze the impact of weather on urban traffic volumes. Almelo is a medium
sized city in the east of the Netherlands. Traffic data is collected by inductive loop detectors at signalized intersections. The data from these loop detectors is processed into traffic flow and occupancy measurements and stored in the database of a traffic information system called Viacontent (www.viacontent.nl). Data validity is checked by means of basic quality checks using minimum and maximum flow thresholds and, when possible, by means of a quality control algorithm based on the principle of conservation of vehicles [10]. For this analysis four signalized intersections are selected. These intersections are located close to residential areas and are used by short distance traffic as well - the proportion of short distance (≤ 7.5 km) traffic is between 50% and 75%-, making them sensitive to substitution of bike trips by car trips.

Traffic data from these intersections is combined with the weather data discussed in section 2.1. Matched pair analysis is used to analyze the differences in traffic flows between dry and wet periods (Andrey et al, 2003). Traffic volumes on each wet period are compared with data from exactly one week prior to or after the measurement (or two weeks prior or after if necessary). For all periods (night, morning, afternoon and evening) and different types of days, the percentage difference between wet and dry periods is determined by taking the median for all detectors for which traffic data is available. The morning is the most important period for analysis, since decisions for a transport mode in the morning have consequences for transport modes later that day.

2.4. Survey

The first and second part of the research provide aggregated information on the changes in the number of trips, the modal split and traffic volumes as a result of rain. They do not provide information about other travel behavioral changes (trip timing, destination choice) and provide no insight in individual decision making. A survey is held to obtain more insight into these individual decisions and other adaptations in travel behavior. Also, the effect of weather forecast on travel behavior is investigated.

In the survey the respondents were asked about their normal daily trip pattern for short trips (number, purpose and mode choice). Subsequently they were asked about the influence of different weather variables and weather forecast on their travel behavior. Finally, some personal characteristics of the respondent (gender, driving license, age, vehicle possession and social activity) make also part of the survey.

The survey was held through the internet. The total number of useful and fully completed questionnaires is 114. Many respondents are employees of the municipality of Almelo.
3. Results

3.1. Impact on short trips using the OVG

Using the OVG data the influence of precipitation and the season on the amount of (short) trips and the modal split was analyzed. It appeared that on wet days on average 2.9% less trips are performed, and 3.4% less short trips. In winter 3.4% less trips are performed than in summer, and 2.3% less short trips. These results are statistically significant at a 0.95 confidence level. Also the modal split is significantly affected by precipitation and seasonal influences. Results are listed in table 1.

<table>
<thead>
<tr>
<th>mode</th>
<th>dry</th>
<th>wet</th>
<th>Δ</th>
<th>summer</th>
<th>winter</th>
<th>Δ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Car driver</td>
<td>20.3</td>
<td>23.4</td>
<td>3.1</td>
<td>20.7</td>
<td>22</td>
<td>1.3</td>
</tr>
<tr>
<td>Car passenger</td>
<td>9.5</td>
<td>11</td>
<td>1.5</td>
<td>10</td>
<td>10.5</td>
<td>0.5</td>
</tr>
<tr>
<td>Bicycle</td>
<td>35.6</td>
<td>31.7</td>
<td>-3.9</td>
<td>36.9</td>
<td>31.8</td>
<td>-5.1</td>
</tr>
<tr>
<td>Public transport</td>
<td>2.8</td>
<td>2.8</td>
<td>0</td>
<td>2.6</td>
<td>3.1</td>
<td>0.5</td>
</tr>
<tr>
<td>other</td>
<td>31.9</td>
<td>31.2</td>
<td>-0.7</td>
<td>29.8</td>
<td>32.5</td>
<td>2.7</td>
</tr>
</tbody>
</table>

Table 1. Modal split (%) for short trips on dry versus wet days and summer versus winter

On wet days, a modal shift from bicycle to car (either driver or passenger) can be observed. Combined with a decrease in total trips on wet days, this results in an increase of car trips of 12% on wet days compared to dry days. The seasonal influence is less prominent though still present. As a result of an increase of modal share of car driver and a small decrease in the number of total trips, 3% more car trips in winter than in summer are observed. The increase of the modal share of the car can be explained by a decrease of the bicycle, that is most dramatic when winter and summer are compared. Finally it is remarkable that the share of public transport seems not to be affected by precipitation. Only in winter the modal share is 0.5% higher than in summer, which means about 20% more trips.

3.2. Impact of weather on urban traffic volume

As explained in section 2, matched pair analysis is used to examine the effect of rain on traffic volumes. Due to the limited space in this extended abstract we only discuss results. The most spectacular changes take place in night and evening periods. However, these results are of little value because traffic volumes are low in general as a result of rain for weekdays a decrease of volumes can be seen in the evening and at night, and an increase of volumes in the morning and afternoon. All weekdays show similar results, only at Wednesday-afternoon volumes decrease when it is raining. This can be explained by the fact that then schools are out. There were not enough
data to perform the analysis for all periods in the weekend. At Saturday an increase of volumes can be seen in the afternoon and a decrease at night. At Sunday the effects are different; when it is raining for traffic volumes a decrease in the afternoon and an increase at night can be seen. Traffic volumes appear to be higher in wet circumstances in morning periods as well as in afternoon periods compared to dry circumstances. The only exception is Wednesday afternoon, which is a free afternoon for kids that go to primary school. In weekends the opposite occurs, especially on Sunday afternoons a significant decrease of traffic flows is reported on wet days. The decrease in traffic flows on wet weekend days and Wednesday afternoon are probably due to a decrease in recreational/leisure trips. From Table 3 the 'wet'-increase in traffic volumes at the analyzed intersections in morning and afternoon periods on weekdays appears to amount approximately 4%. The effect is a bit stronger in the morning than it is in the afternoon. A possible explanation is the fact that mode choice is often made in the morning, and then the same mode must be used for the return trip, independent of the prevailing weather condition. The significance of the results is tested by executing a signed-ranks-test for the deviations of the pair-wise comparison of 'wet' and dry day parts. Negative deviations represent a lower traffic volume on wet day parts compared with dry day parts one (or two) weeks previous or later, whereas positive deviations show a greater traffic volume on wet day parts. As can easily be concluded from the low values of significance levels in Table 4 for the morning and the afternoon, the null hypothesis of equal traffic volumes on wet and dry days is rejected for these periods of the day. Thus, we conclude that a significant increase in traffic volumes is established by precipitation. The lack of a significant effect in evening and night periods is not surprising, because the lower traffic volumes during these periods of day lead to greater relative variations by occasional reasons.

3.3. Results from the survey

The survey results are not completely in line with the results from the OVG analysis and the comparison of traffic volumes in Almelo. Rain and other precipitation-related weather effects are the most important weather factors affecting mode choice as mentioned by the respondents. However, the impact of temperature and seasonal influences appears to be larger from the previous parts of this study. Apparently these factors have more impact on people’s total trip behavior: not only transport mode changes but also the trip’s purpose is affected. Outdoor recreation in summertime is a good example of this phenomenon.

Also the results of the questionnaires show that people change transport mode as a result of rainy weather. Weather is reported as the most important reason to change transport mode. The respondents state they substitute from bicycle into car and/or public transport in cases of rain. Apart from changing one’s transport mode, shortly postponing a short trip is also a possibility to avoid getting wet by rain when making bicycle or walk trips. Nearly half of the respondents (45.6%) sometimes postpones a bicycle trip because of rain at the scheduled departure time. On average the postponement is 15 minutes, depending on the necessity of the trip.
and the necessity of arriving on time. Postponing of a bicycle trip because of the weather occurs on average two times per month. Walking trips are also postponed sometimes, but not as often as bicycle trips. Postponing a trip can sometimes result in total omitting of the trip, especially for non-important trips like certain shopping or leisure trips. 22% of the respondents reports omitting bicycle trip in some cases due to the weather situation. Such a situation occurs on average once per month. Walking trips are more often omitted (on average twice per month), but only by fewer respondents (8.8%). 4% of the respondents omits car trips due to extreme bad weather circumstances (on average once per two months).

Finally, also weather forecasts appear to influence travel behavior, although the effect is smaller than the influence of the actual weather. 23% of the respondents report they change transport mode due to weather forecast whereas 55% changes transport mode when the weather actually changes.

4. Conclusion

For short trips (under 7.5 km) rain has a large effect on trip generation and modal split. Although total trip generation a substantial modal shift from bicycle to the car results in 12% more car trips on wet days.

Also seasonal differences are large, where bicycle use dramatically decreases in winter, and as a result car trips increase.

Apart from transport mode substitution or cancellation of a trip, rain causes postponing of trips as well. Approximately 7% of bicycle trips are postponed until an hour on rainy days. Change of destination also sometimes occurs, but this is only possible for a few trip purposes (e.g. shopping) and therefore is not a frequent effect.

References


FEASIBILITY OF CARSHARE IN MID-SIZE METROPOLITAN AREAS IN THE UNITED STATES

Ardeshir FAGHR1, Ph.D., and Adam CATHERINE

1. Project Overview

The majority of trips in metropolitan areas consist of single occupant vehicle trips, which is the most inefficient and expensive form of transportation for the user, and for the transportation system itself. Many urban areas offer mass transit systems as alternative commuter modes. However, mass transit is not always the most appropriate, or most convenient mode of transportation for every type of trip because most mass transit is scheduled-based, and that schedule is typically centered around peak travel times. Carsharing is a potential solution to filling in mass transit gaps.

Carsharing allows individuals in a metropolitan area to access a fleet of vehicles that are placed throughout the city. Instead of everyone owning their own vehicle, carsharing allows people who don’t need a car all the time to share with others in their area, which redistributes the expenses of owning a car over a larger group. In addition carsharing has shown to reduce the number of vehicles on the road, and reduce pollution.

The Wilmington, Delaware metropolitan area is no different than other larger metropolitan areas in the fact that it experiences a large amount of congestion due to commuters. In addition, it is experiencing a large amount of new residential development around the waterfront areas. If Wilmington is going to keep up with the growing development, and demand for a more

1 Department of Civil and Environmental Engineering, University of Delaware.
Newark, Delaware. USA 19716
efficient transportation system, new methods, programs, and policies must be established. Carsharing could be the City’s solution to augmenting its mass transit system, especially for the areas of new residential development in the city.

The purpose of this study is to examine the feasibility of carsharing within the City of Wilmington, Delaware. This study examines different carshare program development scenarios as well as offer implementation strategies and guidelines for those scenarios.

2. Background

Metropolitan areas are constantly trying to develop new ways to handle the ever increasing demand for single occupant vehicle trips. However; metropolitan areas have many assets that can be utilized to minimize the number of vehicle trips, especially those made by only one person.

Mass transit systems work in metropolitan areas because there are a large amount of people, trying to get in the same general direction at the same times everyday. Typically, the higher the population density in the metropolitan area, the higher the transit ridership. However; mass transit is not always the most appropriate, or most convenient, mode of transportation for every type of trip because most mass transit is scheduled-based, and that schedule is typically centered around peak travel times. For example, a trip to the grocery store, or moving from one location in the metropolitan area to the other, can not be accomplished by utilizing a city’s mass transit system. Sometimes people that rely on mass transit need another mode of transportation for those more difficult trips.

Carsharing offers a potential solution for those more difficult “every once in a while” trips. Carsharing allows individuals in a metropolitan area to access a fleet of vehicles that are placed throughout the city. Instead of everyone owning their own vehicle, carsharing allows people who don’t need a car all the time to share with others in their area, which redistributes the expenses of owning a car over a larger group. Cities enact carsharing programs in order to reduce congestion and pollution, as well encouraging mass transit use.

Carsharing offers the opportunity to reduce individual and social costs of owning a private vehicle. The majority of the cost of a car is due to paying off the loan used to purchase the vehicle, and paying for insurance. In addition there are other expenses related to maintenance, gas, and parking. Carsharing reduces theses costs because many people can use a single car, and so the cost is redistributed. In addition, carsharing does not offer the negative connotations that some forms of mass transit have because it does not limit the users’ independence.
There is substantial opposition to transportation policies that involve financial or lifestyle sacrifices, despite worsening traffic congestion. Carsharing offers some of the benefits of mass transit including:

- An average 44% reduction in vehicle miles traveled per person.
- Every carshare car on the road eliminates between 7 and 20 personal vehicles.
- Each member of a carshare program saves, on average, 100 gallons of gas per year.

In addition to the above benefits, carsharing forces a user to think of using mass transit first. Since carsharing is a pay-as-you-go program, it really forces the user to evaluate the need for paying for private transportation if public transportation is available for that trip. Carsharing is designed for short trips that occur only a few times a week, so it is necessary for the user to plan their trips better. There is a flat fee for using carshare programs in addition to a fee per mile. This forces the user to try to combine all necessary trips into one larger trip, so that multiple fees can be avoided.

3. Problem Statement

The Wilmington, Delaware metropolitan area is no different than other larger metropolitan areas in the fact that it experiences a large amount of congestion due to commuters. However; it is different from other cities, like Washington DC, Philadelphia, or San Francisco, because until recently, it has been a commuter city. This means that few people that work in Downtown Wilmington actually live in the city. However; with the recent development of new condo and apartment buildings, as well as the expected increase in further development, Wilmington is soon going to be a city with a large amount of residents who will be living and working in the city.

If Wilmington is going to keep up with the growing development, and demand for a more efficient transportation system, new methods, programs, and policies must be established. Carsharing could be the City’s solution to augmenting its mass transit system, especially for the areas of new residential development in the city.

This study examines the feasibility of carsharing within the City of Wilmington. It also examines different carshare program development scenarios as well as offer implementation strategies and guidelines for those scenarios.
4. Goals Accomplished

Three major goals for the carsharing program in Wilmington have been undertaken and objectives met:

Goal #1 - To determine the feasibility of a carshare program in the City of Wilmington.

Objective #1.1 – To provide the City of Wilmington with statistics of carshare implementation in other cities, as well as travel demand statistics of the Wilmington metropolitan area.

Objective #1.2 – To examine public opinion and depth of knowledge of carsharing.

Objective #1.3 – To determine whether the current mass transit system in the city is a viable transportation option.

Goal #2 – To examine different carshare program development scenarios.

Objective #2.1 - To determine areas of the city that would be best suited for a carsharing program based.

Objective #2.2 – To evaluate the possibility of implementing a city-wide program versus a program targeted on the new residential development.

Objective #2.3 – To examine typical trips made by residents in the new residential developments and determine if these trips would be appropriate for a carshare program.

Goal #3 – To develop an implementation recommendation and strategy.

Objective #3.1 – To highlight the scenarios that would have the highest probability of success.

Objective #3.2 – To develop a public education program on how to use the public carshare system effectively.

Objective #3.3 – To determine who will manage the carshare system, and the best way to manage the system as a whole.

Cliente
There are two major groups of people that this study has considered. The first, and most important, are the residents of Wilmington, who will actually use the system and determine its success. The second group is the City of Wilmington, the Delaware Department of Transportation, and Transportation Management Association of Delaware, who all will have a hand in the implementation and management of the carshare system. The study has made a major attempt to address all issues that may arise with all the different groups affected by the implementation of a carshare program.

Methods

This study has responded to a few major questions to determine the overall feasibility of a carshare program in Wilmington.

Question 1: Is the resident urban population large enough to support a carshare program?

Question 2: Is the average trip, from home to home, less than 15 miles?

Question 3: Is transit a viable option in the area?

Question 4: What is the public’s attitude and knowledge of carshare programs?

Question 5: Can the program be implemented city-wide?

Question 6: What are the issues related to implementation and management of a program?

Two primary methods for achieving the study goals and objectives, as well as answering the above questions include:

1. The analysis of other carshare programs in the US, which have yielded some information about necessary urban population size, as well as implementation and education strategies, both successful and unsuccessful.

2. Survey of residents in Wilmington. This has consisted of conducting surveys around study target areas, such as transit hubs, the new residential development, and the downtown office area. The surveys have helped to determine who would use the system, where it should be targeted, how the public should be educated.
DIFFERENCES AMONG ROUTE FLOW SOLUTIONS FOR THE USER-EQUILIBRIUM TRAFFIC ASSIGNMENT PROBLEM

Amos LUZON¹ and Hillel BAR-GERA²

Abstract. Many travel forecasting models incorporate the User-Equilibrium (UE) route choice principle. A known limitation of the UE principle is that although UE is sufficient to determine total link flows uniquely, it is not sufficient to determine route flows uniquely. By using two different UE network models as case studies, we seek to evaluate the potential disagreement between different route flow solutions. Our results show that in analyses that require route flows, it is important to make a proper choice of a specific route flow solution, such as the maximum entropy user equilibrium route flow solution.

1. Introduction

User-equilibrium (UE) traffic assignment is widely used in travel forecasting models to determine traffic patterns in urban and regional networks. Assignment results are analyzed in different ways to assist in various decisions during transportation planning processes. Some analyses need only the total flows on every link in the network, while other analyses require more detailed information about route flows. Known examples include the need to determine which portion of the flow on a given link arrives from a specific origin to a specific destination (also known as a select link analysis); deriving OD flows in a sub-region from a regional model [4]; choosing locations for facilities of a specific type (such as truck inspection stations, billboards, gasoline stations, automatic teller machines, etc.) in order to maximize the number of vehicles that will pass at least one facility along their route [3]; assessing the appropriate funding shares of different communities in a transportation improvement project [5]; and others.

An emerging application is the need to support license plates surveys. In these surveys observers in several locations throughout the network record vehicles’ license plates and

¹ Dept. of Industrial Engineering and Management, Ben-Gurion University, Beer-Sheva, Israel, amosl@bgu.ac.il.
² Dept. of Industrial Engineering and Management, Ben-Gurion University, Beer-Sheva, Israel, bargera@bgu.ac.il.
arrival times. Matching the license plates yield information related to route flows and travel times, thus providing an opportunity for validating model route flows. Estimated route flows from past models are also useful in the design of these surveys when choosing the locations for the observers.

A known limitation of the UE principle is that although UE is sufficient to determine total link flows uniquely, it is not sufficient to determine route flows uniquely and as a result there can be many route flow solutions. However, UE models are often used in applications like those mentioned above where the analyses do require route flows. In these cases, the choice of route flows solution is often made arbitrarily, without any specific behavioral assumption or mathematical condition to support the choice. The main goal of this research is to evaluate the potential disagreement between different route flow solutions.

2. Generating UE solutions

Our approach towards spanning the range of all UE route flow solutions was to rely on a sufficient number of route flow solutions. In that respect these solutions should be as different from each other as possible. Therefore, we aim to find "extreme" or "corner" alternative solutions. Using a simple procedure that is based on flow swaps between routes that have alternatives we generate a set of different ("extreme") route flow solutions, all maintaining the UE conditions with identical precision. In that respect, swaps are applied only if (1) they do not change total OD flows; (2) they do not change total link flow; (3) they consider only UE routes. The starting point for our swapping process is the Maximum Entropy User Equilibrium (MEUE) solution, which is considered to be the most likely route flow solution. The method we used to find the MEUE solution is described in [2]. Entropy maximization typically leads to solutions that are relatively in the "center", and in that respect it is a plausible reference point for other UE solutions. The set of alternative solutions is expected to span a major portion of the range of all UE route flow solutions, and thus allow a thorough investigation of the issue of non-uniqueness in UE route flows.

3. Results and analysis

In this work we use two UE models. The first model is for the Chicago region and consists of 387 zones, 933 nodes, 2950 links and 93,513 origin-destination (OD) pairs. (See [1] for the complete dataset.) The total number of used routes in the UE solution is 127,248. There are 74,114 OD pairs connected by one UE route only, meaning that these routes do not have alternatives. The remaining 53,134 UE routes do have alternatives, and their flows change among UE solutions.

The second model is a national model for Israel and consists of 240 zones, 4620 nodes, 11,234 links and 55,842 OD pairs. The total number of used routes in the UE solution is 1,319,446. In this model nearly all routes, 1,305,013, have alternatives, while only 14,433 routes do not have alternatives.
A major focus in our analyses is to evaluate the range of undetermined flow in each route. The flow on a specific route is completely undetermined if there is one UE solution in which the flow on the specific route is zero, while in another UE solution the same route carries the total OD flow. In order to evaluate the range of undetermined flow in each route we compute the minimum and the maximum route flow obtained in all different UE solutions that we generate. The results show that for most of the UE routes that have alternatives the maximum flow equals to the total OD flow and the minimum flow is zero. These dramatic results show that for nearly all UE routes that have alternatives the flow is completely undetermined by the UE conditions, and can be any arbitrary value between zero and the total OD flow.

Additional analyses include examining the ratio of the maximum UE flow on a specific route to its MEUE flow vs. the MEUE flow; and examining the minimum and the maximum flow proportion relative to the total OD flow vs. the MEUE flow. These analyses further emphasize the wide range of UE solutions.

Significant differences between UE solutions are found in more aggregate analyses as well, such as the distribution of users of a given link among OD pairs ("select link analysis"). In this analysis we wish to determine for a specific given link, for which the total UE flow is known, how its flow is distributed among all OD pairs. There are several possibilities for examining the effect of UE route flows on OD-based link flows. The one we find most useful focuses on the relative share of determined and undetermined OD-based link flow. We consider the ratio between the sum of minimum OD based link flows in a certain link to the total link flow as the relative share of link flow with determined OD. These ratios are shown on a map in Fig. 1.

FIG 1: Percentage of link flow for which the OD is the same in all of the randomly generated UE solutions for the Chicago network.

260
We find that in low congestion areas of the network the OD-based link flows can be uniquely determined. However, in congested areas such as within the City of Chicago and its CBD, it is not too rare to find a link on which the OD flow can not be uniquely determined for more than 50% of the flow. In many cases OD-based link flows are analyzed in the context of a very specific project, usually in relatively congested areas. If two analysts, both using state of the art UE solutions, find results that disagree in 50%, either by mistake or deliberately to support a specific agenda, that could result in a fairly controversial situation.

4. Conclusions

The preliminary results of this research demonstrate that in analyses that require route flows it is important to make a proper choice of a specific route flow solution, such as the maximum entropy user equilibrium route flow solution. Additional results for other UE models and additional analyses related to other route-based applications will likely enhance our understanding about the implications of arbitrary choices of route flows.

References


EVALUATION OF ON-STREET PARKING SCHEME USING VIRTUAL REALITY TRAFFIC EXPERIMENT SYSTEM

Shinji TANAKA¹, Masao KUWAHARA²

Abstract. A well-ordered on-street parking scheme in Japanese cities is highly desired to support urban social activities. This research studies the feasibility of making on-street parking space on a surface street in an aspect of efficiency and safety using the virtual reality traffic experiment system, which is an integrated system of driving simulator and traffic simulator. It can examine driving behavior around on-street parking as well as analyze traffic flow. For the safety evaluation, testee experiments by reproducing several patterns of parking spaces were conducted. The analyzed result of obtained driving behavior data will be utilized for the vehicle behavior model in the traffic simulation part.

1. Introduction

In Japan, parking has been regarded to be treated off street and basically prohibited to park on street because it may cause traffic congestions and sometimes traffic accidents in a city. However, since urban business and commercial activities depend on on-street parking greatly, this gap between institution and actual situation leads to a lot of problems like illegal parking and even disordered status of traffic. Therefore, to admit the role of on-street parking and to make parking space on street will be highly desired.

To allow on-street parking on arterial streets, it is quite important to ensure traffic capacity as well as safety between parked vehicles and passing vehicles. In this study, the authors focus on examining the efficiency and safety aspect of on-street parking space using virtual reality traffic experiment system, which is an integrated system of diving simulator and microscopic traffic simulator.

¹ Institute of Industrial Science, University of Tokyo,
² Center for Collaborative Research, University of Tokyo,
4-6-1 Komaba, Meguro-ku, Tokyo, 153-8505, Japan,
{stanaka, kuwahara}@iis.u-tokyo.ac.jp
2. Proposed On-street Parking Scheme

As for traffic congestion in urban area, the bottleneck is almost always at a signalized intersection, because limited amount of green time has to be allocated among different directions of traffic. Therefore, on-street parking around an intersection which reduces the road capacity significantly should be strictly controlled. In other words, there is a possibility to allow parking at a straight section between intersections. According to this idea, the authors propose to make on-street parking space between intersections as shown in Figure 1.

![Figure 1. Proposed on-street parking space](image-url)

This on-street parking scheme is designed inside road space, therefore before entering this section, the width for passing vehicles has to be narrowed, which may sometimes cause dangerous situations. However in the current situation in Japan, there are a lot of disordered illegal parking vehicles almost everywhere including near intersections, this concentrated parking could be a better solution. Anyway, to implement this kind of parking scheme, it is necessary to examine the safety of vehicles’ behavior around this parking section like merging, lane changing etc. And also, the effect of congestion alleviation by this scheme should be evaluated.

3. Virtual Reality Traffic Experiment System

The virtual reality traffic experiment system, which is an integrated system of a 6 axis motion driving simulator and a microscopic traffic simulator, has been developed in Sustainable ITS Project, University of Tokyo. Figure 2 shows the appearance of the current system and the details are explained in reference [1], [2]. The system can reproduce a realistic interaction between a driven vehicle by a testee and surrounding vehicles by microscopic traffic simulator and can observe his driving behavior in several scenarios under fully controlled environment.

For the safety evaluation of the proposed on-street parking scheme, a prototype model of surrounding vehicle behavior at the parking section is newly developed and introduced to this system.
4. Experiment

For the safety evaluation, we designed an experiment that has testees drive in a situation with on-street parking. 3 patterns of on-street parking spaces are prepared and compared as follows.

4.1. Current lane marking (pattern A)

Figure 3 shows a lane marking which does not intend that parking vehicles should be outside of through traffic lanes. This marking actually imitates the current situation where illegal parking occurs.
4.2. Parking lane with lane width reduction (pattern B)

Figure 4 shows a lane marking which creates parking lane by reducing each lane width for through traffic. The number of lanes for through traffic is maintained.

![Figure 4. Lane width reduction (pattern B)](image)

4.3. Parking lane with lane number reduction (pattern C)

Figure 5 shows a lane marking which creates parking lane by reducing number of lanes for through traffic. The lane width for through traffic is maintained or widened.

![Figure 5. Lane number reduction (pattern C)](image)

In the experiment, a testee is asked to drive a road stretch including these parking sections. Driving behavior data in every second are recorded and questionnaire survey is also done after driving. Here is an example of collected data.

**Driving behavior data:** vehicle’s location, speed, acceleration/deceleration rate, steering angle, driver’s eye point etc.

**Questionnaire survey:** subjective evaluation on safety, uneasiness etc.

5. Result

We have conducted the experiment inviting a few dozens of testees. Table 1 shows the number of testees of the experiment.
<table>
<thead>
<tr>
<th>Gender</th>
<th>Male: 32</th>
<th>Female: 14</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>20s: 30</td>
<td>30s: 12</td>
</tr>
<tr>
<td></td>
<td>40s: 2</td>
<td>50s: 2</td>
</tr>
<tr>
<td>Total</td>
<td>46</td>
<td></td>
</tr>
</tbody>
</table>

Table 1. Number of testees

Some of the results are shown in the following sections.

5.1. Driving trajectories

Figure 6 shows the driving trajectories by several testees in the pattern A. In this pattern, most of the drivers change their lane to the right completely although parked vehicles occupy less than half of the first lane. We can see the trajectories vary especially at the entrance point of the parking section.

5.2. Analysis on individual data

Figure 7 shows a sequence of several driving data, that is, steering wheel angle, vehicle speed, acceleration / deceleration ratio and brake pedal stroke associated with vehicle trajectory. From this figure, we can see at the entrance point, the driver takes several actions, such as steering and accelerating / braking, in order to get into the adjacent lane. If these values exceed a certain threshold, it may cause some dangerous situation.
5.3. Analysis on questionnaire

Subjective evaluation by testees is also done using questionnaire survey. Figure 8 shows whether testees felt dangerous or not on each of the patterns of lane marking. Compared with pattern B and C, testees feel more dangerous in the pattern A, a situation of illegal parking.

Figure 9 shows the result asking which was the most preferable pattern of marking and why the pattern was preferred. Drivers who preferred the pattern B raised no need of lane change, while drivers for pattern C find an advantage in wide lane configuration.
6. Conclusion

From the analysis based on the experiment so far, testees feel uneasiness or danger in the pattern A, which is similar to illegal parking situation. Therefore, it would be justified to make some types of parking lane. We are continuing to analyze the driving data in more precise way, for example, considering relative relationship with surrounding vehicles. It may reveal which geometric design of parking lane is desirable and which location of the section potentially become dangerous, and so on.

The result of the data analysis will be used to update the vehicle behavior model in the microscopic traffic simulator of the virtual reality traffic experiment system. Then, it will be used to evaluate proposed parking scheme including traffic flow efficiency analysis.

References


A COMPUTER VISION BASED SYSTEM FOR OFF-LINE QUALITY CONTROL OF A CHEMICALLY COATED COMPONENT OF A DIESEL FUEL INJECTION PUMP

Giuseppe MASTRONARDI¹, Massimiliano RONCONE², Vitoantonio BEVILACQUA³, Lucia CARIELLO⁴

Abstract. A case study for the automatic measurement of depth, quantity and geometry of the pores in the coating layer over a steel mechanical part is presented in this paper. At first the image is divided, by means of a statistics based algorithm, in three horizontal bands: the upper preparation material, the coating layer and underlying steel and coating layer and pore layer inside coating depth is evaluated. Then, by means of a segmentation algorithm and a morphological processing, single pores in form of lines are extracted from the background and their number, length and position are exactly calculated.

1. Introduction

Computer based image analysis has experienced and is experiencing, in the modern manufacturing industry, a very big spread. Possible applications are:

- object shape recognition;
- surface inspection for defects finding;
- materials microstructural characteristics evaluation.

Some of the above mentioned investigation fields are more suitable for in-line analysis methodologies, some other for off-line analysis methodologies. In most cases off-line and destructive strategies are linked together. In the case study presented in these pages, it has been developed an off-line methodology for the analysis of the fine structure of a diesel fuel injection pump chemically coated mechanical component. The images taken from the

¹ DEE Polytechnic of Bari, Via Orabona, 4, - 70125 Bari, Italy, mastrona@poliba.it.
² CVIT S.p.A. - BOSCH Italia, Via delle Ortensie- 19 – 70026 Modugno (Z.I.), Italy, massimiliano.roncone@it.bosh.com
³ DEE Polytechnic of Bari, bevilacqua@poliba.it.
⁴ DEE Polytechnic of Bari, cariello@deemail.poliba.it.
part under investigation metallographic samples are inputs for the developed analysis software (see Figure 1).

![Figure 1. Images taken from metallographic samples of two different parts (magnification 1000x).](image)

In Figure 2a, can be recognized three quite well delimited horizontal bands. The upper and lower bands are respectively the preparation material (bakelite) and the underlying steel base. The band in the middle is actually our concern: the chemically built up coating layer. Within the coating layer is possible to recognize a pore and a pore-free sub-layer. In the pore-free band, deep pores cross almost all the coating.

![Figure 2. (a) Horizontal subdivision of a metallographic sample. (b) Images of the characteristics calculated from sample image in (a); from these images it is possible to extract the necessary informations for the part quality evaluation.](image)

The characteristics to be extracted from presented images that will be used for the incoming quality control, are the followings:
- total thickness of coating;
- thickness of upper pore layer inside the coating;
- density of pores in the above mentioned sub-layer;
- maximal length of deep pores;
- number of deep pores whose length is bigger than a fixed value.

In Figure 2b two images are presented resulting from the statistical analysis (left image) and from skeleton process after segmentation (right image). From the image on the left is possible to calculate the total depth of coating layer and of pore sub-layer. It is possible to obtain the pore density too. The last evaluation is performed comparing the grey level of pore sub-layer with the one of pore-free sub-layer taken as reference. All other needed informations are extracted from the image on the right.

2. **Statistical analysis**

The basis for the following statistical analysis is the assumption that the image pixels, belonging to the same horizontal band, are occurrence of independent random variables
with same \textit{probability density function} PDF and therefore with same mean and standard deviation. The presented hypothesis relies on following statements:

- the whole sample, because of its small dimensions (100[\mu m] long), is equally illuminated;
- the physical aggregation and chemical composition of each band is homogeneous.

Taken as true the previous statements, it could be said that the number and the energy of photons emitted from the surface of each single element of the sample part is similar, in the case they belong to the same horizontal band. But the target of our analysis is just to find the band limits. In other words, the objective is to find a rule that defines to which band each row belongs to. The used criterion is based on the membership of the row mean to a given confidence interval for the mean of the rows belonging to the same horizontal band.

We assume that the pixels of a row are observed values of $g_n$ independent random variables ($n = 1, \ldots, N$ where $N$ is the number of image array columns), with same PDF, mean $\mu$ and variance $\sigma^2$. From the central limit theorem (§ 24.6 in [2]) follows that the random variable $g_m = (g_1 + \ldots + g_N) / N$ ($m = 1, \ldots, M$ where $M$ is the number of rows of image array) is asymptotically normal ($N \rightarrow \infty$) with mean $\mu$ and variance $\sigma^2/N$. It is then possible to apply to the random variable $g_m$ the same well known techniques to find a confidence interval for mean of random variables.

The following procedure is used (§ 24.6 in [2]):

1. choose a confidence interval $\gamma$ (e.g. 95%, 99%, or similar);
2. find out the solution of equation

$$\frac{1}{2}(1+\gamma)$$

from the coded table of $t$ \textit{Student} distribution with $n-1$ degrees of freedom ($n$ is the number of known samples);
3. calculate the mean $\bar{x}$ and variance $s^2$ of samples;
4. calculate $k = sc / \sqrt{n}$ and than the confidence interval for the mean $\mu$

$$\bar{x} - k \leq \mu \leq \bar{x} + k.$$ 

In our case study $n$ is a multiple of $N$. As the number of rows belonging to the band increases, i.e. as the number of samples for the evaluation of mean and variance increases, it reduces the confidence interval amplitude of a factor $\sqrt{N}$ (where $l$ is the number of rows already belonging to examined pump). For the images of Figure 1 has been chosen: $\gamma = 99.8\%$, and therefore, because $n$ is big already at first calculation ($n \approx 900$), then $c = 3.09$. If $b$ is the band the $m$-row could belong to, the interval is given by $\bar{x} - k \leq \bar{x}_m \leq \bar{x} + k$, where $\bar{x}$ and $k$ are referred to the pixels to be assigned to $b$ and $\bar{x}_m$ (calculated as mean of the pixel already assigned to $b$ and of pixel of a $m$-row).

3. \textbf{Segmentation and binary morphology}

The step following the band separation of image is the binary segmentation. The role of the object and background is played respectively by the deep pores and by the white coating. The pores will be rendered as black pixels, the background will be white. Segmentation is a required step before the skeleton process and digital morphology processing.
3.1. **Hyperbolic filter**

Among different segmentation algorithms, the hyperbolic filter \[1\] has shown to be the better for our images. Let consider the pixel \(p_0\) of a gray level image and the pixel \(p'_0\) as the corresponding pixel of resulting binaryzed image. Let \(\overline{p}_0\) be the pixel resulted from convolution between original image and a given convolution array. \(p'_0\) is defined as:

\[
p'_0 = \begin{cases} 
0 & \text{if } p_0 \leq 2\mu - \frac{N}{\overline{p}_0 + 1} \\
1 & \text{if } p_0 > 2\mu - \frac{N}{\overline{p}_0 + 1}
\end{cases}
\]

where

- \(\mu\) is the mean of grey levels of complete image;
- \(N\) is the number of possible grey levels \(8\) [bit/pixel].

The Roberts’ array is the one who gives better results. For our purposes has been used a 5th order Roberts’ convolution array.

3.2. **Zhang-Suen skeleton algorithm**

The step after band recognition and binaryzation of coating band is the one for the skeleton (or essential line) production. By means of the skeleton is possible to extract the number of deep pores and their dimensions. A widely used algorithm is the one prepared by Zhang-Suen \[5\]. It has been used as metering algorithm to judge the quality of other algorithms. It is a parallel method, that is, the new value of a pixel can be calculated just starting from the values coming from the previous iteration.

The algorithm is divided in two steps. During the first one, a pixel \(I(i, j)\) is deleted (or marked to be deleted) if the following conditions have been proven:

1. the connectivity number value is OK;
2. has at least 2 near pixels that belong to background and object;
3. at least one of \(I(i-1, j), I(i-1, j+1)\) and \(I(i, j-1)\) is a background pixel (white);
4. at least one between \(I(i-1, j), I(i+1, j)\) and \(I(i, j-1)\) is a pixel of background.

At the end of this first analysis, the marked pixels are deleted. The following step, checks whether:

1. at least one between \(I(i-1, j), I(i+1, j)\) and \(I(i, j+1)\) is a background pixel;
2. at least one between \(I(i, j+1), I(i+1, j)\) and \(I(i, j-1)\) is background pixel.

The marked pixels are deleted. If at the end of both iterations there are no more pixels marked to be deleted, the program stops; the skeleton has been completed.

In order to improve the quality of the skeleton, the chain Stentiford \[4\] preprocessing→Zhang-Suen algorithm→Holt \[4\] postprocessing has been proven to be very effective.

![Figure 3.](image-url)

*Figure 3. In (a) is presented the central band of the sample image in Figure 1 after binaryzation process. In (b) the result of skeleton process.*
After skeleton process further improvements of the image under analysis have been realized by means of simple algorithms of digital morphology. All horizontal lines have been deleted, the pore branches have been separated, etc. The results related to the image in Figure 3 are presented in Figure 4.

![Figure 4](image)

Figure 4. In (a) the horizontal parts of the skeleton have been suppressed, in (b) the branches of the skeleton have been separated. At last in (c) the small circles like the one pointed out, in (b), from the arrow have been removed.

The fundamental algorithms that have been used in Figure 4 are listed below. The first one, that has been used to delete horizontal lines, outputs a boolean TRUE in the case the examined pixel is the central pixel of one of the masks presented in Figure 5, otherwise FALSE. In case of TRUE the examined pixel is marked for deletion.

![Figure 5](image)

Figure 5: Templates to identify horizontal lines.

![Figure 6](image)

Figure 6. In (a) the mask used for binary erosion. The pixel marked with X is the origin pixel. In (b) the skeleton element to be deleted.

![Figure 7](image)

Figure 7. All steps described in the Figures 3 and 4 applied to the second sample image of Figure 1. In this case, the skeleton of pores shorter than 15 pixels has been deleted.

The second one is a recursive algorithm that separates the pore branches. It is needed to count the single pores. The policy of branch separation is to separate the shorter branch from the longer one. The third algorithm is the one that eliminates the short pores. The
pores shorter than a user fixed length are deleted. It is recursive algorithm that calculates the length of pores and deletes the shorter ones. Last processing is that one used in Figure 4b in order to delete the circular artefacts like the one pointed out by the arrow. A binary erosion algorithm is used. The structuring element is presented in Figure 6a, it has actually the shape of the artefacts to be deleted. The eroded image is combined with not eroded image in order to eliminate the artefacts.

The Figure 4c is finally ready to correctly identify the position, the length and the number of deep pores inside the coating layer of the part under analysis.

The results obtained from the application of all steps described in the Figures 3 and 4 to the second sample image of Figure 1, are represented in Figure 7.

4. Conclusions

The developed procedure has been tested on many (> 10) images of different metallographic samples. With same camera, microscope and frame grabber for the image acquisition, the software was able to perform a correct analysis (recognise the horizontal bands and find out pores) in the 99% of the examined cases. This software based analysis becomes therefore a valid objective methodology for the evaluation in quality classes on incoming parts. The main advantage is the elimination of the operator subjectivity.

A combination of 1st order statistical analysis, segmentation techniques and basic digital morphology algorithms has been effectively used to face the problem. The most important step in the development of this software has been the coating layer recognition. Actually the statistical approach as first analysis step could be used thanks to the horizontal image distribution. In order to let the procedure be more robust a step “0” can be implemented. The role of the step “0” should be to turn the image to the horizontal position.

References

AN ANT COLONY OPTIMIZATION APPROACH FOR THE CAPACITATED VEHICLE ROUTING PROBLEM WITH SIMULTANEOUS DELIVERY AND PICK-UP

Bülent Çatay

Abstract. We propose an Ant Colony Optimization (ACO) algorithm to the NP-hard Vehicle Routing Problem with Simultaneous Delivery and Pick-up (VRPSDP). In VRPSDP, commodities are delivered to customers from a single depot utilizing a fleet of identical vehicles and empty packages are collected from the customers and transported back to the depot. The objective is to minimize the total distance traveled. The algorithm is tested with the well-known benchmark problems from the literature. The experimental study indicates that our approach produces comparable results to those of the benchmark problems in the literature.

1. Introduction

The Vehicle Routing Problem with Simultaneous Delivery and Pick-up (VRPSDP) is a variant of the VRP where the vehicles are not only required to deliver goods but also to pick up some goods from the customers. Customers receiving goods are called linehauls while customers sending goods are called backhauls. The objective is to minimize the total distance traveled by the vehicles and/or the number of vehicles used subject to maximum distance and maximum capacity constraints on the vehicles. VRPSDP may be classified into three categories: (i) Delivery First, Pick-up Second: the vehicles pick up goods after they have delivered their goods; (ii) Mixed Delivery and Pick-up: linehauls and backhauls can occur in any sequence on a vehicle route; and (iii) Simultaneous Delivery and Pick-up: the vehicles simultaneously deliver and pick-up goods [11]. Delivery First Pick-up Second and Mixed VRPDP problems are jointly referred to as the VRP with backhauling (VRPB). Delivery First Pick-up Second and Mixed VRPDP problems are jointly referred to as the VRP with backhauling (VRPB). Each customer has either a delivery demand or a pick-up to be satisfied. Products to be delivered are loaded at the depot while picked up products are transported back to the depot. A set of vehicle routes has to be designed so that all

1 Sabanci University, Faculty of Engineering and Natural Sciences, Tuzla, Istanbul, 34956, Turkey, catay@sabanciuniv.edu
customers are serviced exactly once and no “pick-up customer” is visited before any other “delivery customer” on the same route.

In this paper, we address VRPSDP with a fleet of uniform vehicles each having the same capacity. The delivery and pick-up items are identical in the sense that each unit consumes the same amount of vehicle capacity. Delivery and pick-up locations are unique and the delivery to a customer is only allowed from the depot. The objective is to serve all customers with the minimum total distance.

Although VRP has been intensively studied in the literature research on VRPSDP is scant. The problem is first introduced by Min [10], where book distribution and recollection activity between a central library and 22 remote libraries at a county in Ohio, with limited number of trucks and available capacity, is discussed. Min utilizes a cluster-first route-second approach and solves Traveling Salesman Problems (TSP) to optimality as sub-problems. Halse [9] studies VRPSDP as well as many other special case problems in the VRP literature. Cases with a single depot and multiple vehicles and number of nodes varying between 22 and 150 are addressed. Lagrangean relaxation and column generation approach is utilized. A cluster-first route-second type heuristic is developed in which nodes are first distributed to vehicles and then the problem is solved using 3-opt algorithm.

Angelelli and Mansini [1] study the VRPSDP with time windows constraints. They implement a branch-and-price approach based on a set covering formulation. Gendreau et al. [8] study the VRPSDP for a single vehicle case and develop two heuristics. Casco et al. [3] introduce a load-based insertion procedure for VRPB where the insertion cost for backhaul customers is based on the load still to be delivered. Salhi and Nagy [12] modify this method by allowing backhauls to be inserted in clusters rather than one by one. This procedure is also capable of solving simultaneous problems.

Dethloff [6] presents insertion-based heuristics using four different criteria. He develops 40 instances to test his algorithm. He also compares his results with those of [12] and reports an improvement on Min’s [10] problem. In the problem structure in [12], nodes are represented as disjoint delivery or pick-up nodes so repetitive servicing is allowed. Besides, the problem puts a limit on the maximum route length and introduces multiple depots rather than a single depot case.

In this paper, we propose Ant Colony Optimization (ACO) algorithm for the VRPSDP introducing a new visibility function. To our knowledge, no ACO based approach has been previously proposed for this problem. Only Wade and Salhi [25] address the Mixed VRPB and Reimann et al. [20] address the VRPB with time windows using an ACO based approach. The remainder of the paper is organized as follows: In Section 2 description of the problem is provided. Section 3 is devoted to the discussion of the proposed ACO approach. Section 4 discusses computational experiments and numerical results. Finally, concluding remarks and future research issues are presented in Section 5.

2. Problem Description

The problem deals with a single depot distribution/collection system servicing a set of customers using a homogeneous fleet of vehicles. Each customer requires two types of service: a delivery and a pick-up. The critical feature of the problem is that both activities have to be carried out simultaneously by the same vehicle. Products to be delivered are
loaded at the depot and products picked up are transported back to the depot. The objective is to find the set of routes servicing all the customers with the minimum total distance [1].

From a practical point of view, VRPSDP models situations such as the distribution of bottled soft drinks, LPG tanks, laundry service of hotels where the customers are typically visited only once but for a double service and grocery stores where reusable specialized pallets/containers are used for the transportation of merchandise. Also, regulations may force companies to take responsibility for their products throughout their lifetime.

Mathematically, VRPSDP is described by a set of homogenous vehicles \( V \), a set of customers \( C \), and a directed graph \( G(N, A) \). \( N = \{0, \ldots, n+1\} \) denotes the set of vertices. Each vehicle has capacity \( Q \) and each customer \( i \) has delivery and pick-up requests \( d_i \) and \( p_i \) respectively. The graph consists of \(|C|+2 \) vertices where the customers are denoted by \( 1, 2, \ldots, n \) and the depot is represented by the vertices 0 and \( n+1 \). \( A = \{(i, j): i \neq j\} \) denotes the set of arcs that represents connections between the depot and the customers and among the customers. No arc terminates at vertex 0 and no arc originates from vertex \( n+1 \). A cost/distance \( c_{ij} \) is associated with each arc \((i, j)\). Finally, \( Q, d_i, p_i, c_{ij} \) are assumed to be non-negative integers.

If \( P \) is assumed as an elementary path in \( G \), \( P = \{0 = i_0, i_1, \ldots, i_p, i_{p+1} = n + 1\} \), a feasible solution for our problem can be represented by a set of disjoint elementary paths originating from 0 and ending at \( n+1 \). These paths visit every customer exactly once while satisfying the capacity constraints. Thus, the pick-up demands already collected plus the quantities to be delivered must not exceed the vehicle capacity. The objective is to minimize the total distance traveled by all the vehicles. (Refer to Dethloff [6] for the mathematical model)

Anily [2] proves that the VRPB is \( NP \)-hard in the following way: If \( P_j = 0 \) \((j \notin J)\) or even \( P_j = D \) \((j \notin J)\) then the problem reduces to the VRP which is known to be \( NP \)-hard. Thus, VRPB is also \( NP \)-hard. VRPB may be considered as a special case of the VRPSDP where either the delivery demand \( D_j \) or the pick-up demand \( P_j \) of each customer equals zero [12]. Hence, VRPSDP is also \( NP \)-hard.

3. Description of the ACO based approach

ACO is based on the way real ant colonies behave to find the shortest path between their nest and food sources. While walking ants leave an aromatic essence, called pheromone, on the path they walk. Other ants sense the existence of pheromone and choose their way according to the level of pheromone. Greater level of pheromone on a path will increase the probability of ants following that path. The level of pheromone laid is based on the path length and the quality of the food source. It will increase when the number of ants following that path increases. In time all ants are expected to follow the shortest path.

ACO simulates the described behavior of real ants to solve combinatorial optimization problems with artificial ants. Artificial ants find solutions in parallel processes using a constructive mechanism guided by artificial pheromone and a greedy heuristic known as visibility. The amount of pheromone deposited on arcs is proportional to the quality of the solution generated and increases at run-time during the computation. ACO was first introduced for solving the TSP [5]. Since then many implementations of ACO have been proposed for a variety of combinatorial optimization problems such as quadratic assignment problem, scheduling problem, sequential ordering problem, and vehicle routing
problems. The interested reader is referred to [7] for details of ACO metaheuristic. In what follows is the description of the algorithm.

3.1. Initialization

An initial amount of pheromone \( t_0 \) is deposited on each arc. In the literature, it has been observed that \( t_0 = \frac{1}{n L_{\text{init}}} \), where \( L_{\text{init}} \) is the length of an initial feasible route and \( n \) is the number of customers, can generate the good routes. The initial route is constructed by starting at the depot and then selecting the not yet visited closest feasible customer as the next customer to be visited. A customer is infeasible if it violates the capacity. If no feasible customer is available then the route is terminated at the depot and a new route is initiated.

Since the distance of customers to the depot is an essential characteristic of the tour length we incorporate this in calculating the visibility considering the savings of serving customer \( i \) and \( j \) on the same route instead of serving them on different tours. The distance between two customers is introduced into the visibility function through the use of Clarke and Wright savings measure [4]:

\[
s_{ij} = d_{ij} + d_{0i} - d_{0j}
\]

\( d_0 \) (\( d_0 \)) denotes the distance between customers \( i \) and \( j \) (the depot). The higher savings value favors visiting customer \( j \) after customer \( i \) while the longer distance value prevents it. Thus, the savings per unit distance traveled between customers measures the attractiveness of visiting customer \( j \) after customer \( i \). The visibility of selecting customer \( j \) after customer \( i \) is then computed as follows:

\[
\eta_{ij} = \begin{cases} 
\frac{s_{ij}}{d_{ij}}, & \text{if } s_{ij} \geq 1 \\
1/d_{ij}, & \text{otherwise}
\end{cases}
\]

Since a high value of \( \eta_{ij} \) indicates that visiting customer \( j \) after customer \( i \) is a desired choice the tour length is expected to be shorter if the probability of moving from customer \( i \) to customer \( j \) increases with \( \eta_{ij} \). Furthermore, a candidate list is used to reduce the computation time. The candidate list of each customer is formed as follows: in ACO, visiting customer \( j \) after the current customer \( i \) is based on the amount of both the pheromone trails \( t_{ij} \) and the visibility \( \eta_{ij} \) on arc \((i,j)\). Therefore, at each customer \( i \) candidate set \( O(i) \) is formed by taking \( k \) feasible customers with the largest attractiveness \( \Phi_{ij} \).

3.2. Route construction process

An ant is positioned at each customer and each ant constructs its own tour by successively selecting a not yet visited feasible customer. The choice of the next customer is based on its attractiveness values, which is a function of the pheromone trails and the visibility:

\[
\Phi_{ij} = \left[ t_{ij} \right]^a \left[ \eta_{ij} \right]^b
\]

\( a \) and \( b \) are parameters to control the relative weight of trail intensity \( t_{ij} \) and visibility \( \eta_{ij} \).

Using the following equations (4) and (5) each ant may either visit the most favorable customer or randomly select a customer to visit based on a probability distribution \( p(i,j) \).

\[
p(i,j) = \begin{cases} 
\arg \max_{\min \in \mathbb{P}(i,j)} \Phi_{ij}, & \text{if } q \leq q_0 \\
\frac{p(i,j)}{P(i,j)}, & \text{otherwise}
\end{cases}
\]
\[
P(i, j) = \begin{cases} 
\frac{\varphi_j}{\sum_{\alpha \in \Omega(i)}} & \text{if } j \in \Omega(i) \\
0 & \text{otherwise}
\end{cases}
\]

where \( q_0 \) (\( 0 = q_0 = 1 \)) is a parameter to control exploitation versus exploration.

### 3.3. Local and global updates

In order to reduce the probability of repeatedly selecting a customer, the amount of the pheromone on the arc is reduced through evaporation. The reduction is made by applying a local updating rule given in equation (6):

\[
\tau_{ij} = (1 - p) \tau_{ij} + \rho \tau_{ij} \tau
\]

where \( \tau \) (\( 0 = \tau = 1 \)) is the trail persistence parameter. If no feasible customer is available due to the vehicle capacity constraint then the depot is chosen and a new route is started. This process is executed until all customers have been visited.

Once all ants construct their tours, the best \( \tau \) tours are chosen and the global updating rule is applied as follows:

\[
\tau_{ij} = (1 - p) \tau_{ij} + \sum_{r=1}^{\tau} \Delta \tau_{ij} + \lambda \Delta \tau_{ij}^*
\]

If an arc is used by the \( r \)-th best ant, the pheromone value on arc \((i, j)\) will be increased by \( \Delta \tau_{ij}^* = (\lambda - r)/L^* \), where \( L^* \) is the tour length of the \( r \)-th best ant. The pheromone level on the arcs of the best solution is also increased by \( \Delta \tau_{ij}^* = 1/L^* \), where \( L^* \) is length of the best tour. The steps of the algorithm are summarized in Figure 1.

```plaintext
compute visibility
initialize pheromone
while (max number of iterations is not reached)
    for each ant i
        select the next customer to visit from candidate list O(i)
        update vehicle capacity and candidate list
        perform local pheromone trail update
    end for
    perform global pheromone trail update
    save the best route
end while
```

Figure 1. Pseudo-code of the algorithm.

### 4. Experimental study

In this section, the proposed ant colony system based approach for VRPSDP is tested on the benchmark problem instance(s) of Min [10] and Dethloff [6]. The algorithm is coded using C++. The parameters were set according to initial experimental runs to different problems: \( k = \text{(No of Customers)} / 2 \), \( q_0 = 0.80 \), \( a = 0.75 \), \( \beta = 4 \), and \( \tau = 15 \). The number of iterations is set to 5000 and \( \rho = 0.15 \).
<table>
<thead>
<tr>
<th>Problem</th>
<th>Dethloff's Best</th>
<th>Avg Soln</th>
<th>Best Soln</th>
<th>% Imp</th>
</tr>
</thead>
<tbody>
<tr>
<td>Min</td>
<td>89</td>
<td>89,8</td>
<td>88</td>
<td>1.14%</td>
</tr>
<tr>
<td>SCA3-0</td>
<td>689</td>
<td>705,9</td>
<td>697,0</td>
<td>-1.15%</td>
</tr>
<tr>
<td>SCA3-1</td>
<td>765,6</td>
<td>781,5</td>
<td>772,4</td>
<td>-0.88%</td>
</tr>
<tr>
<td>SCA3-2</td>
<td>742,8</td>
<td>744,0</td>
<td>742,6</td>
<td>0.02%</td>
</tr>
<tr>
<td>SCA3-3</td>
<td>737,2</td>
<td>731,0</td>
<td>727,9</td>
<td>1.27%</td>
</tr>
<tr>
<td>SCA3-4</td>
<td>747,1</td>
<td>766,6</td>
<td>758,9</td>
<td>-1.55%</td>
</tr>
<tr>
<td>SCA3-5</td>
<td>784,4</td>
<td>803,8</td>
<td>799,8</td>
<td>-1.92%</td>
</tr>
<tr>
<td>SCA3-6</td>
<td>720,4</td>
<td>725,5</td>
<td>719,2</td>
<td>0.17%</td>
</tr>
<tr>
<td>SCA3-7</td>
<td>707,9</td>
<td>742,3</td>
<td>731,5</td>
<td>-3.23%</td>
</tr>
<tr>
<td>SCA3-8</td>
<td>807,2</td>
<td>820,9</td>
<td>799,7</td>
<td>0.93%</td>
</tr>
<tr>
<td>SCA3-9</td>
<td>764,1</td>
<td>763,4</td>
<td>746,2</td>
<td>2.40%</td>
</tr>
<tr>
<td>SCA8-0</td>
<td>1132,9</td>
<td>1102,8</td>
<td>1087,5</td>
<td>4.17%</td>
</tr>
<tr>
<td>SCA8-1</td>
<td>1150,9</td>
<td>1232,2</td>
<td>1230,0</td>
<td>-6.43%</td>
</tr>
<tr>
<td>SCA8-2</td>
<td>1100,8</td>
<td>1164,7</td>
<td>1143,1</td>
<td>-3.70%</td>
</tr>
<tr>
<td>SCA8-3</td>
<td>1115,6</td>
<td>1264,9</td>
<td>1243,6</td>
<td>-10.29%</td>
</tr>
<tr>
<td>SCA8-4</td>
<td>1235,4</td>
<td>1331,6</td>
<td>1286,9</td>
<td>-4.00%</td>
</tr>
<tr>
<td>SCA8-5</td>
<td>1231,6</td>
<td>1272,4</td>
<td>1266,2</td>
<td>-2.73%</td>
</tr>
<tr>
<td>SCA8-6</td>
<td>1062,5</td>
<td>1181,9</td>
<td>1177,5</td>
<td>-9.77%</td>
</tr>
<tr>
<td>SCA8-7</td>
<td>1217,4</td>
<td>1304,2</td>
<td>1293,0</td>
<td>-5.85%</td>
</tr>
<tr>
<td>SCA8-8</td>
<td>1231,6</td>
<td>1350,7</td>
<td>1326,6</td>
<td>-7.16%</td>
</tr>
<tr>
<td>SCA8-9</td>
<td>1185,6</td>
<td>1230,7</td>
<td>1189,4</td>
<td>-0.32%</td>
</tr>
<tr>
<td>CON3-0</td>
<td>672,4</td>
<td>668,6</td>
<td>662,6</td>
<td>1.48%</td>
</tr>
<tr>
<td>CON3-1</td>
<td>570,6</td>
<td>606,4</td>
<td>600,5</td>
<td>-4.97%</td>
</tr>
<tr>
<td>CON3-2</td>
<td>534,8</td>
<td>549,5</td>
<td>538,3</td>
<td>-0.66%</td>
</tr>
<tr>
<td>CON3-3</td>
<td>656,9</td>
<td>644,4</td>
<td>636,5</td>
<td>3.20%</td>
</tr>
<tr>
<td>CON3-4</td>
<td>640,2</td>
<td>630,9</td>
<td>628,7</td>
<td>1.84%</td>
</tr>
<tr>
<td>CON3-5</td>
<td>604,7</td>
<td>618,9</td>
<td>610,7</td>
<td>-0.99%</td>
</tr>
<tr>
<td>CON3-6</td>
<td>521,3</td>
<td>549,3</td>
<td>546,3</td>
<td>-4.58%</td>
</tr>
<tr>
<td>CON3-7</td>
<td>602,8</td>
<td>634,8</td>
<td>625,1</td>
<td>-3.57%</td>
</tr>
<tr>
<td>CON3-8</td>
<td>556,2</td>
<td>581,6</td>
<td>578,0</td>
<td>-3.77%</td>
</tr>
<tr>
<td>CON3-9</td>
<td>612,8</td>
<td>615,4</td>
<td>615,3</td>
<td>-0.40%</td>
</tr>
<tr>
<td>CON8-0</td>
<td>967,3</td>
<td>1011,1</td>
<td>1000,0</td>
<td>-3.27%</td>
</tr>
<tr>
<td>CON8-1</td>
<td>828,7</td>
<td>845,7</td>
<td>840,5</td>
<td>-1.40%</td>
</tr>
<tr>
<td>CON8-2</td>
<td>770,2</td>
<td>837,4</td>
<td>820,0</td>
<td>-6.07%</td>
</tr>
<tr>
<td>CON8-3</td>
<td>906,7</td>
<td>970,6</td>
<td>960,0</td>
<td>-5.56%</td>
</tr>
<tr>
<td>CON8-4</td>
<td>876,8</td>
<td>909,2</td>
<td>896,3</td>
<td>-2.17%</td>
</tr>
<tr>
<td>CON8-5</td>
<td>866,9</td>
<td>946,6</td>
<td>923,8</td>
<td>-6.16%</td>
</tr>
<tr>
<td>CON8-6</td>
<td>749,1</td>
<td>827,3</td>
<td>799,2</td>
<td>-6.27%</td>
</tr>
<tr>
<td>CON8-7</td>
<td>929,8</td>
<td>959,1</td>
<td>934,2</td>
<td>-0.47%</td>
</tr>
<tr>
<td>CON8-8</td>
<td>833,1</td>
<td>887,7</td>
<td>876,5</td>
<td>-4.96%</td>
</tr>
<tr>
<td>CON8-9</td>
<td>877,3</td>
<td>914,8</td>
<td>898,2</td>
<td>-2.33%</td>
</tr>
</tbody>
</table>

Table 1. Average improvements with ACO compared to Dethloff’s results.
Min [10] reported the objective function value for his problem as 94. Dethloff [6] reported his best solution for Min’s problem as 91 with a computation time of 0.27 seconds. Dethloff also reported the best known solution as 89 after 100 hours of computing time on a Pentium III 500 Mhz processor using XPRESS-MP. Our proposed algorithm obtained has been able to improve the solution to 88 with a computation time of 50 seconds on a Pentium IV 2.6 Gzhz processor.

Dethloff performed 10 experiments for each data and published the average travel distances. We performed 5 runs for each instance, compare the averages to those of Dethloff’s, and compute the gap as \((\text{Dethloff Avg/ACO Avg}) - 1\). We observe that ACO algorithm outperforms in 9 out of 40 problem instances. The average computation time for all problem instances is about 15 minutes. Table 1 reports Dethloff’s best solution in comparison with our average and best solutions as well as the gap for all problem instances. We observe that our algorithm performs better for the SCA instances where the customers are scattered. It is also worth noting here that these results are directly obtained from the ACO algorithm and they may be further improved by using a local search heuristic.

5. Conclusion and future research directions

We address the VRPSDP which has a growing practical relevance in the reverse logistics literature. The computational complexity of the problem necessitates good heuristic solution procedures. We tackle the problem using an ACO based solution approach equipped with a new visibility function. The experimental analysis reveals some improvements to the previously published results in the literature. In addition, our algorithm has provided the best solution in a well-known problem instance. On the other hand, the computation times are larger compared to the other heuristics whose progresses have been declared to be in seconds.

Future work in this area may be dedicated to apply a local search heuristic to further improve the solutions obtained through ACO algorithm. The author is currently investigating 2-opt and 3-opt algorithms. Another future direction may involve attempting to further reduce the number of user controlled parameters and to improve the speed of the procedure. Parallel computing techniques may be utilized to reduce the computational efforts. The approach may be extended to apply to other types of VRP as well.

References


IMPLEMENTATION OF AN ASYMMETRIC NETWORK EQUILIBRIUM PROBLEM WITH DETAILED REPRESENTATION OF UNSIGNALIZED AND SIGNALIZED URBAN INTERSECTIONS.

Marino LUPI¹, Federico RUP1¹, Guido ROSSI¹

Abstract. This paper discusses the implementation of an asymmetric network equilibrium model with detailed representation of unsignalized and signalized urban intersections. A software has been developed to solve the deterministic user equilibrium (DUE) problem which takes into account real urban intersections in their detailed configurations. During the first phase this software was tested on a “toy” network and then on the real network of Villafranca (a town near Verona Italy). The comparison between the equilibrium flow patterns resulting from the model and some traffic counts on the Villafranca network confirms that the model is good.

1. Introduction

This paper describes the implementation of an asymmetric user equilibrium route choice model with non-separable cost functions. The study is focused on a detailed representation of unsignalized and signalized intersections in the network. This kind of representation can highlight in the best way the complex interactions between different traffic streams competing for the use of limited road capacities and allows the best possible determination of the link average delays of the flows approaching the intersection; the analysis procedures have followed the gap acceptance theory. The interactions among different traffic streams sharing some lanes on the same approach, and, above all, among conflicting traffic streams using different approaches, require the use of non-separable cost functions. One of the distinctive features of this paper is related to the computation of average delay for unsignalized intersections and, particularly, for modern roundabouts. Such computation is carried out in a detailed way which is not usual in ordinary assignment software. The equilibrium assignment problem needs to be formulated and solved as an asymmetric equilibrium problem. Our first target was the development of an asymmetric equilibrium assignment software for studying real urban intersections, both signalized and unsignalized, and particularly roundabouts in their detailed configurations. During the first phase of our

¹ DISTART TRASPORTI, University of Bologna, viale Risorgimento 2, 40136 Bologna, marino.lupi@unibo.it
work the software was tested on a “toy” network and subsequently on the real network of Villafranca near Verona.

2. Detailed cost functions for unsignalized intersections in assignment problems

During this study particular attention was devoted to modelling the delay of unsignalized intersections in urban network assignment problem. An example of unsignalized T-intersection is represented in Fig. 1.

![Figure 1. Right-turn manoeuvre from a minor street in a shared lane.](image)

The average control delay $d$ (sec/veh) for the right-turn stream from a minor street can be evaluated by means of the following expression:

$$d = d_q + d_s + 5$$  \hspace{1cm} (1)

where: $d_q$ (sec/veh) is the average queue waiting time (the time elapsed since joining the queue to the time when the vehicle is at the stop line ready to carry out its manoeuvre); $d_s$ (sec/veh) is the average service time (in this case it represents the average time necessary to carry out its manoeuvre as the vehicle is at the stop line). The constant value of 5 sec/veh is included in equation (1) to account for the deceleration of vehicles from free-flow speed to the speed of vehicles in a queue and the acceleration of vehicles from the stop line to free-flow speed (HCM, 2000).

To give an example of the method used, the determination of the average delay for a minor street right-turn stream, in a shared lane first and then in an exclusive lane are reported.

In the case of the shared lane, $d_q$ could be determined by means of the equation (HCM, 2000):
where: $v_{RT}$ is the minor street right-turn stream flow rate in shared lane (veh/hr); $v_{LT}$ is the minor street left-turn stream flow rate in shared lane (veh/hr); $T$ (hr) is the analysis time period; $c_{SH}$ is the capacity of the shared lane (veh/hr):

$$c_{SH} = \frac{v_{RT} + v_{LT}}{v_{RT} + v_{LT} + c_{m,RT}} + \frac{c_{m,RT}}{c_{m,LT}}$$

where: $c_{m,RT}$ is the movement capacity for minor street right-turn stream as if it had its own separate lane (veh/hr); $c_{m,LT}$ is the movement capacity for minor street left-turn stream as if it had its own separate lane (veh/h). The movement capacity $c_{m,x}$ for a given stream $x$ can be calculated correcting the potential capacity $c_{p,x}$ for the same stream with an adjustment factor (according to the suggestions in HCM 2000). The classic Harders’ formula has been used to determine the potential capacity $c_{p,x}$ for a general minor street stream $x$:

$$c_{p,x} = \frac{3600}{v_{c,x}}$$

where $v_{c,x}$ is the conflicting flow rate (veh/hr) for minor street stream $x$; $t_c$ (sec) is the critical gap for minor street stream $x$; $t_f$ (sec) is the follow-up time for minor street stream $x$. In equation (4) $v_{c,RT}$ is equal to $v_{p1}$ and $v_{c,LT}$ is equal to $v_{p1} + v_{p2}$ (Fig. 1).

The service time $d_s$ (sec/veh) can be calculated by means of the following equation:

$$d_s = \frac{3600}{c_{m,RT}}$$

where $c_{m,RT}$ is the minor street right-turn stream movement capacity which can be determined again correcting the minor street right-turn stream potential capacity with an adjustment factor; the conflicting volume $v_{c,RT}$ = $v_{p1}$ must be included in equation (4). It should be noted that the average control delay $d$ (equation (1)) for the minor street right-turn stream in a shared lane has been subdivided in the calculus of the average queue waiting time $d_q$ and in the calculus of the average service time $d_s$. The cost function for $d_q$ is a non-separable cost function which depends on the minor street right-turn flow rate but also on the minor street left-turn flow rate (equation (2)), as the lane is shared. But $d_s$ is also dependent on $v_{p1}$ and $v_{p2}$ because in equation (2) there is the shared lane capacity $c_{SH}$ which depends on $c_{m,RT}$ and on $c_{m,LT}$ (equation (3)). These variables depend, in their turn, on $v_{p1}$ and $v_{p1} + v_{p2}$, respectively. From the point of view of the conflicting volume, the average service time $d_s$ instead, depends only on $v_{p1}$ (Fig. 1). In the end the minor street right-turn average delay in a shared lane is expressed by a non-separable cost function in which the independent variables are the left-turn flow rate and the flow rates of the major street in both directions $v_{p1}$ and $v_{p2}$, in addition to the right-turn flow rate.
In the case of the right-turn stream from a minor street in an exclusive lane, $d_q$ could be determined by means of an equation being similar to (2):

$$d_q = 900T \left\lfloor \frac{V_{RT}}{C_{m,RT}} - 1 + \left\lfloor \frac{V_{RT}}{C_{m,RT}} - 1 \right\rfloor^2 \right\rfloor \left( \frac{3600}{C_{m,RT}} \right) \left( \frac{V_{RT}}{450T} \right)$$

(6)

the meanings of the terms in (6) are the same as in (2) and (3). Potential capacity $C_{p,RT}$ could be determined by means of Harders’ formula:

$$C_{p,RT} = V_{c,RT} \frac{e^{-V_{c,RT}T_{c,RT}/3600}}{1 - e^{-V_{c,RT}T_{c,RT}/3600}}$$

(7)

in which $V_{c,RT}$ is equal to the volume $V_{p1}$ only (Fig. 2). The movement capacity $C_{m,RT}$ can be calculated, again, correcting the potential capacity $C_{p,RT}$ (7) with an adjustment factor (HCM 2000).

The service time $d_s$ (sec/veh) can be calculated by means of equation (5).

The cost function for $d_q$ is a non-separable cost function which depends only on the minor street right-turn flow rate and not also on the minor street left-turn flow rate (as in equation (2)) as the lane is exclusive. But $d_q$ is also dependent on $V_{p1}$ because in equation (6) there is the movement capacity $C_{m,RT}$ for minor-street right-turn stream which depends on $V_{p1}$. From the point of view of the conflicting volume, the average service time $d_s$, instead, depends only on $V_{p1}$ (Fig. 2). In the end the minor street right-turn average delay is expressed by a non-separable cost function in which the independent variable is the flow rate of the major street $V_{p1}$, in addition to the right-turn flow rate $V_{RT}$.
3. Application of the model

During the first phase, the algorithm was tested on a “toy” network formed by 40 links, 31 nodes and 3 centroids (nodes at which trips originate and/or terminate). For this network we have considered one unsignalized T-intersection, one modern roundabout and one signalized intersection in a detailed representation. The simulations developed on the “toy” network were focused on solving the computational problems occurring in the implementation of the diagonalization algorithm with the detailed representation of the intersections and the detailed cost functions used. In order to study the effectiveness of the flow assignment when the volume to capacity ratio is close to 1, we have considered a set of 5 OD demand matrixes, the elements of which were gradually increased. In the second phase the algorithm was applied to the real road network of Villafranca characterized by 55 intersections, 37 of which are unsignalized and, among these latter ones, 6 are modern roundabouts. The O/D matrix, which has been provided by the Municipality of Villafranca (Verona), corresponds to the private transport demand of the morning peak hour (8:00 – 9:00). This matrix was initially estimated basing on 1991 census data and a sample of interviews; then the O/D matrix was improved using traffic counts carried out over the years 1999-2001.

For each link of the network the flow rate in terms of veh/h, the flow vs capacity ratio and the control average delays were calculated. Some factors of the cost functions, such as critical gap, follow-up time and impedance factors, were calculated basing on field observations.

The simulated flow rates were compared with the traffic counts made on 68 links (Figure 3) using Hi-Star portable traffic analyzer.

![Figure 3. Distribution of measured and simulated flow rates](image)

This comparison with the aim at verifying the effectiveness of the model was developed by calculating the mean square error (MSE) and the root mean square error (RMSE%) of the measured link flow vector \( f^* \) and the simulated link flow vector \( \tilde{f} \):

\[
MSE(f^*, \tilde{f}) = \frac{1}{N} \sum_{i} (f^*_i - \tilde{f}_i)^2 = 881.89
\]  \hspace{1cm} (6)

\[
RMSE\% = \left( \frac{MSE(f^*, f)}{\sum f_i} \right)^{0.5} = 6.59\%
\]  \hspace{1cm} (7)

287
where $N$ is the number of the counted links. The results of the comparison (Fig.2, equations (6) and (7)) highlight that the proposed model fits well the measured link flow vector as shown by the low values of MSE and RMSE%. The effectiveness of the model was also validated by comparing the average control delays $d^*$, calculated using HCS 2000 and referring to the measured link flow rates, with the average control delays $d$, calculated using detailed non-separable cost functions, but referring to the simulated link flow rates. This comparison, made on 4 unsignalized intersections with the higher flow rates, shows that $d$ fits well $d^*$.

4. Conclusions

This paper describes the implementation of an asymmetric user equilibrium route choice model with non-separable cost functions. We have developed an asymmetric equilibrium software in Visual Fortran code for studying real urban intersections in their detailed configurations and particularly unsignalized intersections. The procedures adopted for modelling the average delay of each link at an unsignalized intersection consider the delay as a sum of two terms: the average queueing delay and the average service delay. The former one is the average delay related to the time necessary to reach the stop line after joining the end of the queue; the latter one is the average time necessary to carry out the manoeuvre. In particular, we have considered different conflicting flows for queueing delay and for service delay, according to the right of way priority of each traffic stream at the intersection. We have also determined the potential capacity, the movement capacity and the shared lane capacity for each link at an intersection according to the methodology outlined in the Highway Capacity Manual 2000, after setting the critical gap, the follow-up time, and the impedance factor. The developed software was tested first on a “toy” network and then on the real network of Villafranca near Verona. The computational results of assignments were compared with the real flows recorded in 68 links of the Villafranca network. These comparisons highlighted the fact that the model framework fits well the real one. Moreover, the effectiveness of the model was supported by the comparison of the average control delays, calculated by means of the Highway Capacity Software 2000 for unsignalized intersections, and basing on the real flows measured, with the simulated average control delays.

References


ANALYSIS ON THE MECHANISM OF CONGESTION ON THE Merging SECTION OF URBAN EXPRESSWAY USING VIDEO IMAGE DATA

Fumitaka KURAUCHI¹, Nobuhiro UNO, Akito HIGATANI and Yasunori IIDA

Abstract. To understand the reasons of traffic congestion, the relationship between macroscopic traffic characteristics such as space mean velocity and microscopic traffic characteristics such as individual vehicle manoeuvres should be analysed all together. By the recent progresses on the development of image processing technologies, it is technically possible to obtain both macroscopic and microscopic flow indices from video image data automatically. This study discusses about the mechanism of the congestion occurred on the merging section of the urban expressway by video data. The video survey was conducted on the merging section of the Hanshin Expressway in Osaka, and by the automatic vehicle recognition system, vehicle trajectories are extracted. Traffic flow indices such as traffic volume, space mean density and velocity are then calculated. By these indices, this paper attempts to understand the mechanism of congestions.

1. Introduction

Accurate estimations of traffic flow indices such as volume, density and velocity are indispensable for understanding the traffic flow characteristics. These indices are essential for transport facility planning, traffic management, and roadway engineering. From the viewpoint of traffic flow theory, a better understanding of individual drivers’ behaviour is believed to lead to improvements on representation of traffic flow phenomena. At the same time, recent studies [1], [2] have shown that the quality of the current empirical data, with respect to collection methods, are questioning traditional theories on traffic flow modelling. The use of video cameras for collecting fuller and more stable traffic parameters is one promising method to verify both issues.

The increasing use of video cameras to observe road traffic conditions has provided the opportunity of applying vehicle tracking technologies to calculate traffic flow indices.

¹ Department of Urban Management, Kyoto University, Katsura Campus, Nishikyo, Kyoto, 615-8540, Japan, kurauchi@trans.kuciv.kyoto-u.ac.jp
Many methodologies have been proposed to obtain vehicle trajectories automatically, and vehicle identification by video image processing technology can be a powerful method to obtain traffic flow data ([3] for example). Recently, the authors have evaluated the methodology to measure traffic flow characteristics such as traffic volume, density and velocity through time from video image data [4]. We have concluded that the image processing technique has a good performance at estimating space mean velocity. Although there are still some errors in estimating traffic volume and density, we can at least discuss the relative change in traffic flow by the estimated indices. This study attempts to understand the mechanism of congestion occurrence at the merging section of the urban expressway by using video data.

2. Traffic Observation System

2.1. Automatic Vehicle Recognition System

This study utilises the automatic vehicle recognition system developed by Sumitomo Electric Industries Co., Ltd., and the detail specification of the system is out of the scope of this study. Basically, the system extracts the vehicles from the video data by comparing the current image and background image created by the images of former time slices. The image processing algorithm provides locations of the recognised vehicles every 0.2 second. During data processing, each different vehicle recognised is associated with a number (named ID) and its longitudinal and horizontal positions were automatically recorded along time.

2.2. Calculation of Traffic Flow Indices

The fundamental traffic flow indices are velocity, density and flow rate. In this work we present a way of averaging those characteristics proposed by Edie [5]. Edie derived consistent definitions of flow, density and velocity that can be applicable to all kinds of measurement. In Edie’s derivation, each vehicle trajectory through a given area \((A=x_1t_1)\) of space and time is considered as a vector \((x_i, t_i)\), where \(x_i\) is the distance travelled and \(t_i\) is the time taken by the \(i\)th vehicle within the measurement area. Thus, average flow \((q)\), density \((k)\) and velocity \((u)\) are calculated by averaging such vectors in space and time according to the following equations.

\[
q(A) = \sum x_i / A
\]

\[
k(A) = \sum t_i / A
\]

\[
u(A) = \sum x_i / \sum t_i = q(A) / k(A)
\]
3. Data Observation and Processing

The data used in this study are obtained from video cameras installed on tall buildings located at the merging section of the Moriguchi Line and the Loop Line of Hanshin Expressway, Osaka (Figure 1). 10 video cameras are installed and the video was taken from the 25th to the 31st of March, 2005. The length of the road section observed was about 200 metres in total. Severe traffic congestion frequently occurs at this merging section, and usually long queue grows toward the upstream of Moriguchi Line. By the existing findings [4], the flow indices are calculated by the area of 5 metres x 1 minute. Figure 1 illustrates the research site. Vehicle trajectory data obtained from 5 video cameras around the merging section are used in this study. The 1st and 2nd lanes are regarded as the Moriguchi Line and 3rd, 4th and 5th lanes as the Loop Line. To explicitly compare the traffic flow of the Moriguchi Line and the Loop Line, traffic flow indices are calculated separately even after merging. Note that the unit for y-coordinate is 10cm.

![Figure 1. Research Site.](image)

4. Traffic Flow Analysis by the Obtained Data

From the vehicle trajectories, traffic volume, space mean density and velocity for one week are calculated and analysed. Because of the limitation of the space, we will further discuss the traffic condition on the 29th of March. Figure 2 illustrates the transitions of space mean velocity. Solid line represents the velocities on the Moriguchi Line, and dotted line represents the velocities on the Loop Line. At the most downstream sections (y=2500), the fluctuations of the space mean velocity on both lines are almost same and therefore traffic conditions on this section is uniform among lanes. At the section of y=2000, the values of space mean velocity is still almost the same but especially from 14:00 to 17:00, the velocity...
Figure 2. Transitions of Space Mean Velocity (3/29)
of the Moriguchi Line is a little slower. At the location of y=1500 (20m upstream of the merging section), the value of space mean velocity fluctuates enormously especially during the night time. We have not identified the reason why it happened, but it seems that the velocity on the Loop Line is higher than that of the Moriguchi Line around this section. Then at the location of y=1000, the variation of the Loop and the Moriguchi Line is not similar any more. Especially after 12:00, the average velocity of 50 km/h has been maintained on the Loop Line, but on the Moriguchi Line, the space mean velocity sometimes fell down to 20 km/h. This phenomenon gets clearer at y=500. On the Loop Line, the congestion only occurred around 9:30 to 10:30 whereas the congestion continued on the Moriguchi Line from 7:00 to 13:00 and 14:00 to 16:30. In conclusion, on the 29th of March, the deceleration first occurred around 7 o’clock at further downstream of the study area, and it propagated up to the merging section. When the congestion reached to the merging section, the different phenomena were observed among 2 lines. On the Moriguchi Line, the space mean velocity decreases in accordance with the deceleration propagated from the downstream section. On the other hand, deceleration did not propagate to the further upstream of the Loop Line.

To understand the reason why the deceleration propagated only to the Moriguchi Line, the spatiotemporal transitions of space mean density is explored. As is shown in the Figure 2, the space mean velocity on the Moriguchi Line recovered temporarily around 1300, and then it dropped again around 1400. Figure 3 represents the spatiotemporal transition of the space mean density during that time. The figure suggests that the key difference of the traffic condition was that the density did not change in the case of the Moriguchi Line whereas it increased on the Loop Line. This is because there are three lanes on the Loop Line and in the case of the Loop Line, the density could increase and the same traffic flow rate could be maintained even the deceleration occurred. However on the Moriguchi Line, since the density could not increase any more, the traffic flow level decreased in accordance with the deceleration. By these effects, once the deceleration happens, the traffic flow from the Moriguchi Line decreases when the congestion occurs. However the Moriguchi Line, since the density could not increase any more, the traffic flow level decreased in accordance with the deceleration. By these effects, once the deceleration happens, the traffic flow from the Moriguchi Line decreases when the congestion occurs. However once the congestion occurred at around 7:00, the merging rate dropped down to 0.25 and it did not improve until around 19:00, when the congestion diminished. We also found that this is because of the lower usage of the 1st lane compared with the 2nd. It can be concluded that some traffic measures to encourage the usage of the 1st lane for recovering the merging rate of traffic volume from the Moriguchi Line is needed to relax the congestion at this section.

5. **Congestion Mechanism on This Merging Section and Recommended Strategies**

We have explored the spatiotemporal transitions of traffic flow, and the mechanism of congestion at this section can be summarised as Figure 5. With some traffic flow turbulences either by lane changing or merging/diverging behaviour, the deceleration occurs from the downstream section. These turbulences are the triggers of the congestion. When the decrease in velocity expands to the merging section, traffic flow entering from Moriguchi Line decreases and the congestion occurs on this Line.
(a) Loop Line

(b) Moriguchi Line

Figure 3. Spatiotemporal Transition of Space Mean Density

Figure 4. Transitions of Merging Rate of Traffic Volume from Moriguchi Line
Heavy traffic volume

Disturbances

Back propagation of deceleration

deceleration reaches to the merging section

- Deceleration from further downstream
- Deceleration to find the gap for lane changing
- Conflict at merging section

Disturbances

- Increase of the density (the drivers start using the 3rd lane)
The flow rate keeps unchanged
- Strength of merging decreases

Strength of merging increases

The congestion remains!!

Figure 5. Mechanism of the Congestion at This Merging Section.

Once the congestion occurred in the Moriguchi Line, the merging rate from this Line become lower to around 0.25 to 0.40, and the congestion remains on the Moriguchi Line until the traffic flow rate from the Loop Line decreases. To mitigate congestion occurring here, we should discuss how to reduce the trigger of the congestion and to recover the merging rate from the Moriguchi Line separately. For recovering the traffic flow rate from the Moriguchi Line during the congestion, it is important to improve the inequality in lane usage on the Moriguchi Line. Also to discourage to use the third lane of Loop Line may contribute to increase the traffic volume entering from the Moriguchi Line since the vehicles flowing from the second lane can easily merge when the traffic flow on the third lane is low. To do this, lane use guidance to encourage the usage of the first lane might be effective. To mitigate the conflicts of the vehicles which might become a trigger of the congestion, more advanced technologies such as vehicle-to-vehicle communication might be needed. Possible strategies are summarised as Table 1.

<table>
<thead>
<tr>
<th>Objective</th>
<th>Measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>To prevent the occurrence of the trigger of the congestion</td>
<td>Controlling individual vehicle manoeuvres by ITS</td>
</tr>
<tr>
<td></td>
<td>Changing lane configuration</td>
</tr>
<tr>
<td></td>
<td>Information provision to discourage the use of the 3rd lane of Loop Line</td>
</tr>
<tr>
<td></td>
<td>Vehicle-to-vehicle communication for co-operated merging</td>
</tr>
<tr>
<td>To improve the flow rate of Moriguchi Line after the occurrence of congestion</td>
<td>Information provision</td>
</tr>
<tr>
<td></td>
<td>- to encourage using the 1st lane of Moriguchi Line</td>
</tr>
<tr>
<td></td>
<td>- to discourage using the 3rd lane of Loop Line</td>
</tr>
<tr>
<td></td>
<td>Controlling gaps among groups of vehicles by ITS</td>
</tr>
<tr>
<td></td>
<td>Inflow control / ramp metering</td>
</tr>
</tbody>
</table>

Table 1. Possible Strategies for the Congestion Mitigation.
Summary

This study attempted to understand the mechanism of the congestion at the merging section on Hanshin Expressway by using image data. By the video image data, spatiotemporal transition of traffic flow can be observed and also the traffic condition can be visually confirmed. This study used the video data obtained at the merging section of Hanshin Expressway and the mechanism of the congestion occurrence at the section was analysed. Consequently, it is suggested that we should discuss separately how to reduce the incident which can be a trigger of the congestion, and how to recover from the congested situation. The possible strategies for handling these two issues are also summarised. Although the result might be only valid on this specific merging section, the analysis using video image is a promising methodology to understand the traffic flow more in detail.

Acknowledgments

This study is the outcome of the “ITS Implementation Project” of Japan Society for Civil Engineering supported by National Institute for Land and Infrastructure Management. The authors are also thankful to Sumitomo Electric Industries Ltd. to provide image processing technologies and Hanshin Expressway Company, Ltd. for data observation.

References


ON THE CONVERGENCE OF WEIGHTING METHODS IN MULTI-ATTRIBUTE ASSESSMENT OF TRANSPORT PROJECTS

Paolo DELLE SITE\(^1\) and Francesco FILIPPI\(^1\)

Abstract. The paper compares three weighting methods in multi-attribute models: ratio with swings, AHP scale with swings and trade-off. An experiment is set up for a decision problem on transport projects. Methods are assessed with respect to convergent validity of weights and of ranks of alternatives, and perceived difficulty to provide judgments. Trade-off turns out to be an outlier for convergent validity and is perceived as most difficult.

1. Introduction

Multi-attribute value theory (MAVT) and analytic hierarchy process (AHP) are well known multi-criteria methodologies. Both use a functional approach. AHP is widely used also for transport applications (see e.g. Delle Site and Filippi [4]). In AHP, in the version originally proposed by Saaty, elicitation of weights of the criteria cannot be anchored to alternatives as consistency with the mathematical structure of the model would require anchoring to alternatives which can be ill-defined (for a discussion of the issue see Dyer [5] and Ferrari [6]). In MAVT criteria scales are set up with reference to well-defined anchor points: this makes it possible both the scores of the alternatives on the criteria and the weights of the criteria be normalized with respect to the same range over which the alternatives vary on each criterion.

There are different weight elicitation methods which are consistent with the classical additive model used in MAVT. Theoretical rigour requires all methods be implemented with explicit reference to the attribute ranges defined by the anchor points. With ratio weights (von Winterfeldt and Edwards [12]) and semantic scales, such as the AHP scale (Saaty [9]) and the MACBETH scale (Bana e Costa and Vasnick [1]), the weight eliciting questions are derived from an interpretation of weights as swings and the respondent is

\(^1\) DITS, Dipartimento Idraulica Trasporti Strade, University of Rome “La Sapienza”, Via Eudossiana 18, 00184 Rome, Italy, paolo.dellesite@uniroma1.it
asked to compare swings among different attributes. The trade-off method (Keeney and Raiffa [7]) uses questions where the respondent is asked to adjust attribute outcomes to achieve indifference among alternatives.

If different weigh elicitation methods can be used within the same modeling framework, the research question arises whether different methods yield significantly (in statistical sense) different results in terms of weights and of ranks of alternatives. This type of investigation is helpful to detect behavioral biases. Another question is how respondents perceive the methods in terms of difficulty. Several experimental studies have addressed these questions in different applied decision contexts (Bell et al. [2]; Borcherding et al. [3]; Pöyhönen and Hämäläinen [8]; Schoemaker and Waid [10]). Correlation was used as a measure of the convergence degree between data sets of results. None of those studies, however, has carried out experiments on decisions relating to transport projects.

The paper presents the results of an experiment aimed at assessing the convergence properties and perceived difficulty of different weight elicitation methods in additive MAVT models. Methods compared include ratio weighting with swings, AHP scale with swings, trade-off. The experiment is carried out for a decision between layout alternatives of a metro line.

2. Attribute weighting

Weight elicitation methods must be consistent with the underlying mathematical model. The model here is the classical additive model of MAVT:

\[ v(x) = \sum_{i=1}^{n} w_i \cdot v_i(x_i) \]

where:
- \( v(x) \) value function of the alternative
- \( x \) attribute vector of the alternative
- \( w_i \) weight, or scaling constant, of attribute \( x_i \)
- \( v_i(x_i) \) single-attribute value function (score) on attribute \( x_i \)
- \( n \) number of attributes
- \( x_i^- \) worst outcome of attribute \( x_i \) (anchor point)
- \( x_i^+ \) best outcome of attribute \( x_i \) (anchor point)

The interpretation of weights as swings is based on the following. Given the two alternatives:

\[ x^1 = [x_i^-; \ x_i = x_i^+ k \neq i] \]
\[ x^2 = [x_i = x_i^-; \ x_i = x_i^+ k \neq i] \]

we get:

\[ v(x^2) - v(x^1) = w_i \cdot v_i(x_i^+) - w_i \cdot v_i(x_i^-) = w_i \]
Ratio weighting and weighting with AHP scale are elicitation methods which can be used consistently with this interpretation. In ratio weighting 100 points are assigned to the attribute with the highest value swing. The other swings are valued by judgment in terms of percentage of the highest value swing. In weighting with the AHP scale for each pair of attributes the dominance of the swing in the first attribute on the swing in the second attribute is assessed by judgment according to the Saaty semantic scale.

The interpretation of weights as trade-offs is based on the following. Given the alternative:

\[ x^* = [x_i = x_i^*; \ x_j = x_j^*; \ x_k = x_k^* \quad k \neq i, j] \]

we can search the outcome \( x_i^* \) of the attribute \( x_i \) that makes indifferent to \( x^* \) the following alternative:

\[ x^b = [x_i = x_i^*; \ x_j = x_j^*; \ x_k = x_k^* \quad k \neq i, j] \]

We get

\[ \frac{w_i}{w_j} = \frac{1}{1 - v(x_i^*)} \]

In trade-off weighting the outcome of the attribute \( x_i \) is adjusted by judgment \( n-1 \) times in order to assess the \( n-1 \) ratios \( w_i/w_j \).

3. Experimental design

The experiment considers three weight elicitation methods: ratio weighting with swings, AHP scale with swings, and trade-off.

To formulate the hypotheses of the experiment we define convergent validity as the degree to which items that should be correlated theoretically are correlated in reality. The hypotheses of the experiment are:

- HP1: convergent validity between weights elicited with different methods depends on the methods,
- HP2: convergent validity between ranks of alternatives obtained with different weight elicitation methods depends on the methods,
- HP3: weight elicitation methods differ in the perceived ease to provide judgments.

The decision problem relates to the layout alternatives for the fourth line of the metro network in Rome. Three layout alternatives are considered. Criteria include one cost attribute, construction costs, and three benefit attributes, travel time, safety and air quality. All attributes have quantitative descriptors appraised by the planning unit of the municipality. Descriptors of benefits are yearly reductions, on the without-case, of respectively, hours spent travelling, accidents resulting in injury or death, and particulate matter emissions.

Single-attribute value functions are set up using local scaling (i.e. 0 and 1 scores are assigned to alternatives in the set of those to be ranked in the currently addressed decision problem) and linearity assumptions with respect to the quantitative descriptor of the attribute.

Attribute weights are elicited in the experiment with a questionnaire. Seventeen experts
in transport planning, including academics and practitioners, participated in the experiment. The questionnaire includes an introduction where the decision problem is explained and basic information on the alternatives are provided. These include descriptors for the four attributes used in the multi-criteria model. The questionnaire includes then three blocks of questions aimed at eliciting weights. Each block uses one elicitation method.

In the ratio and AHP methods the subjects are confronted with the change from the worst to the best outcome of each attribute. In the trade-off block the attribute that subjects are asked to adjust is construction costs and the questions are formulated by asking the willingness to pay of the society to achieve improvements in each of the benefit attributes. Trade-off is equivalent in this instance to a pricing out method. At the end of each block the subject is asked to rate (from 0 to 10) the elicitation method in terms of ease to provide judgments.

Normalised weights, derived from the three elicitation methods, are calculated for each subject. Ranks of the three layout alternatives, resulting from the different weight vectors, are finally calculated according to the additive model for each subject. For HP1 within-subject inter-method Pearson correlation coefficients are calculated and then averaged over subjects. Rank correlation coefficients are used instead for HP2. Statistical significance is assessed for all hypotheses using Wilcoxon matched pairs signed-rank tests.

4. Results

4.1. Hypothesis 1

HP1 is supported. Table 1 shows average correlations for pairs of methods with reference to weights. The correlation of ratio with AHP scale is significantly higher than the correlations of each of these two methods with trade-off (respectively p<0.0043 and p<0.0122).

Another interesting result from the analysis of weights is that the average weight of the cost attribute in the ratio method (average 0.21) and that in the AHP scale method (0.20) are lower than the average weight of the cost attribute in the trade-off method (0.48). This is significant at p<0.00016 and p<0.0010 respectively. Subjects take into account the range of cost variation to varying extent according to the elicitation method.

The finding that the cost attribute has a higher weight in the trade-off method can be explained in terms of the compatibility principle (Tversky et al. [11]). This states that stimulus components that are compatible with the response are weighted more heavily than those that are not, presumably because the former are accentuated. In the case here the effect is induced by the response scale as the trade-off method is implemented as a pricing out method.

The finding can also be seen as a manifestation of a loss aversion effect, well known in prospect theory, when it is recognised that money is given lower value if the question is about reducing costs borne (i.e. a gain) as in ratio and AHP scale, it is given higher value if the question is about bearing higher costs and willingness to pay (i.e. a loss) as in trade-off.
4.2. Hypothesis 2

HP2 is also supported. Table 2 shows average correlations for pairs of methods with reference to the ranks of the alternatives. The correlation of ratio with AHP scale is significantly higher than the correlations of each of these two methods with trade-off (respectively \( p<0.0371 \) and \( p<0.1162 \)).

<table>
<thead>
<tr>
<th></th>
<th>Ratio</th>
<th>AHP scale</th>
<th>Trade-off</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ratio</td>
<td></td>
<td>0.67</td>
<td>0.02</td>
</tr>
<tr>
<td>AHP scale</td>
<td>-</td>
<td>-</td>
<td>0.10</td>
</tr>
<tr>
<td>Trade-off</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 1. Average within-subject inter-method correlation of weights

<table>
<thead>
<tr>
<th></th>
<th>Ratio</th>
<th>AHP scale</th>
<th>Trade-off</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ratio</td>
<td></td>
<td>0.65</td>
<td>0.06</td>
</tr>
<tr>
<td>AHP scale</td>
<td>-</td>
<td>-</td>
<td>0.18</td>
</tr>
<tr>
<td>Trade-off</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 2. Average within-subject inter-method correlation of ranks of alternatives

4.3. Hypothesis 3

HP3 also is supported. Average and median ratings of ease to provide judgements (scale 0-10, 0 difficult, 10 easy) are 6.36 and 6.5 for ratio, 5.86 and 6 for AHP scale, 4.78 and 5 for trade-off. Ratio and AHP scale are rated significantly higher than trade-off (respectively \( p<0.026 \) and \( p<0.116 \)). The percentage of subjects who rated ratio as easiest is 64%, that of subjects who rated trade-off as most difficult is also 64%.

4.4. Insights into the application of multi-attribute models to transport projects

Some participants stressed they would have felt more comfortable if additional information had accompanied the descriptors of time and air quality benefits. While health damage effects are suggested unanimously for air quality, the recommendation is less firm for time benefits. A solution could be the distribution of travel time savings per trip, i.e. how many travellers save how many minutes, in the average weekday.

5. Conclusions

The paper has presented the results of a weight elicitation experiment within a decision
problem relating to transport projects. Results are in agreement with those found in the literature for other decision problems. Results suggest that selection of the weight elicitation method is of relevance in multi-attribute models as this affects both the weights and the resulting ranks of the alternatives. The findings here add to the body of experimental research documenting that weights are “constructed” in the elicitation process rather than “uncovered”.

References


ANALYSIS OF THE PUBLIC TRANSPORT SUPPLY IN A LOCAL LEVEL

Guillem ALSINA, Magín CAMPOS1, Francesc ROBUSTÉ

Abstract. A good planning for a suitable investment in public transport is needed. In this task it is important to have tools and data that can improve the analysis of the current state and possible alternatives. The present article shows a simple but effective methodology to know the quality that is being offered by certain public transport network.

1. Quality of the public transport supply

The basic stages of an unimodal journey in public transport can be summarized in: access to the bus stop (1), wait for the bus (2), accommodation into the vehicle (3), travel across the route bus until the final bus stop (4) and reach the destination point from the bus stop (5).

The quality that the user demands refers, basically, to only one attribute of the service based on the stage of the journey it is in. The main attributes associated to each stage are: bus stops density (1), frequency (2), comfort (3), speed (4) and bus stops density (5).

Summarising, we try to measure the quality offered in a public transport network considering the quality of these four basic aspects. To achieve this objective, four indicators have been defined, and each one evaluates the quality of one of the four attributes (Table 1): TAI, SAI, IQI and RQI. These indicators have to be an estimation the maximum of precise, objective and simple. Finally, the quality of the four attributes will be weighed to establish a global quality that will be called Global Service Indicator (GSI).

1 Center for Innovation in Transport (CENIT), Barcelona, Spain, magin.campos@upc.edu
<table>
<thead>
<tr>
<th>Stage of the journey</th>
<th>Main attribute of the public transport supply</th>
<th>Associated indicator</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 &amp; 5 Bus stop access</td>
<td>Bus stops density</td>
<td>Territorial Accessibility Indicator (TAI)</td>
</tr>
<tr>
<td>2 Wait for the bus</td>
<td>Frequency</td>
<td>Service Accessibility Indicator (SAI)</td>
</tr>
<tr>
<td>3 Accommodation into the vehicle</td>
<td>Comfort</td>
<td>Internal Quality Indicator (IQI)</td>
</tr>
<tr>
<td>4 Travel across the route bus</td>
<td>Speed</td>
<td>Route Quality Indicator (RQI)</td>
</tr>
</tbody>
</table>

Table 1 – Main attribute in each stage of the journey and its associated indicator.

2. Application field

This methodology can be applied to different geographic fields and public transport networks. Nevertheless, as an example, it has been applied to the city of Barcelona and its conventional bus network which is managed by a state-owned company (Transports Metropolitans of Barcelona, TMB).

2.1. Territorial accessibility indicator (TAI)

This first indicator measures the access facility towards the bus stops or from them. Logically the offered quality in this attribute will be proportional to the density of the bus stops. For that reason, the TAI indicator is defined as the number of bus stops by square kilometre, as it is showing in the equation (1).

\[
TAI_i = \frac{NBS_i}{A_i} \cdot 10^6
\]

Where,
- \( TAI_i \): Territorial Accessibility Indicator of the zone \( i \), bus stops/km\(^2\)
- \( NBS_i \): number of bus stops that zone \( i \) contains, bus stops
- \( A_i \): surface of the zone \( i \), m\(^2\)

2.2. Service accessibility indicator (IAS)

The second indicator tries to be an estimation of the user facility to accede to any bus expedition. That is to say, the greater frequency of expeditions, the smaller delay and
therefore better quality is received by the user in this stage of the journey. For this reason the indicator IAS of a certain zone is defined as the average of the frequencies (or number of expeditions per hour) at the bus stops that belong to this zone (Equation 2).

\[
SAI_i = \frac{1}{m} \sum_{k=1}^{m} \left( \sum_{j=1}^{n} \frac{ExpDay_j}{HoursDay_j} \right)
\]

(2)

Where,  
- **SAI** : Service Accessibility Indicator of the zone i, expeditions/hour  
- **ExpDay** : number of daily expeditions of the route j, expeditions/day  
- **HoursDay** : daily period of circulation of the route j, hours

2.3. Internal quality indicator (IQI)

One way to measure the comfort enjoyed by the passengers is to relate the number of free seats that are offered daily with the demand of travellers. Consequently the index of the comfort quality that is offered in a zone is just the average comfort defined for the different bus stops within this zone (Equation 3).

\[
IQI_i = \frac{1}{n} \sum_{k=1}^{n} \frac{FreeSeats_k}{BusJourneys\_CA_k}
\]

(3)

Where,  
- **IQI** : Internal Quality Indicator of zone i, free seats per passenger  
- **FreeSeats** : daily vacancies offered in the bus stop k, free seats  
- **BusJourneys\_CA** : daily number of bus journeys from the area covered by the bus stop k, journeys

2.4. Route quality indicator (RQI)

The last of the four indicators analyses the quality offered during the route. In this stage the bus users expect a good commercial speed that allows them to arrive as rapidly as possible at their destiny.

Once the commercial speed of each route is known a weighed average by the expeditions is used in a particular bus stop. For a zone i it is just deduced the average of the different bus stops that it contains (Equation 4).
\[
RQI_i = \frac{1}{m} \cdot \sum_{j=1}^{n} \frac{\sum_{k=1}^{m} V_j \cdot \text{ExpDay}_j}{\sum_{j=1}^{n} \text{ExpDay}_j}
\]

Where,
- \( RQI_i \): Route Quality Indicator of the zone \( i \), km/h
- \( \text{ExpDay}_j \): daily number of expeditions of the route \( j \), expeditions/day
- \( V_j \): commercial speed of the route \( j \), km/h
- \( n \): number of routes that give service at bus stop \( k \), routes
- \( m \): number of bus stops that belong to zone \( i \), bus stops

2.5. Global service indicator (GSI)

Global Service Indicator (GSI) with the purpose of synthesizing the quality measured in the different stages in a single indicator is defined. To calculate it two previous steps are needed: the normalization of the indicators in each zone and their weighing according to its importance (Equation 5).

\[
IGS_i = \alpha_{IAT} \cdot IAT_i^N + \alpha_{IAS} \cdot IAS_i^N + \alpha_{ICI} \cdot ICI_i^N + \alpha_{ICR} \cdot ICR_i^N
\]

Where,
- \( IGS_i \): Global Service Indicator of zone \( i \)
- \( N \): superscript that indicates the standardization of the indicator
- \( \alpha_j \): weight of the indicator \( j \)

3. Conclusions

A new methodology has been developed that allows to analyse the quality offered in a public transport network by comparing zones of a particular geographic scope. The main advantage of this procedure is that it requires few data and in addition these data are easy to obtain by the operator of the network.

This methodology has been applied to the conventional network of buses in the municipality of Barcelona by means of its division in 250 zones. The obtained results allow to conclude that this network offers a different quality depending on the zone where the user is (figure *1). In spite of having better global quality in the zones of more population the differences among zones nearby, and even adjacent, are remarkable.
This study serve to locate the deficit of quality based on the area of Barcelona and know in what attribute or attributes of the quality is necessary to act in a high-priority form.
A SIMULATION TOOL FOR THE EVALUATION OF STRATEGIES FOR PRODUCTIVITY IMPROVEMENT OF CONTAINER TERMINALS

Antonio BARBARA¹, Cristiana DEGANO², Pierluigi FICARA³, Flavio MARANGON⁴, Pasquale MODICA⁵, Giulio VALFRE⁶

Abstract. The optimization of activities associated with intermodal terminal operations and the consequent reduction of operative costs are the most important problems to face for increasing the flexibility and the dynamic capacity of terminals. This paper proposes a simulation model for container terminal analysis developed within SITRAC project. The simulation model has been developed using process-oriented approach, and using Extend software package in order to study the current state of a container terminal and possible future expansions to handle increased throughput. In particular, the main issue is the way in which the use of terminal resources can be modeled to analyze the current behavior of a terminal and to evaluate possible future situations.

1. Introduction

An intermodal terminal is a critical transfer point for moving cargo through the transportation and distribution process, but transfer between vessels and inland transportation is one of the weakest, least efficient, and most costly links in the intermodal
chain. Then, the intermodal terminals represent a critical factor within the distribution network that must be faced in order to prevent the penalty costs and the financial loss through the logistics chain as a whole.

The transfer of cargo among various modes of transport (concept of intermodality) has become an integral part of the maritime industry.

In the freight transportation context, the management of terminal is a very interesting topic that researchers have to deal with. In the related literature, most works concerning with intermodal freight terminal management are devoted to the investigation of models and strategies able to optimize: i) the performances of the terminal in term of cost-savings and quality improvements in the handling systems [1]; ii) the allocation of storage space for containers in order to minimize the total number of re-handles, so as the number of containers handling facilities to reduce the turn-around time of container carriers in the terminal [2, 3, 4].

Furthermore, a significant research stream has developed regarding the modelling of intermodal freight terminals, so as to design suitable management and control strategies [5, 6, 7]. Holguin-Veras and Walton [8], for instance, created a micro-simulation tool that captures the geometry of the system and the interactions with the operational policies. The proposed simulation is also able to model the time required for the movement and operation of yard cranes, gantry cranes and other equipment.

This paper proposes a simulation container model developed in the context of SITRAC (Ricerca su Simulatori a Supporto dello Sviluppo di una Rete di Trasporto Intermodale basata sul Cabotaggio), a project funded by MIUR and whose partners are D’Appolonia, ENEA-ENETEC, Uniontrasporti.

One of the two primary objectives of SITRAC is to provide an agile and flexible tool for modeling freight terminals in order to collect quantitative information related to freight handling time and costs associated with depending on the adopted technological solution so as the organizational one. Then, the main focus of the project is to create a flexible and easy-to-use simulation tool in order to understand the dynamical behavior of freight terminals, to identify and quantify improvement potentials, and to test and verify how improvement potential can be realized.

In investigating ways in which maritime freight terminal can improve efficiency, this paper outlines the container terminal simulation model, developed in the SITRAC context, and provides components architecture that have been developed within Extend simulation environment. In particular, the simulation model is related to a real life detailed processes concerning the inbound and outbound container handling in a container terminal. Furthermore, the model is calibrated and fitted to the characteristics of the Salerno Container Terminal (SCT), Italy, used as a case study.

2. The SITRAC tool

As said above, the aim of SITRAC is to provide a simulation tool which allows to evaluate and improve operational and financial performance of a freight terminal under realistic conditions. Based on this simulation tool, the terminal’s management should be able to evaluate different possible alternatives and to identify optimization potentials.

The maritime container terminal in itself includes a multitude of interacting factors. While analytical models offer a quick and general approach for representing a given problem, they lose in detail and flexibility. Therefore, simulation modeling is better suited for the random and complex environment of a container terminal, especially when several parameters and scenarios need to be investigated.
With no loss of generality, and to comply with the characteristics of the real system that has been studied, the physical and the simulation models of the considered container terminal are hereafter shortly described.

2.1. The physical model: resources and processes

The design of the physical model of the terminal requires the definition of all entities, which compose a container terminal service process. Then, all the resources, which physically make up the system, have been defined. Similar considerations have been made about dynamic features of the system: all different kinds of operating processes, which can be in such a system, have been formalized.

A. Resources

In this work, the container terminal is thought as divided into areas, each one associated with a kind of container handling operation. The areas are the following: the operation area, also called quay, the yard, and the in/out gates.

As the container movements inside the terminal required for the loading and discharging process are carried out by mobile vehicles and crane installations, each area is equipped with its own material handling system, made up of a set of resources capable of performing more than one operation at a time. More specifically, the quay is equipped with a set of own resources

\[ P = \{ p^i_k; \quad i = 1, \ldots, p; \quad k = 1, \ldots, \alpha_i \} \]

quay cranes, able to load/unload the containers to/from the ships. Quay cranes are of \( p \) different types, and there are \( \alpha_i \) quay cranes of each type \( i = 1, \ldots, p \). Furthermore the quay is equipped with a bounded area, called buffer area, where containers are placed while waiting for the material handling facilities when not immediately available.

The yard, is equipped with a set of dedicated resources

\[ T = \{ t^i_k; \quad i = 1, \ldots, t; \quad k = 1, \ldots, \gamma_i \} \]

yard cranes, able to load/unload the container in the yard. Yard cranes are of \( t \) different types, and there are \( \gamma_i \) yard cranes of each type \( i = 1, \ldots, t \). Like the quay it has a bounded area for containers waiting to be moved.

The in/out gates are resources themselves; it is assumed that the set of gates is divided into three subsets gathering the input gates to the terminal, the output gates, and the gates that can be used to both enter and leave the terminal, according to its current setting \[9\].

Formally, the terminal is equipped with a set of shared material handling vehicles, with fixed load capacity

\[ V = \{ v^i_k; \quad i = 1, \ldots, \nu; \quad k = 1, \ldots, \delta_i \} \]

where \( v^i_k \) is the \( k^{th} \) vehicle of type \( i \) and \( \delta_i \) the number of vehicles of type \( i \).
B. Processes

The considered container terminal can be viewed as an intermodal node that supports two different modes of transport: sea and road. Each container has to be moved from its origin to its destination inside the terminal through a fixed sequence of operations. Then, containers are handled in the terminal according to two different processes, essentially distinguished into two typologies: export and import. Container requiring an export process arrives to the terminal by land, and leaves it by sea, whereas container undergoing an import process arrives by sea and leaves by land. Furthermore, each process consists of a sequence of operation which have to be executed in the fulfillment of the constraints represented in the corresponding precedence-relation graph as depicted in Figure 1.

This means that, the containers, upon their arrival at the terminal, flow through its areas according to the processes they are associated with. Consequently, the system is characterized, at each instant time, by the presence of a certain number of containers flowing through the plant, following a specified routing which requires them to “visit” the areas in a given sequence, and requiring handlings to be performed on them from one or more resources at each visited areas [10]. In fact, any handled container is associated in an exclusive mode with one or more resources as regards each task belonging to its operating cycles. This means that different process operations cannot hold the same resource(s) simultaneously.

Then, the modeled terminal consists of resources to which the containers have to access as they flow through the system.

![Figure 1 Precedence relation graph of an export process](image)

2.2. The simulative model

According to the concept of operating processes highlighted in the physical model of the terminal, a process-oriented approach has been considered in order to generate the
simulation model. Furthermore, process-oriented approach clashes with the modeling logic of Extend simulation software that also highlights the Discrete Event System characteristics of the considered container terminal [11].

The model proposed in the SITRAC project belongs to the timed model class and to the one usually defined as simulative models. Furthermore, the analysis that has been done is longitudinal: this means that the system is described by the containers (clients) viewpoint.

Furthermore, at any instant time, the system state describes the tasks in process, the resources that containers currently hold and will acquire before the processing completion, and, finally, the sequence of all the resources necessary to complete each operating cycle.

In the Extend environment a system has to be modeled as a sequence of tasks, generally represented by blocks connected by lines. In the developed simulation model each block represents a specific physical entity, a resource (i.e., quay crane, yard crane, stacker, in/out gate), performing a task, while the double lines connecting blocks represent the containers' flow as depicted in Figure 2, representing the simulation model as whole.

Furthermore, a set of hierarchical blocks have been developed. As depicted in Figure 3, each hierarchical block is made of other blocks connected like they would be in the model. These blocks represent a portion of the model we call subsystems. Then, the simulation model representing the functioning of the system described above can be seen to be made up of subsystems, i.e., quay operation module, each one modelling a task belonging to the import or export operating sequence. The subsystems can be viewed as single discrete-event plants interacting with each other.

One of the main important characteristics of the simulation model is that, at any instant time, the system state describes the tasks in process, the resources that containers currently hold and will acquire before the processing completion, and, finally, the sequence of all the resources necessary to complete each process.

Once the simulation model and its computer implementation have been completed, a number of tests were conducted in order to validate the simulation model and compare different scenarios. To examine the impact of the operation scenarios, a base case model of the SCT has been created and has been validated by comparing its results with the real measurements of the Salerno Container Terminal. Furthermore, simulations are performed and compared in four scenarios using performance indices such as: i) average waiting time; ii) yard crane utilization; iii) quay crane utilization; iv) waiting time at the berth for the beginning of loading and unloading operations; v) waiting time at the gate for the beginning of inbound and outbound operations.

3. Conclusions

The developed simulation tool represents a decision support system able to evaluate the most useful actions (i.e., management policies, handling resources) in order to improve terminal attraction, that is the improvement of the overall system productivity.

Particularly, it provides the possibility in changing the configuration and the characteristics (i.e., processing time, set-up time, etc.) of each element in the modeled terminal in order to represents new technologies’ impact and new logistics’ policies.

Particularly, the mail goals of the simulation model are: i) evaluation of alternative operation scheduling policies; ii) evaluation of different resource allocation procedures. Furthermore, the flexibility of the model allows the considerations of many alternative scenarios, thus making it an excellent tool to incorporate cost analysis features.
References


AN ENSEMBLE REGRESSION APPROACH FOR BUS TRIP TIME PREDICTION

João M. MOREIRA*
Jorge Freire de SOUSA†
Alípio M. JORGE‡
Carlos SOARES§

Abstract. This paper is about bus trip time prediction in mass transit companies. We describe the motivations to accomplish this task and how it can support operational management on such companies. Then, we describe a Data Mining framework that recommends the expected best regression algorithm(s), from an ensemble, to predict the duration of a given trip. We present results that show the advantage of using an ensemble regression approach.

1. Introduction

In recent years, several mass transit companies all over the world have done large investments in Advanced Transportation Management Systems [9]. These investments appeared mostly as a consequence of the enormous evolution in the areas of information and communication systems. The new Operational Control Systems (including GPS) are among the most well known examples of this evolution. They store location data per time unit as well as all the relevant records for operational control, such as, information about the bus, the driver, the crew duties and the vehicle duties, etc. This type of information is particularly important for the optimization of resources, namely, drivers and vehicles.

2. Motivation and description of the study

Mass transit companies use both the time of the planned and the effective duties in various ways. Of course, the closer the real duty is to the planned one the better, because the use of the resources and the quality perceived by the clients will both be improved. The vehicle

*Faculty of Engineering, University of Porto, Portugal, jmoreira@fe.up.pt
†Faculty of Engineering, University of Porto, Portugal, jfsousa@fe.up.pt
‡Faculty of Economics, University of Porto, Portugal, amjorge@liacc.up.pt
§Faculty of Economics, University of Porto, Portugal, csoares@liacc.up.pt
and the crew duties are defined as a sequence of trips, with a certain frequency and each one with a given duration. If the real trip time is longer than planned and the lag between consecutive trips is not enough to avoid overlapping, this will have consequences on the amount of extra money the company has to pay to the driver as well as on the image of the company. Trip time prediction may be used in four different ways:

- In the long term (several months): for the timetable definition;
- In the medium term (it can go from few days to several months, depending on the company’s culture): for the logic definition of crew duties and rostering;
- In the short term (few days): for the assignment of each duty to a driver;
- In runtime (minutes): to provide information such as expected arrival time, via SMS or electronic boards.

Trip time prediction may be useful in different ways depending on the type of operational management and control made by the company. There are bus lines over which the control is made by timetable, like the ones used by commuters, for instance; there are others managed and controlled by frequency, as it is the case of the entirely urban ones, at least during peak hours. The study on the factors explaining trip time prediction shall not be undertaken in this paper. Although this is also very important for mass transit companies, namely for negotiation with decision makers on transport policies, it is a different problem, not necessarily solved with the same techniques.

Our case study is STCP (www.stcp.pt), the public bus operator in Porto, Portugal.

## 3. Objectives and methodology

The broader aim of our research project is to design a decision support system for the operational management of mass transit companies in all the areas discussed in section 2. In this paper we focus on trip time prediction for the short term, i.e., the assignment of the duties to the drivers 3 days in advance. The goal is to reduce the cost of overtime pay to drivers by cutting the difference between planned and real trip time.

Considering the on line arrival of every new trip, the input flow can be regarded as a data stream. These data are continuously changing, thus requiring a continued Data Mining (DM) approach. The DM models proposed adapt to new available data, without human intervention. The whole DM project is approached using the CRISP-DM methodology ([7]). This methodology comprises 6 phases: business understanding, data understanding, data preparation, modelling, evaluation and deployment. The methodology is continuously applied, since the business goals may change with time.

## 4. The architecture model

The architecture model we are developing (figure 1) is an ensemble regression framework, i.e., it uses an ensemble of algorithms / parameter sets (a&ps), in order to select, for each
value to predict, an a&ps to train and to predict the trip time. It has three main components: pre-processing, recommendation and prediction components.

Figure 1. The architecture model

The main goal of the pre-processing component is the daily update of the predictions for the next coming days obtained by the different a&ps that have been previously selected. The objective of the pre-selection of $n$ a&ps (discussed in the next section) is to minimize the set of a&ps used by the model without loosing prediction accuracy. The pre-tested algorithms were random forests ([1]), projection pursuit regression ([2]) with three different smoother methods (super smoother, spline and generalized cross-validation spline), and support vector machines ([8]) with three different kernels (linear, radial and sigmoid). The pre-processing component has two tasks: training and prediction. Each one of the a&ps are trained and the new values from the historical database are predicted (see, for example, [3]). These new predictions are stored in the pre-processed database. The pre-processing component runs everyday, typically, during the night.

The recommendation component uses ensemble regression. There are two approaches to implement it: combination or dynamic selection ([6]). The first one combines predictions of several recommended a&ps while dynamic selection chooses just one a&ps. Both approaches use the pre-processed information to recommend the a&ps (one or more). The recommendation is done according to the input value we want to predict. Figure 1 assumes the dynamic selection approach.

The prediction component uses the output of the recommendation component to train the model(s) and uses the trained model(s) to predict the value for the new trip.

5. A study on the pre-selection of $n$ a&ps

The use of several algorithms instead of just one for trip time prediction is motivated by previous experimental work which showed that the best algorithm for each region of the instance space is different from region to region [5]. In the present work we use the regression algorithms referred in the previous section with several parameter sets. 1238 was the total number ($t$) of a&ps tested. The result for each a&ps is the variation index of the trip time calculated as $variation\_index = \sqrt{\frac{\sum(p_i - r_i)^2}{m} / \frac{\sum(r_i)}{m}}$ where $p_i$ and $r_i$ represent,
respectively, the predicted and the real trip times for \( i = 1, \ldots, m \), where \( m \) is the number of observations. The best overall result was 9.92\% and was obtained using random forests.

The hypothesis for our experiments was that an ensemble of regressions could improve predictions. What we present is an heuristic for the pre-selection of an ensemble of \( n \) a\&ps (figure 2). We also present the influence of the \( n \) value on the best possible variation index for the ensembles obtained with our heuristic. Figure 3 presents the average for 10 runs for each different value of \( n \). Obviously, the values for \( n = 1 \) and for \( n = 1238 \) are unique. The heuristic obtains local minimum but the standard deviation is quite small for all values of \( n \).

\[
\text{Input: } M(t \times m), \text{ a matrix with the squared differences between the predictions and the real value of the } m \text{ trips for } t \text{ different a\&ps (in our case } t=1238 \text{ and } m=1796) \\
n, \text{ the number of a\&ps to be selected from } M
\]

\[
\text{Output: } sn, \text{ a set of } n \text{ a\&ps}
\]

1. \( sn = n \) different random values between 1 and \( t \)
2. \( \text{in.set} = \) matrix with the \( sn \) rows of \( M \)
3. \( \text{out.set} = \) matrix with the \( M \) rows

DO

FOR all the rows of \( \text{out.set} \)

1. add the \( \text{in.set} \) row of the \( \text{out.set} \)
2. \( \text{new.eval.value} = \) calculates the sum of the of the \( m \) squared differences selecting for each trip the minimum squared difference from the \( n \) a\&ps from the \( \text{in.set} \)

ENDFOR

in.set = adds to the \( \text{in.set} \) the \( \text{best.in} \), i.e., the a\&ps with the minimum \( \text{new.eval.value} \)

FOR all the rows of \( \text{in.set} \) (now with \( n+1 \) rows)

6. Subtracts the row from the \( \text{in.set} \)

7. \( \text{calculate the } \text{new.eval.value} \text{ as in step 5} \)

ENDFOR

out.set = removes from the \( \text{in.set} \) the \( \text{best.out} \), i.e., the a\&ps with the minimum \( \text{new.eval.value} \)

UNTIL (\( \text{best.in} = \text{best.out} \))

Figure 2. Heuristic for the pre-selection of \( n \) a\&ps

![Figure 2](image1.png)

Figure 3. Mean variation index oracle for \( n \) a\&ps

With this approach we expect to be able to reduce meaningfully the variation index in comparison to the use of just one a\&ps.

320
6. Current status of the framework development

Up to now we have conducted extensive tests with different a&ps. This preliminary work is necessary for a pre-selection of the a&ps and for obtaining the basic statistics needed as input for the heuristic. The selection of the a&ps set size ($n$) will be evaluated after the implementation of the recommendation component ([4]). This component is currently being developed. The software for the updating of the pre-processed database will be implemented soon.

Acknowledgments

This work was partially supported by FCT (POCT/TRA/61001/2004) and FEDER e Programa de Financiamento Plurianual de Unidades de I&D.

References


WILLINGNESS TO PAY FOR ACCESS TIME SAVING: SOME EVIDENCES

Angel IBEAS¹, Luigi DELL’OLIO², Jose Luis MOURA³

Abstract. This paper analyses a logit multinomial discrete choice model using data obtained from Stated Preferences on the willingness to pay for parking places. The case study used an underground car park yet to be constructed. Willingness to pay is shown to be very sensitive to variations in the specifications of the discrete choice model as well as between different types of model because the users only reason for choosing between parking in the street or in the car park is the distance to their final destination and they ignore (it turns out to be irrelevant) another important variable which is the time they are going to be parked.

1. Introduction

In town planning the politics of parking have an important bearing on journey management. In order to study this problem it is important to define the criteria behind why users choose between one type of parking or another. Discrete choice models are an important tool for modelling the users choice between parking in the street or in a pay car park.

These models are often very good at simulating possible user behaviour for projects that are still in the planning stage and, therefore, not yet available.

¹ Universidad de Cantabria, Avda. De los Castros s/n, Codigo 39005, Santander, Spain, ibeasa@unican.es
² Universidad de Burgos, Campus San Amaro Edificio "La Milanera" C/ Villadiego s/n 09001 Burgos, Spain, ldellolio@ubu.es
³ Universidad de Cantabria, Avda. De los Castros s/n, Codigo 39005, Santander, Spain, mourajl@unican.es
The objective of this paper is to calibrate logit multinomial discrete choice models using data from stated preferences to model how user choice would change with the construction of a new underground car park and calculate their willingness to pay to reduce journey time to their final destination. Several relevant studies can be found in the international literature, such as that of Van der Goot [1] analysing different types of parking, and those of Hunt [2] and Polak J.W. & Vythoulkas [3], Ergün [4] and Gillen [5].

2. Context of Choice and Data Collection

A discrete choice model is used to assess willingness to pay in exchange for a shorter journey time to final destination. A "context of choice" must first be defined. This "context of choice" is a dilemma for the user - either park in the street (assuming that street parking is free) or in the soon to be constructed underground car park thereby saving time and/or money.

A survey of Stated Preferences was carried out putting different possible settings to 250 users who parked in the street close to the area where the underground car park would be constructed.

To define the possible settings the area studied was divided into the zones shown in Figure 1. Zone 1 is where the new underground car park would be constructed and zone 2 is where street parking would be banned as the area will be partially pedestrianised.

The survey considered the following two settings depending whether the final destination of the user was in Zone 1 or Zone 2:
Final Destination Located in Zone 1:
“Imagine that you had to make the same journey and had the possibility of parking in an underground car park between 0 and 150 metres from your final destination and that given traffic conditions you couldn't find street parking any closer than 300-350 metres from your destination. Please choose one of the following choices”:

<table>
<thead>
<tr>
<th>If the cost of parking was:</th>
<th>Would you make the journey?</th>
<th>Would you use the car park?</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.8 €/h</td>
<td>Yes - No</td>
<td>Yes - No</td>
</tr>
<tr>
<td>1 €/h</td>
<td>Yes - No</td>
<td>Yes - No</td>
</tr>
<tr>
<td>1.5 €/h</td>
<td>Yes - No</td>
<td>Yes - No</td>
</tr>
</tbody>
</table>

Table 1. Setting zone 1

Final Destination Located in Zone 2:
“Imagine that you had to make the same journey and had the possibility of parking in an underground car park between 150 and 300 metres from your final destination and that given traffic conditions you couldn't find street parking any closer than 150-200 metres from your destination. Please choose one of the following choices”:

<table>
<thead>
<tr>
<th>If the cost of parking was:</th>
<th>Would you make the journey?</th>
<th>Would you use the car park?</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.8 €/h</td>
<td>Yes - No</td>
<td>Yes - No</td>
</tr>
<tr>
<td>1 €/h</td>
<td>Yes - No</td>
<td>Yes - No</td>
</tr>
<tr>
<td>1.5 €/h</td>
<td>Yes - No</td>
<td>Yes - No</td>
</tr>
</tbody>
</table>

Table 2. Setting zone 2

The following variables were considered during data collection and the later calibration of the discrete choice models:

- Sex: Binary variable, 1 for Male, 0 for Female.
- Age: Continuous Variable indicating exact age of user.
- Resident: Binary variable, 1 for resident, 0 if not.
- Vehicle occupation: Number of passengers in the car surveyed
- Age of Vehicle: Age of vehicle in years.
- Income: Monthly income of person surveyed.
- Frequency of journey: Discrete variable, 1 for rare, 2 for monthly, 3 for weekly and 4 for daily.
- Time car is parked
2.1. Elimination of out-layers

In most surveys a high proportion of people, who for whatever reason, don't answer the questions in a consistent way, producing errors when evaluating the models. It is therefore very important to carefully look at the answers in order to weed out the inconsistencies.

In this study 20 answers were rejected because they were considered to be inconsistent with the normal behavior of the user.

Most of them represented choices which were not justifiable and are shown in the graph in Figure 2. Figure 2 is a dispersion graph showing aligned points in two bands.

Axis “x” is the difference \( \Delta d = d_{\mu} - d_{\nu} \), in which \( d_{\mu} \) is the distance from the future underground car park to the final destination of the person surveyed, \( d_{\nu} \) the distance from the street parking location on the day of the survey to the final destination. The “y” axis shows the final choice of the person surveyed using a value of 1 if the choice is to park in the underground car park and 0 if the decision is to park in the street.

The data was ignored if, even though \( \Delta d = d_{\mu} - d_{\nu} > 0 \), the choice was to park in the underground car park, thereby getting a better adjustment for the models used.

3. Results and Conclusions

In order to calculate willingness to pay in exchange for a reduced journey time to final
destination several discrete choice models were calibrated using trial and error. Model 2 was chosen as the best (Table 3). Willingness to pay was calculated using the following equation:

\[ DAP = \frac{\partial V_i}{\partial t_i} = \frac{\partial V_i / \partial c_i}{\partial V_i / \partial t_i} \]

or:

\[ DAP = \frac{\partial V_p / \partial t_{pf}}{\partial V_p / \partial TAR} \]

where:

- \( V_p \) = function of usefulness of parking in the underground car park.
- \( t_{pf} \) = journey time to final destination from underground car park (hours).
- \( TAR \) = Cost in €/h of underground car park.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant [parking in the street]</td>
<td>-1.41330</td>
<td>-2.402350</td>
<td>1.1354288</td>
</tr>
<tr>
<td></td>
<td>-1.725</td>
<td>-2.220</td>
<td>2.871</td>
</tr>
<tr>
<td>Access Time to final destination</td>
<td>-2.897973</td>
<td>20.823492</td>
<td>21.06899</td>
</tr>
<tr>
<td></td>
<td>-2.139</td>
<td>3.172</td>
<td>3.365</td>
</tr>
<tr>
<td>Cost (TAR)</td>
<td>-3.233058</td>
<td>-3.7439401</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-4.129</td>
<td>-4.225</td>
<td></td>
</tr>
<tr>
<td>TAR* Time parked</td>
<td>-0.15694966</td>
<td></td>
<td>-2.337</td>
</tr>
<tr>
<td>[Access Time to final destination]²</td>
<td>-273.691342</td>
<td>-275.364954</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-3.713</td>
<td>-3.000</td>
<td></td>
</tr>
<tr>
<td>Log likelihood function</td>
<td>-104.8002</td>
<td>-96.4243</td>
<td>-104.6008</td>
</tr>
<tr>
<td>No coefficients</td>
<td>-165.6622</td>
<td>-165.6622</td>
<td>-165.6622</td>
</tr>
<tr>
<td>Constants only</td>
<td>-119.8729</td>
<td>-119.8729</td>
<td>-119.8729</td>
</tr>
<tr>
<td>Willingness to pay</td>
<td>2.689</td>
<td>1.899</td>
<td>6.906</td>
</tr>
</tbody>
</table>

Table 3 Calibrated Models

The following conclusions can be drawn from these results:
1. Model 2 is clearly the best.
2. Willingness to pay varies greatly with model specification.
3. Model 2, taking into account the quadratic effects of the access time to the final destination, produces a better adjustment than model 1.
4. Model 3 results in a willingness to pay much greater than both models 1 and 2.

Conclusion 4 can only be explained by the fact that the users are not taking into account the time they are going to be parked when answering the question - they give priority to access time to final destination. Obviously, this means that when choosing whether to use the underground car park or not, they are prepared to be parked irrespectively for how long and ignoring the associated cost (this aspect increases the willingness to pay value). Finally, the inclusion of other variables in the model can generate inconsistencies when determining willingness to pay, and in many cases it is probably better to simplify the specifications of the model and only use variables that we a certain the users are sure of in their answers, thus producing a more prudent model.

References

THE UNCERTAINTY OF DELAYS AT OPTIMAL PRE-TIMED 
AND VEHICLE ACTUATED SIGNALS

Francesco VITI, Henk J.Van ZUYLEN

Abstract. Although extensive research has been carried out in the past to assess the performance of control signals under different service mechanisms, little interest has been given to the way these control methods deal with the variability of the traffic demand. This paper proposes a probabilistic approach to calculate delay and queue length distributions in time, which enables one to compare the effects of simple pre-phased controls - based on delay minimization - and of vehicle actuated controls, which look at the actual headway distribution of vehicles without accounting explicitly for vehicle delays. This study confirms the superiority of gap-based mechanisms over delay-based controls also in terms of delay variability, given their property to adapt green and cycle lengths according to the real amount of traffic arriving during each cycle.

1. Introduction and problem description

Dynamic traffic signal control systems are designed with the scope of increasing network efficiency and safety. On the other hand, signal controls strongly affect the capacity of an intersection and, on a larger scale, to a part of a network. The importance of this performance measure is confirmed by its use in the definition of intersection Level of Service in the Highway Capacity Manual [1].

One of the determinants of delays is the excessive demand. The dependency of delays on the number of arrivals has been analyzed and several models have been developed in the past, mostly based on heuristics (e.g. [2], [3], [4], [5]). The estimation of such delays is, in all models, based on average demand conditions, and it is simply an expectation value. In spite of the road capacity, which has relatively little fluctuations (for example due to adverse weather conditions, lane closures, accidents etc.), the arrivals at one intersection can be very different from one day to another. The uncertainty reflected to the delay that a traveler may experience can be therefore very large. Figure 1 shows travel time observations collected at an arterial road in the city of Delft, the Netherlands. The stretch of road is interrupted by two intersections with vehicle actuated signals. Data was processed

---

1 Delft University of Technology, faculty of Transport and Planning, Stevinweg 1 2628CN Delft, f.viti@ctg.tudelft.nl and h.j.vanzuylen@ct.tudelft.nl
using camera detection. When the demand is low travel times do not increase considerably. During the morning peak several observations resulted in very large and spread travel times. Therefore, during this period a traveler may drive at free flow during one day and experience the double of the expected travel time the day after.

Figure 1: travel times observed at an urban road in Delft, the Netherlands

Recently, transport policies are more and more asking attention to the analysis of the causes and the effects of travel time uncertainty and they call for solutions to reduce it. Very little consideration has been given in the past to the estimation of the queue and delay variability. Newell [6] formulates mathematically this problem using renewal theory. This approach inspired the work of Olszewski [7], who investigated the queue length distribution in time using a Markov Chain process.

In this paper we examine how different traffic control mechanisms cope with the variability of delays. To do so, we formulate the overflow queuing process as a Markov Chain for optimal pre-timed controls (section 2), which are based on delay minimization, and for vehicle actuated controls (section 3), which are gap-based. Comparison between these two control mechanisms is made in section 4 in terms of expectation value and of standard deviation. Finally conclusions are given in section 5.

2. The Markov model for optimal pre-timed controls

The signalized intersection system is governed by a cyclic mechanism, which allows the use of a discrete time analytical process instead of the continuous approach. Van Zuylen [8] described a Markov model for queues at isolated intersections assuming Poisson arrivals and normally distributed saturation flows. Olszewski [7] independently developed the idea of applying the Markov chain technique to signal control problems. The model described in this section is not different from the ones developed by these authors.

The stochastic queuing process is defined in its discrete time case to be a sequence of stochastic variables in time, where a state $Q_{t+1}$ is described by the previous state $Q_t$ and the number of arrivals $a_t$ and departures $d_t$ during the interval $[t, t+1]$, according to the simple relationship $Q_{t+1} = \max\{Q_t + a_t - d_t, 0\}$. Assumed the probability distribution of the arrivals known and deterministic service rate, one can compute the probability of the queue length in time by first computing the transition probability from one cycle to another.
\[ q_j(t) = \Pr(i = j + a_i - d_i) \quad \forall \quad j \geq i - d_i, \quad a_i \in [0, a_{\text{max}}] \]  

(1)

which represents the probability that the queue length moves from a state \( j \) at time \( t - 1 \) to state \( i \) at time \( t \). The probability of a zero overflow queue comprises all cases where \( j + a_i - d_i \leq 0 \), while if a maximum value for the queue \( Q_{\text{max}} \) is assumed (which can represent the maximum number of vehicles that can buffer at the road section without creating spillback effects), this value will comprise all values for which \( j + a_i - d_i \geq Q_{\text{max}} \).

The queue length probability at time \( t \) is therefore given by the following formula (2):

\[ \Pr(j, t) = \sum_{i=0}^{Q_{\text{max}}} \Pr(i) \cdot q_j(t) \]  

(2)

Figure 2: The distribution of overflow queue lengths in time

Figure 2 shows an example of overflow queue length distribution in time for \( x=0.95 \) and with zero initial queue. Even if the probability of having a zero overflow queue remains the largest chance, it reduces to only 50%.

Figure 3: individual delay and optimal cycle length

The computation of delays is done by using the probabilistic model proposed by Olszewski [9] and it is left out in this paper.
The presented model enables one to correctly evaluate the dynamic and the stochastic character of overflow queues at any loading condition. A direct consequence can be seen in the computation of the optimal cycle length for each demand condition. For example, the optimal cycle computed by Webster [2] is derived by assuming the overflow to be in equilibrium. This state can be reached in conditions of demand very close to capacity only after several cycles (in the example above after around 50 cycles). This means that the Webster delay formula overestimates the individual delay. As it can be seen in figure 3, this error reflects in the computation of the optimal cycle length, which is with the Markov model smaller than the one computed with the Webster’s optimal cycle formula. This error increases the closer the demand is to the capacity.

3. Markov model for vehicle actuated controls

Actuated control phase plans are in general determined by the headway distribution of the arrivals at the intersection. The basic mechanism is to extend the guaranteed green time (say, a minimum value) until the arrival time of two consecutive vehicles is larger than a certain threshold. This green time is usually constrained to be smaller than a maximum value. Therefore, the assigned green times and the delay incurred are stochastic variables too. The computation of green times is then subdivided into three parts: the green time given to serve the vehicles queuing up during the red phase, the one given to the vehicles queuing while the green phase is started and the green time extension given to vehicles arriving in sequence with short headways after the queue has been cleared.

If $a_r(t)$ is the arrival rate (in vehicles per second), and $r_i(t)$ is the red time at the previous cycle for stream $i$, one can compute the probability of a certain number of vehicles $k$ queuing up during the red phase (of length $\rho$) as:

$$P(Q^r(t) = k) = \int_{\rho/\tau}^{\infty} \left( P(a_r(t) \cdot \rho = k) \cdot P(r_i(t) = \rho) \right) d\rho$$

(3)

The probability of a green time $g^i(t)$ needed to clear the queue at the end of the red phase to be a value $l$ is therefore given by:

$$P(g^i(t) = l) = \sum_{k=1}^{\infty} P(Q^r(t) = k)$$

(4)

While clearing the queue formed during the red phase, other vehicles may join the queue. These vehicles are computed by replacing the probability of red time in formula (3) with the green time of formula (4). The probability of green time due to all vehicles in queue $g^0$ is thus given by computing the joint probability of green due to vehicles arriving during the red phase and the ones arriving during the green phase (see also [10]).

The probability of green time extension is computed by computing all sequences of vehicles with headway shorter than the unit extension $T$. If one computes the probability distribution of a sequence of $n$ vehicles at times $0 < t_1 < t_2 < ... < t_n = t$, $P(n,t)$, the probability of observing this sequence with $t_2 - t_1 < T$, $t_3 - t_2 < T$, etc. is given by the following formula:

$$P(t_{\text{ext}} = t) = \sum_{n=0}^{\infty} P(t_1 < t_2 < ... < t_n = t) \cdot P(n,t) \quad \text{s.t.} \quad t_1 < T, t_2 - t_1 < T, ..., t_n - t_{n-1} < T$$

(5)
The probability of having an extension of exactly \( t \) seconds is then given by:

\[
P_t(g^e_i(t) = t) = \sum P(t_{ext} = t) \cdot P(g_{\text{max}} - g_i^e \geq t)
\]  

(6)

The probability of a total green time \( g^t \) is finally given by computing the joint probability of green given to clear the queue and the green time extension.

Overflow queues are likely to occur only when the intersection is oversaturated and the maximum green extension is met. The corresponding probability is computed by the following formula:

\[
P(Q_q(t) = q) = \sum_{k-g_{\text{max}} \leq q} P(Q(t) = k)
\]  

(7)

Since an eventual overflow queue should be cleared in the next green phase, formula (3) should also consider that, apart from the arrivals, also the eventual overflow queue should be served (see [10] for details). Last step is to derive the probability distribution of the red times at the previous cycle. The corresponding probability of a red time to be a certain value \( r(t) = s \) is thus computed with the following formula:

\[
P(r(t) = \rho) = P(\sum_{j=1}^{\rho} g_{\text{ext}}(t-1) + TL = \rho)
\]  

(8)

### Figure 4: expected green time and overflow queue for different demand conditions

#### 4. Comparison of delay estimates

The application of the Markov model to these two control mechanisms enables one to analyze their efficiency also in terms of variability, according to the scope of this paper. Table 1 presents the results of the Markov models in terms of expectation value and standard deviation of the individual delay for different demand conditions. The corresponding degrees of saturations range from 0.7 to 0.99 with pre-phased controls.

<table>
<thead>
<tr>
<th>flow</th>
<th>500</th>
<th>535</th>
<th>570</th>
<th>605</th>
<th>640</th>
<th>675</th>
<th>710</th>
</tr>
</thead>
<tbody>
<tr>
<td>E[W] optimal</td>
<td>14.1</td>
<td>15.2</td>
<td>17.1</td>
<td>19.5</td>
<td>22.6</td>
<td>26.6</td>
<td>32.1</td>
</tr>
<tr>
<td>(\sigma[W]) optimal</td>
<td>4.7</td>
<td>6.5</td>
<td>8.1</td>
<td>9.9</td>
<td>12.1</td>
<td>13.9</td>
<td>16.2</td>
</tr>
<tr>
<td>E[W] actuated</td>
<td>19.5</td>
<td>20.9</td>
<td>22.4</td>
<td>24.2</td>
<td>26.1</td>
<td>28.3</td>
<td>30.6</td>
</tr>
<tr>
<td>(\sigma[W]) actuated</td>
<td>8.7</td>
<td>9.1</td>
<td>9.3</td>
<td>9.7</td>
<td>10.1</td>
<td>10.5</td>
<td>10.9</td>
</tr>
</tbody>
</table>

Table 1: expected delay and standard deviation for \( s=1800 \) veh/s
It should be said that if a larger demand is loaded the optimal cycle method will result in too large cycles and the delay increases more steeply, since overflow queues will appear systematically. The vehicle actuated control can still give low delays and standard deviations for a much larger demand (as one can see from figure 4). This is easy to understand: low demands for pre-timed controls result in an unused part of the green time, while this does not occur at vehicle actuated controls because green times are variable too.

5. Conclusions

This paper analyzed the behavior of optimal pre-phased controls and vehicle actuated controls in terms of expected delay and standard deviation. To do so, the individual delay at these traffic controls is modeled using Markov renewal theory, which enables one to compute a full probability distribution of the overflow queue and delay in time.

Conclusions from the study point at the superiority of vehicle actuated controls, during conditions of demand near capacity, for two reasons: 1) it outperforms the optimal pre-phased control in terms of both expectation value and standard deviation and 2) it enables one to serve a higher total demand at the intersection. This conclusion can be helpful information for practitioners and policy makers; for an intersection, which does not have to serve a large demand, the simple fixed control can suffice, while it is recommended to use a responsive control like the vehicle actuated control if this demand is close to the capacity.

References


TWO LANE HIGHWAYS - A MICROSCOPIC TRAFFIC SIMULATION MODEL

Caroline C. PECKER, Helena B. B. CYBIS

Abstract. Simulation models are very important as supporting tools for traffic analysis. Due to its complexity, there are not as many traffic models to represent two-lane highways. This work aims to describe a microscopic simulation model for two-lane rural roads designed to reproduce scenarios characterized by high vehicle flow in both directions at same period, high percentage of overloaded vehicles on upgrades and aggressive driver behavior.

1. Introduction

Microscopic simulation models are useful tools for road analysis studies. However, given its complexity, just a small number of models are able to represent the interaction between vehicles in a two-lane highway. Commercial simulation models that allow simulation of two-lane highway are TRARR [1], TWOPAS [2] and PARAMICS [3]. The last version of TRARR was released on 1991 and technical support is no longer provided. TWOPAS was originally developed during the 70’s and has been updated constantly. It was used to develop the two-lane highway capacity analysis of HCM2000 and is included as the traffic analysis module at the Interactive Highway Safety Design Model (IHSDM) [4]. PARAMICS is a generic model developed to simulate urban and interurban areas and is able to represent two-lane highways.

In general terms, these models are adequate to represent normal traffic conditions. However, high vehicle flow in opposite directions at same period, high percentage of overloaded vehicles riding uphill and aggressive driver behavior produce extreme traffic conditions that deserve special consideration. This paper presents a microscopic traffic simulation model for two-lane highways intended to represent traffic behavior under these extreme conditions.

The model overview is described in the following section. Section 3 present the results from a case study and the last section present the concluding remarks.

1 LASTRAN, School of Engineering – UFRGS, Praça Argentina n. 9 sala 402 – Porto Alegre, Brazil, (carol, helenabc)@producao.ufrgs.br
2. The model overview

The model was developed to represent traffic flow on different road and traffic conditions. Regarding highway geometry, the model takes into account the effect of grades, horizontal curves, passing and climbing lanes. Traffic control features include the definition of passing and no-passing zones as well as zones with different speed limits. The vehicle characteristics are length, engine power, weight and others physical and mechanical specifications. Besides the vehicle information, it is possible to define inherent driver characteristics, such as desired speed, maximum desired acceleration and car-following and lane-changing/overtaking decisions parameters.

The simulation software was developed in C++. A graphical interface allows a visual representation of the road during the simulation process (see figure 1). It is also possible to select probe vehicles and specify road sections for data collection purposes.

2.1. Car-following logic

The car-following model adopted was proposed by Gipps [5]. The Gipps logic has already been implemented in AIMSUN [6], developed at the Universitat Politecnica de Catalunya, Spain. This car-following model is ruled by two functions: acceleration and deceleration. The first function \(u_{ac}(t+T)\) represents the vehicle intention to reach a desired speed constrained by its acceleration restriction (see equation 1). The second function \(u_{dc}(t+T)\) reproduces the constraints inflicted by the preceding vehicle when trying to drive at the desired speed (see equation 2).
\[
\begin{align*}
  u_n^{acc}(t + T) &= u_n(t) + 2.5aT \left(1 - \frac{u_n(t)}{U_n}\right) \sqrt{0.025 + \frac{u_n(t)}{U_n}} \quad (1) \\
  u_n^{des}(t + T) &= \sqrt{b^2T^2 - b[2(x_n(t) - L_n) - x_n(t)] - u_n(t)T - \frac{u_n(t)}{b'}} \quad (2)
\end{align*}
\]

where \(u_n(t)\) is the speed of vehicle \(n\) at time \(t\), \(u_{n-1}(t)\) is the speed of vehicle \(n-1\), \(T\) is the reaction time, \(U_n\) is the desired speed of vehicle \(n\), \(a\) is the maximum acceleration for vehicle \(n\), \(b\) is the maximum deceleration rate desired by vehicle \(n\) \((b < 0)\), \(L_n\) is the effective length of vehicle \(n\), \(x_n(t)\) is the position of vehicle \(n\) at time \(t\), \(b'\) is an estimation of the desired deceleration of vehicle \(n-1\). The relation between \(b\) and \(b'\) influence the lane capacity. The distribution of desired speeds affects the position and shape of the upper arm of the speed-flow curve.

According to Gipps, the speed of the follower vehicle at time \(t+T\) is:

\[
u_n(t+T) = \min(u_n^{acc}(t+T), u_n^{des}(t+T)) \quad (3)
\]

In order to provide a better representation of the acceleration behavior in the modeling environment, equation 1 was replaced by a vehicle dynamics model. It takes into account the road geometry and physical and mechanical features of vehicles to calculate the maximum acceleration the driver can impose his vehicle. The acceleration is based on the resultant force between traction and resistance acting on the vehicle. The vehicle dynamics model is described in [7].

A realistic fleet scenario description is essential to obtain sound results from the vehicle dynamics model. The research presented in [8] describes this model validation for Brazilian truck engines. Most of the engines are old and low power. The overloaded vehicles are 20% to 30% of the fleet [9] and the load excess is 25% per axle. These fleet characteristics and aggressive driver behavior lead to singular performance in two-lane roads.

### 2.2. Lane-changing/overtaking maneuvers

The lane-changing/overtaking maneuvers were defined by two different events: ‘overtaking using opposite lane’ and ‘changing to auxiliary lane’.

The process of ‘overtaking using opposite lane’ is based on three variables: (i) the time gap the overtaking vehicle takes to reach the leader on the opposite lane, (ii) the headway between the two vehicles preceding the overtaking vehicle, and (iii) the time to accomplish the overtaking maneuver. During the overtaking process, all other involved vehicles keep constant speeds and do not overtake or change lanes.

Vehicles traveling through a road section with passing/climbing lanes behave according to the rules from the ‘changing to auxiliary lane’ process. The road section is segmented in two zones and vehicles’ behavior depends on which zone they are.

In the first zone, the ‘lane-changing zone’ any vehicle slower than its follower is supposed to move to the auxiliary lane as long as it finds an adequate gap. Vehicles in the auxiliary lane, delayed by their leaders, may move back to the main lane, provided they comply with a set of rules ensuring they will not cause major disruption to traffic flow.

The second zone is the ‘emergency zone’ (see Figure 2). It is the area closer to the end of the auxiliary lane. Vehicles on the main lane are not allowed to move to the auxiliary
lane. They adjust their speed according to vehicles on the auxiliary lane in order to allow these vehicles to return to the main lane.

Figure 2. Passing lane scheme and emergency zone.

These rules have been devised to represent environments with large number of heavy and slow vehicles and aggressive drivers that force their way back to the main lane. Parameters to adjust this model include the extension of the emergency zone and car following deceleration parameters.

3. An application example

A case study modeling a two-lane road segment with and without climbing lane allowed the comparison of the two roads performances. The road segment located on a mountainous area had a speed limit set to 80 km/h. The analysis was performed on an uphill section, defined as a no-passing zone. The road had a uniform 5% grade and extension of 1,200 meters.

Figure 3 illustrates the average speed fluctuation according to the flow rate. Figure 3a, presents the simulation results for the performance of the road without climbing lane and traffic flow composed by 25% of heavy vehicles. Figures 3b and 3c presents the simulation results for the road segment with a climbing lane and traffic flow composed by 25% and 40% of heavy vehicle respectively.

Figure 3a allows identifying the impact of trucks on the average speed of the uphill section. All vehicles average speed is reduced to the trucks average speed. Cars performance is constrained by heavy vehicles. According to Figures 3b and 3c, it is possible to identify that the construction of a climbing lane would greatly affect car’s average speed.
Figure 3: Speed-flow curves.
4. Concluding remarks

This model structure was devised to provide a flexible modeling structure to allow the calibration of different scenarios. The vehicle dynamics model allows the representation of an assorted fleet of vehicles with different physical and mechanical characteristics.

Modeling the driver behavior during the passing maneuvers, essentially on segments with auxiliary lanes, was devised to adequately reproduce the performance on Brazilian roads. This tool can be very useful considering the limited number of tools intended to represent two-lane highways, especially for developing countries scenarios.

References


TIME WINDOWS FOR SCHEDULED TRIPS IN MULTIPLE DEPOT VEHICLE SCHEDULING

Natalia KLIEWER*, Stefan BUNTE, Leena SUHL†

Abstract. In this paper we consider time windows for scheduled trips in multi-depot vehicle scheduling problem arising in public bus transportation. To solve this problem by branch-and-cut using standard optimization libraries we formulate a network flow problem based on the extended time-space-network. Considering time windows increases the problem complexity and makes large practical instances hard to solve. To avoid this, trips that are critical for schedule operations should be determined and only these trips should be provided with time windows. We propose two different techniques to identify critical trips. Shifting of these trips is likely to lead to savings in the number of vehicles and/or operational costs.

Keywords: Multiple Depot Vehicle Scheduling, Time Windows, Time-Space-Network, Public Transportation

1. Introduction

We consider the multiple-depot multiple-vehicle-type vehicle scheduling problem (MDVSP in the following), which arises in public transportation. This problem consists in assigning buses to cover a given set of timetabled trips with consideration of multiple depots for vehicles and vehicle type groups. Furthermore it considers depot capacities and constraints on number of vehicles of certain vehicle type. Each timetabled trip can be served by a vehicle (bus) belonging to a given set of vehicle types - vehicle type group. Each vehicle has to start and end its work day in one of the given depots. After serving one timetabled trip, also called service trip, each bus can serve one of the trips starting later from the station where the vehicle is standing, or it can change its location by moving unloaded to another station, through an empty movement, or deadhead, in order to serve the next loaded trip starting there. In addition, the maximal number of parking slots for each depot as well as maximal fleet size for each vehicle type is given.

An optimal schedule is characterized by minimal total schedule operation costs. The costs include fixed costs for each used vehicle as well as variable operational costs consisting

*corresponding author
†DS & OR Lab, University of Paderborn, Warburger Str. 100, D-33098 Paderborn, Germany. {kliewer,bunte,suhl}@dsor.de
of distance-dependent travel costs and time-dependent costs for time spent outside the depot. All cost components depend on the vehicle type. Since the fixed cost components are usually orders of magnitude higher than the operational costs, the optimal solution always involves a minimal number of vehicles.

The greatest difficulty by solving MDSVP is about handling an immense amount of possible connections between compatible trips (Two trips, which can be served subsequently by one vehicle are compatible). In [7] and [6] a new modeling approach for MDSVP has been presented, which avoids the drawback of explicit consideration of all possible connections as used in other existing network flow ([8], [3]), quasi-assignment ([2]) or set partitioning ([9]) models for MDSVP. This approach is based on a multi-layered time-space-network (TSN in the following) in which possible connections between groups of compatible trips are aggregated. Thus the number of compatibility arcs in the network decreases drastically (stable by 97 to 99 %) compared to traditional approaches without losing any feasible vehicle schedule. This kind of modeling makes it possible to solve the corresponding multi-commodity flow model using the Branch&Cut method of a standard MIP-solver (like [4] and [10]) in appropriate time even for large real instances.

To close the gap between theoretical approaches for MDSVP and real planning problems arising in public transport companies, further practical requirements should be considered. For example customers can ask for minimization of line changes for each bus over the working day or for the returning of a bus to its depot over the day if worthwhile. The TSN-based model allows for the integration of such practical aspects (i.e. see [5]).

The consideration of time windows for scheduled trips in MDSVP (MDSVP-TW in the following) is one highly relevant aspect in resource scheduling of bus companies. Time windows of several minutes should be considered for some or all service trips, especially for school bus trips. Modeling of time windows with continuous variables for this kind of problems leads to models that are difficult to solve. But since the departure and arrival times in a timetable are just accurate to some period of time (as a rule one minute - particularly in Germany), it is possible to consider discrete intervals for trip departure times, for instance 09.59 or 10.01 for time window of -1 to +1 minutes relating to initial scheduled departure time of 10.00. Thus we assume discrete values for possible start and arrival times for trips on MDSVP-TW as e.g in [1], which uses discrete time windows in a heuristic approach for VSP.

This paper presents time windows consideration in TSN-based modeling of MDSVP. It is organized as follows: Section 2. describes the extension of the TSN-model with time
2. Consideration of time windows for trips

In the TSN-model, the potential vehicle activities are represented as arcs between time-space-points (possible arrival or departure events). All time-space-points in a station or in a depot are organized as timelines - possible events ordered according to time and connected with waiting arcs. Other arc types represent serving a service trip, deadhead or moving from or to the depot. To consider time windows for service trips, we introduce a new kind of arc, the so-called time window-arcs in the following way:

**Multiplication of arcs:** The time window-arcs represent multiplications of original service trip-arcs. For service trips with given time windows we insert time window-arcs, each of them representing a trip displacement of a certain amount of time. Figure 2 shows an example with three service trips and a time window for trip b of ±2 minutes. A latening of trip b for one or two minutes would make trips a and b compatible. Each new arc requires an additional flow variable in the mathematical model. No additional constraints are needed; the existing cover constraints only have to be enhanced by variables corresponding to new arcs. They assume that the sum of all arc flows belonging to a special trip over different network layers is equal to one, thus covering each service trip exactly once.

**Model size reduction:** Inserting additional arcs leads to a growing mathematical model, for which solution times can get very high. Therefore we should avoid insertion of a time window-arc if it don’t enable new trip compatibilities compared to the original service trip-arc without shifting. We developed techniques to filter out such useless arcs.

Figure 3 contains an example for identifying useless time window-arcs. A service trip a from station 1 to station 2 is provided with a time window of ±2 minutes. In addition, there are two arrivals in station 1 and two departures in station 2 (s. picture (1)). Picture (2) shows inserted time window-arcs. Now, all "earlier" time window-arcs should be checked for achieving a new connection in their arrival station (here - station 2). Analogous, "later" time window-arcs should be checked for providing a new connection in their departure station (here - station 1). Picture (3) shows that in this case two of four time window-arcs could be useful. In similar way we proceed with additional arcs for depot trips.
Penalty for trip displacement: In providing trips with time windows bus companies want to save resources, but they also want to keep the given timetable as untouched as possible. To make sure that a displacement only takes place if a saving is offered, we extend the model with penalty costs for service trip displacement in the following way: Before optimizing the model we slightly raise costs for time window-arcs, so that the service-arc without displacement has the lowest cost and the connected time window-arc costs are increased the further they are from the original service trip. Formally, the cost $c_a$ for a time window-arc $a$ is $c_a = c_s + |t_a| \cdot \epsilon$ with $t_a$ as displacement of $a$, $c_s$ as associated service-arc and $\epsilon$ as a low cost value which is dominated by all other cost coefficients. By varying the cost ratio of $\epsilon$ concerning the other model costs, we control the trade-off between cost saving and overall displacement. In our tests we use a setting ensuring that displacement only takes place if a vehicle could be saved.

3. Determining the set of displaceable trips

Our first, straightforward approach is to let all trips to be displaceable in a certain given range - global time window. This leads to the abundance of variables in the model. Furthermore, the user is not able to control the displacement for individual trips. Anyway, the overall displacement is theoretically interesting as it offers an upper bound and gives us the chance to compare and evaluate identification methods for critical trips. In real-life planning, it is practical to approve only a subset of trips for displacement. For example, it could be desirable to let all standard timetabled trips unapproved and to involve only special trips like school buses or amplification trips in rush hours. That way, the user gets full control of displacement and the model size stays small.

We encountered the problem that the user is interested in getting the most saving, but has not the knowledge which trips he has to approve for displacement. Therefore we have developed two heuristic approaches - a trip-shortening heuristic, and a cutting-heuristic. The intention of these heuristics is to identify a so-called critical set of trips. The goal thereby is to get a preferably small set that achieves high savings comparable to the global time windows and is faster to solve.

Trip-shortening heuristic: This heuristic uses a kind of what-if analysis with shortened trips. For this purpose we first generate a cloned timetable with shorten trip times and solve the MDVSP with shortened trips. Then we find out the set of "bottleneck"-trips
whose shortening led to new connections that would be infeasible without trip shortening. Afterwards we solve the MDVSP-TW with time windows for trips recovered in this way.

**Cutting-heuristic:** Unlike the trip-shortening heuristic cutting-heuristic needs no preliminary optimization process. Instead, the critical set of trips is extracted from the information of the timetable. The cutting-heuristic assumes the existence of load peaks in the timetable. As a rule, there are characteristic peaks in the morning and in the afternoon, when school buses are used and job rush-hour takes place. Thus we determine the set of trips that have to be served in the peak times as a critical trip set and provide only these trips with time windows. The size of the set can be controlled by user-defined parameters.

### 4. Computational Results

Tables 1 and 2 present computational results on three instances of the cities Saarbruecken, Halle and Munich provided by PTV AG. The complexity of an instance depends on the number of scheduled service trips and furthermore on the average group size ($\mathcal{G}$), which means the average number of legal vehicle type-depot-combinations for the timetable trips. The width of time windows vary from 1 to 6 minutes, the results for cutting heuristic are given for two different parameter settings. Missing entries means optimizations runs dropped put after exceeding the time limit of 24 hours. All run times are given for ILOG CPLEX 9.0 using dual simplex algorithm to solve the LP-relaxations under Microsoft Windows XP. The cutting heuristic outperforms the trip shortening heuristic on this benchmark. In nearly all experiments the cutting heuristic obtains the same amount of savings as the global time windows approach in only a fraction of the global approach runtime.

<table>
<thead>
<tr>
<th>Instance</th>
<th>Time window:</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Saarbruecken</td>
<td>– Global</td>
<td>14</td>
<td>39</td>
<td>106</td>
<td>310</td>
<td>1648</td>
<td>2563</td>
</tr>
<tr>
<td></td>
<td>– Shortening</td>
<td>174</td>
<td>175</td>
<td>172</td>
<td>184</td>
<td>191</td>
<td>195</td>
</tr>
<tr>
<td>$\mathcal{G} = 2$</td>
<td>– Cutting(10/10)</td>
<td>10</td>
<td>30</td>
<td>70</td>
<td>246</td>
<td>1302</td>
<td>2602</td>
</tr>
<tr>
<td></td>
<td>– Cutting(50%)</td>
<td>8</td>
<td>28</td>
<td>28</td>
<td>277</td>
<td>327</td>
<td>475</td>
</tr>
<tr>
<td>Halle</td>
<td>– Global</td>
<td>1600</td>
<td>3600</td>
<td>7847</td>
<td>11688</td>
<td>17109</td>
<td>21075</td>
</tr>
<tr>
<td></td>
<td>– Shortening</td>
<td>190</td>
<td>703</td>
<td>621</td>
<td>791</td>
<td>1751</td>
<td>2620</td>
</tr>
<tr>
<td>$\mathcal{G} = 3.2$</td>
<td>– Cutting(10/10)</td>
<td>268</td>
<td>483</td>
<td>644</td>
<td>1873</td>
<td>1522</td>
<td>2145</td>
</tr>
<tr>
<td></td>
<td>– Cutting(50%)</td>
<td>313</td>
<td>592</td>
<td>1132</td>
<td>2421</td>
<td>3115</td>
<td>3883</td>
</tr>
<tr>
<td>Munich</td>
<td>– Shortening</td>
<td>50114</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>$\mathcal{G} = 13$</td>
<td>– Cutting(50%)</td>
<td>80908</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

Table 1. Runtime (CPU seconds on XEON 2.20 GHz with 2 GB RAM)
<table>
<thead>
<tr>
<th>Time window (minutes):</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Saarbruecken</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>– Global</td>
<td>0(0)</td>
<td>2.0(16)</td>
<td>4.1(81)</td>
<td>8.2(300)</td>
<td>12.2(770)</td>
<td>14.3(1087)</td>
</tr>
<tr>
<td>– Shortening</td>
<td>0(0)</td>
<td>0(0)</td>
<td>0(0)</td>
<td>2.0(22)</td>
<td>2.0(39)</td>
<td>4.1(145)</td>
</tr>
<tr>
<td>– Cutting(10/10)</td>
<td>0(0)</td>
<td>2.0(14)</td>
<td>4.1(76)</td>
<td>8.2(291)</td>
<td>12.2(749)</td>
<td>14.3(1062)</td>
</tr>
<tr>
<td>– Cutting(50%)</td>
<td>0(0)</td>
<td>2.0(14)</td>
<td>2.0(16)</td>
<td>8.2(299)</td>
<td>8.2(279)</td>
<td>8.2(310)</td>
</tr>
<tr>
<td><strong>Halle</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>– Global</td>
<td>4.4(50)</td>
<td>7.0(127)</td>
<td>7.8(225)</td>
<td>8.7(291)</td>
<td>9.6(372)</td>
<td>11.3(569)</td>
</tr>
<tr>
<td>– Shortening</td>
<td>2.6(18)</td>
<td>6.1(372)</td>
<td>7.0(279)</td>
<td>7.8(376)</td>
<td>7.8(413)</td>
<td>9.6(1012)</td>
</tr>
<tr>
<td>– Cutting(10/10)</td>
<td>3.5(26)</td>
<td>7.0(125)</td>
<td>7.8(231)</td>
<td>8.7(283)</td>
<td>9.6(344)</td>
<td>9.6(378)</td>
</tr>
<tr>
<td>– Cutting(50%)</td>
<td>4.4(50)</td>
<td>7.0(127)</td>
<td>7.8(226)</td>
<td>8.7(300)</td>
<td>9.6(362)</td>
<td>11.3(598)</td>
</tr>
<tr>
<td><strong>Munich</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>– Shortening</td>
<td>0(0)</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>– Cutting(50%)</td>
<td>1.8(7)</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

Table 2. Saving of number of vehicles in % (total displacement in minutes)

5. Conclusions

The proposed model allows easy to consider different special cases - time windows neither have to be symmetric, nor have to be connected. Therefore, additional restrictions, like only allowing displacement if a minimum displacement range is fulfilled, are easy to extend without adding new restrictions to the mathematical model.

After extensions have been made to the network and the mathematical model, it turns out that real-life instances are still very large and thus still take a long time to solve. In order to reduce the model size even further, we proposed to allow changes not for each trip, but to reduce the size of the set of changeable trips instead. Two heuristics are proposed for selecting critical trips that are likely to improve the solution when changed.

The tests on real problem instances show savings in the number of used vehicles even with short time windows for service trips. Global time windows for all trips of a timetable provide the largest savings, but require a long solution time. Comparable results can be achieved in a much shorter time by using one of proposed heuristic procedures.

References


AN ALGORITHM FOR THE PICK-UP AND DELIVERY PROBLEM WITH REAR-LOADING CONSTRAINTS

Giovanni RIGHINI*, Federico FICARELLI†

Abstract. We study the variation of the pick-up and delivery problem with given origin/destination pairs in which loading and unloading operations must be sequenced according to a last-in-first-out policy. This constraint models the situation of a rear-loaded vehicle, in which the packages picked-up must be delivered in reverse order. We present an exact optimization algorithm, based on dynamic programming, where states are fathomed by dominance tests as well as lower bounds and we report on computational results.

1. Introduction

In this paper we consider the pick-up and delivery problem with rear-loading constraints (PDPRL), a variation of the well-known pick-up and delivery problem (PDP) that asks for a minimum length tour starting from a depot, visiting a given set of customers and going back to the depot. Each customer corresponds to an origin-destination pair and each origin must be visited before the corresponding destination. Since the pairings between origins and destinations are known in advance, the problem is also sometimes referred to as the single-vehicle dial-a-ride problem (1-DARP). A review on pick-up and delivery problems can be found in a paper by Savelsbergh and Sol [8]. Exact optimization algorithms for the PDP have been presented by Kalantari et al. [4] and by Ruland and Rodin [7]. Heuristics have been presented by Healy and Moll [3], Renaud et al. [6] and Renaud et al. [5].

The rear-loading constraint allows the vehicle to deliver only the last item picked up; hence the vehicle is loaded and unloaded according to a last-in-first-out policy. Heuristic algorithms for the PDPRL have been presented by Cassani [2] and Carrabs et al. [1]. No exact algorithms for the PDPRL have been published so far at the best of our knowledge.

*DTI - Università degli Studi di Milano, via Bramante 65, 26013 Crema, Italy. righini@dti.unimi.it
†DTI - Università degli Studi di Milano, via Bramante 65, 26013 Crema, Italy. fficarelli@crema.unimi.it
We present here an exact optimization algorithm based on dynamic programming and we report on computational results. The effectiveness of the algorithm relies on dominance tests made efficient by suitable data-structures and on bounding tests based on a simple combinatorial lower bound.

2. The problem

The PDPRL is formally defined as follows. We are given a set $N$ of cardinality $N$ of *customers*; each customer is represented by an origin/destination pair: a vehicle must visit each customer, picking up a unit load at the origin and delivering it at the corresponding destination. The problem is represented by a digraph $G(V, A)$, whose vertex set $V$ is made by the depot, numbered 0, and by the $N$ origin/destination pairs, such that $|V| = 2N + 1$. Each arc $(u, v) \in A$ has an associated cost $c(u, v)$ and the overall cost of the tour must be minimized. We assume the triangle inequality holds.

3. A dynamic programming algorithm

We developed a path-growing dynamic programming algorithm, in which paths starting from the depot are iteratively augmented by appending vertices to them. The extensions considered are only those complying with the rear-loading constraints.

A *state* corresponds to a feasible path from the depot to a vertex and is represented by a quadruple $S = (s, D, u, C)$, where:

- $s$ is the ordered set of origins whose corresponding destinations have not yet been visited; it represents the state of the stack correspondent to the load on board of the vehicle;
- $D$ is the set of the destinations already visited;
- $u$ is the last vertex reached;
- $C$ is the cost incurred.

State *transitions* correspond to the operations of appending an additional vertex to a path. The extension rules are the following.

- Visit to an origin: let $o$ be the origin visited. The new state is $S' = (s', D', u', C')$ such that $s' = \text{push}(s, o)$, $D' = D$, $C' = C + c(u, o)$ and $u' = o$. By \text{push}(s, o) we indicate the result of inserting vertex $o$ on top of the stack $s$.

- Visit to a destination: let $d$ be the destination visited. The new state is $S' = (s', D', u', C')$ such that $s' = \text{pop}(s)$, $D' = D \cup \{d\}$, $C' = C + c(u, d)$ and $u' = d$. By \text{pop}(s) we indicate the result of deleting the top element from the stack $s$. We
remark that the only destination the vehicle can visit is the one corresponding with the origin on top of the stack.

All states are organized in a data-structure made by a \((N + 2) \times (N + 2)\) matrix \(M\) whose row and column indices represent respectively the number of origins and the number of destinations visited. All states stored in cell \(M(\text{row}, \text{col})\) correspond to paths visiting \(\text{row}\) origins and \(\text{col}\) destinations. Since the number of destinations visited can never exceed the number of origins visited, only half of the matrix is actually used. Every transition from a state in cell \(M(\text{row}, \text{col})\) generates new states to be stored either in cell \(M(\text{row} + 1, \text{col})\) or in cell \(M(\text{row}, \text{col} + 1)\). The initial state \(S_0 = (\emptyset, \emptyset, 0, 0)\) is stored in cell \(M(0,0)\) and the final content of cell \(M(N+1,N+1)\) is the optimal solution.

Dominance criteria are used to delete dominated states. A state \(S' = (s', D', u', C')\) dominates a state \(S'' = (s'', D'', u'', C'')\) only if \(D'' \subseteq D', s'' = s', u'' = u', C'' \geq C'\) and at least one of the inequalities is strict. To make the dominance test fast, when a new state \(S'\) is generated in cell \(M(\text{row}, \text{col})\), we restrict the comparison to the states in the same diagonal of \(M\), since they are the only states for which condition \(s' = s''\) can be true. All states in cells with lower row and column indices are candidate to be dominated by \(S'\), whereas all states in cells with higher row and column indices are candidate to dominate \(S'\). For each cell of matrix \(M\) the states are partitioned according to the last vertex reached. Therefore in constant time a subset of states is selected for the dominance test, according to vertex \(u'\). Finally a hash function allows to quickly discard states with a stack different from \(s'\). Hence the search is done on a list of remaining states and for each of them the two inequalities on \(C\) and \(D\) are checked.

Another method to reduce the number of states is bounding. A lower bound is associated to each path and if it is found to be greater than or equal to the incumbent upper bound, the corresponding path is discarded. The lower bound is given by the cost of returning to the depot visiting the destinations corresponding to the origins in the stack \(s\) in the right order. Let us indicate such path by \(\hat{P}\). Owing to the triangle inequality this cost must certainly be incurred in any case. Moreover we add to the cost of path \(\hat{P}\) the maximum insertion cost among all the vertices of the customers not yet visited. The insertion cost of a vertex is defined as the minimum among the insertion costs correspondent to all possible positions along the path \(\hat{P}\).

The search policy can be chosen among different alternatives. In our experiments we found convenient to adopt a breadth-first search. In this case dominance tests are effective only between states of the same cell of matrix \(M\), because when a new state \(S'\) is generated all states potentially dominated by \(S'\) in previous cells have already been generated and therefore there is no advantage in deleting them; on the other side all states that can dominate \(S'\) from successive cells on the same diagonal have not yet been generated. Dominance tests are however effective at reducing the number of states even if the tests are restricted to only one cell of \(M\). This approach is made viable also by the effectiveness of the lower bound described above.
To make bounding effective the algorithm needs a heuristic solution to be computed initially. This can be done in many different ways: in our experiments we used a fast local search algorithm presented in [2].

4. Computational results

Computational experiments have been carried out on a PC 1.4 GHz, 512 MB RAM with Linux operating system. Algorithms have been coded in C programming language. The problem instances for the tests have been obtained from the TSPLIB, where the Euclidean instances have been rounded up and origin/destination pairs have been defined at random.

Table 1 reports the outcome of some of our experiments. Column \( V \) reports the number of vertices (excluding the depot), column \( GEN \) reports the number of states generated, column \( DOM \) reports the number of states deleted by the dominance test, column \( BND \) reports the number of states fathomed owing to their lower bound, column \( DEL/GEN\% \) reports the percentage of states deleted (for either dominance or bounding), column \( Time \) reports the computing time in seconds, column \( k \) reports the dimension of the hash vector and column \( \overline{m} \) reports the average length of the state lists in matrix \( M \).

<table>
<thead>
<tr>
<th>( V )</th>
<th>( GEN )</th>
<th>( DOM )</th>
<th>( BND )</th>
<th>( DEL/GEN% )</th>
<th>( Time )</th>
<th>( k )</th>
<th>( \overline{m} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>1265</td>
<td>144</td>
<td>651</td>
<td>62.85</td>
<td>-</td>
<td>1</td>
<td>5.04</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.010</td>
<td>10</td>
<td>2.05</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.010</td>
<td>100</td>
<td>2.05</td>
</tr>
<tr>
<td>12</td>
<td>6882</td>
<td>787</td>
<td>3780</td>
<td>66.36</td>
<td>0.010</td>
<td>1</td>
<td>12.50</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.010</td>
<td>50</td>
<td>3.66</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.020</td>
<td>100</td>
<td>3.66</td>
</tr>
<tr>
<td>14</td>
<td>40464</td>
<td>4783</td>
<td>29401</td>
<td>84.48</td>
<td>0.090</td>
<td>10</td>
<td>7.72</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.090</td>
<td>100</td>
<td>6.42</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.120</td>
<td>300</td>
<td>6.42</td>
</tr>
<tr>
<td>16</td>
<td>95856</td>
<td>9170</td>
<td>63257</td>
<td>75.56</td>
<td>0.330</td>
<td>1000</td>
<td>8.46</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.320</td>
<td>2000</td>
<td>7.40</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.510</td>
<td>3000</td>
<td>7.40</td>
</tr>
<tr>
<td>18</td>
<td>826128</td>
<td>64848</td>
<td>569845</td>
<td>76.83</td>
<td>2.100</td>
<td>1000</td>
<td>25.80</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2.000</td>
<td>3000</td>
<td>8.30</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2.100</td>
<td>5000</td>
<td>8.30</td>
</tr>
<tr>
<td>20</td>
<td>1343910</td>
<td>189376</td>
<td>698265</td>
<td>66.05</td>
<td>20.450</td>
<td>5000</td>
<td>28.40</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>19.987</td>
<td>8000</td>
<td>9.40</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>20.340</td>
<td>10000</td>
<td>9.37</td>
</tr>
</tbody>
</table>

Table 1. Experiments with the dynamic programming algorithm.
From the results reported it appears that dominance and bounding are quite effective in fathoming a large percentage of states. A suitable tuning of the hashing parameter \( k \) may influence the average length of the lists of states that must be scanned to perform the dominance test. However the computing time is not strongly affected by this choice.

Besides solving at optimality PDPRL instances with a dozen customers, this dynamic programming approach paves the way to the development of pricing algorithms for multi-vehicle pick-up and delivery problems with rear-loading constraints by column generation: in that case each vehicle is not forced to visit all customers but rather a prize is collected for each customer visited. Applying our exact optimization approach to the multi-vehicle pick-up and delivery problem will be a future development of the work presented here.

References


A DYNAMIC ACTIVITY BASED TRAVEL DEMAND MODEL TO ASSESS STOCKHOLM CONGESTION CHARGING SCHEME
- ESTIMATION USING MARKOV CHAIN MONTE CARLO AND REINFORCEMENT LEARNING

Anders KARLSTROM∗

Abstract. In this paper we develop a new method for estimating activity-based models using approximate dynamic programming.

1. Introduction and motivation

In this paper we will report results from ongoing research which aims at developing and estimating a dynamic activity based model which is computationally efficient and microeconomic sound to enable welfare economic assessment. The objective is two fold. First we want to partially assess the welfare economic impact of the currently implemented congestion charges in the Stockholm metropolitan area. We want to assess welfare impact taking into account changes in travel frequencies, mode choice, trip chaining, departure time, and rescheduling in an activity based travel demand model. Second, we want to use this model to assess proposed small changes in the congestion charging scheme, i.e. for forecasting purposes.

State-of-the-art microeconomic (discrete choice) models for transport demand are predominantly static in nature, in the sense that they are inadequate to capture timing. In our application, time of departure is an important potential adjustment for individuals as the congestion charge is differentiated in time. Although departure time choice models have been proposed and estimated, they are either only partial (only models time for departure to work, for instance) or non-dynamic in that they do not reflect an environment where individuals take decisions sequentially in an uncertain environment. For instance, if travel time to work is uncertain, this uncertainty is resolved after having arrived to work in the morning, and the rest of the day can be replanned in light of this information.

In fact, also theoretical models used for derivation of measures of travel time savings are inherently static in nature. We argue that we need to allow for sequential decision making in an uncertain environment if we are to derive welfare measures for travel time uncertainties.

∗Royal Institute of Technology, Stockholm, Sweden. andersk@infra.kth.se
In our ongoing work, we have suggested to use a markov decision process (MDP) modelling framework. From a theoretical perspective, the MDP framework provides a solid microeconomic framework to assess valuation of changes in travel time and travel time uncertainties, taking into account not only mode choice and destination, but also rescheduling of the activity pattern and travel departure choices. In our case, if the individual as a response to a congestion charge starts earlier in the morning (to avoid the congestion charge), this may result in an earlier travel home from work in the afternoon. A decreased travel uncertainty due to less congestion may result in a later departure for a segment of the population, while rescheduling may occur in the case of unusual delays. As a final example, if an individual changes mode from car to public transportation in response to a congestion charge, this may result that the shopping activity on the way home from work (by car) now is replaced by a shopping trip from home later in the evening. From a theoretical perspective, all these examples can easily be accommodated in our the MDP framework.

2. A dynamic microeconomic framework

In this section we will develop the microeconomic framework for sequential decision making in an uncertain environment. We will here use the principle of dynamic programming (DP) for solving markov decision problems (MDP). In the following we will formulate the activity pattern problem for one individual as an MDP.

Let us first describe the concept of a (discrete time) markov decision process, in which the decision maker being in state \( s_t \) in time \( t = 1, \ldots, T \) takes an action \( a_t \) that will determine the immediate utility \( u(s_t, a_t) \) and also determine the distribution of the next period’s state \( s_{t+1} \), represented by a Markov transition probability matrix \( p(s_{t+1} \mid s_t, a_t) \) (that may or may not be known in advance by the individual). The individual seeks a decision rule \( \delta(s_t) \) that solves

\[
V(s) = \max_{d(s_t)} \mathbb{E}\left\{\sum_{t=0}^{T} \beta^t u(s_t, a_t) \mid s_0 = s\right\}
\]  

(1)

where \( \mathbb{E} \) denotes expectation with respect to the stochastic process \( s_t, a_t \) induced by the decision rule \( d = (d_1, \ldots, d_T) \).

3. Estimation

In a maximum likelihood or GMM context, estimating a dynamic programming econometric model is much more difficult than finding the optimal solution for one single agent, since we will have to solve the problem for each individual in the sample for different parameters in each iteration. To overcome the obstacles we will use different techniques: (i) MCMC estimation, (ii) value function approximation (iii) actor-critic algorithms (iv) randomization.

Estimation of dynamic programming (DP) econometric models has grown popular in labor economics (e.g. French, [7]) and, more recently, in marketing science (e.g. Imai et al, [11]). One method is the nested fixed point poly-algorithm (NFXP) due to Rust. In this algorithm, the DP problem for one individual is solved by policy iteration for a given
parameter vector, and a hill-climbing algorithm is used to find the maximum likelihood parameter vector, using Newton-Kantorovich iterations to achieve fast convergence towards the maximum likelihood solution.

However, we can improve the computational efficiency if we can solve the DP problem only \textit{approximately} in each iteration. For instance, Aguirregabiria and Mira [1] develop one such algorithm in which fix point and hill-climbing iterations are swapped in comparison with the NFXP algorithm.

Imai et al [11] showed that it is possible to use MCMC techniques, that recently has grown popular in the marketing science field, in the context of estimating a discrete choice dynamic programming model. One advantage with MCMC methodology is that it does not involve optimization methods, while giving more efficient standard errors of estimated parameters for finite sample. Of course, Bayesian methodology (and theory) becomes feasible. In the discrete choice econometric literature, simulation methods has brought together the bayesian strand of methods with the methods of the classical approach. For instance, Huber and Train [8] demonstrate how the classical maximum likelihood estimates can be replicated by the hierarchical Bayes.

Hence, the suggested approach is to only \textit{approximately} solve the DP problem by only using a limited value iteration in a MCMC estimation framework. Following Nagel and Marchal [16], and in particular Charypar and Nagel [4], we explore other methods from computer science and machine learning literature to approximately solve for the value function. These methods include simulation methods using reinforcement learning with value function approximation.

Why do we need the value function? These represents the continuation pay-off by the individual, and they are in our case a crucial part of the utility function at a given decision moment. They are in turn given by the Bellman equation, given immediate rewards, which are the fundamental representation of the preferences. Without preference parameters, we will not be able to back out the value(s) of time, and welfare economics will become infeasible. However, if we were to believe value(s) of times that were given from some other source, it may be possible to estimate a model for the policy functions directly, without relying on the value functions. However, even in this case, experience from the machine learning literature shows that is is beneficial to learn the value functions at least approximately to speed up convergence (Konda and Tsitsikis, [13]). So, we do need the value functions, at least approximations thereof. On the other hand, having an actor-based method may be benefcial for fast implementation (rather than estimation) of the model, an issue that is left for further research.

Given that the individual acts in an environment that can be formalised by the MDP modelling framework, we apply methods that approximately solve the dynamic programming problem. In fact, we do not have to accurately solve for policies which is rarely of never visited. Moreover, some mandatory activities such as sleeping, school and working shrinks the number of visisted states considerably. Therefore, approximately solving the the dynamic programming problem may be considerably less difficult than solving it exactly for each state.

Recently, Charypar and Nagel [4] have tested a reinforcement learning algorithm, known as the tabular Q-learning algorithm. They show that the algorithm is computationally feasible for solve realistic scheduling problems of the same size as tested in, for instance
Jonsson and Karlstrom [9]. In this paper we demonstrate that the algorithm can be extended, and at the same time estimated using real world travel survey data.

A particular function approximation that has been successfully employed in the activity based literature is the decision tree. In our framework, we could use state aggregation to arrive at policies that are based on decision trees, which is convenient when we want to enforce spatial and temporal constraints. Although some activity based models do not do this, it is desirable that individuals cannot behave in disagreement with the physical environment, for instance travel more than 24 hours a day, or travel faster than the environment allows. Hägerstrand’s time geography provides a conceptual framework for understanding the constraints. However, as noted by Miller [14], variable travel time is not adequately integrated by the traditional the concept of the space-time prism. Wu and Miller allows the prism to be time dependent, but further theoretical development is needed to allow for stochastic travel times.

Finally, randomization is an important feature of our estimation methodology. In a technical sense, randomization can actually break the curse of dimensionality in dynamic programming problems with finite action space (as ours), see Rust [23]. In any practical sense, the possibility of sample along the most probable paths, rather than enumerating all possible paths is an important ingredient to make the activity based scheduling problem tractable in real world problems. Imai et al [11] shows how randomization can be helpful by sampling new states in each iteration when approximating the value function. At the same time, we are able to use information from similar individuals to increase accuracy and speed.

One important aspect in our choice of activity based model is scalability. In Karlstrom [12] and Jonsson and Karlstrom [9] we demonstrate that the one-day activity scheduling problem easily can be solved for one individual using standard dynamic programming backward induction techniques. Charypar and Nagel also report on the feasibility of realistic scheduling problems. However, using this technique to model households with more than one person is intractable, either using backward induction by dynamic programming, or tabular Q-learning. If there is a million states for one individual (which is quite sufficient for a realistic model), introducing another household member would give $10^{12}$ states, which is not tractable to handle with standard dynamic programming backward induction methods. In the proposed framework, value function approximation and simulation methods are essential tools.

Finally, MCMC methods, as well as maximum likelihood estimation, are embarassingly parallel, and lends itself easily to parallel implementation. We will use a simple (but, in practical terms, effective) infrastructure for distributed computing with a heterogeneous cluster.

References


AN AGENT-BASED APPROACH TO COMBINED ROUTE AND MODE CHOICE

Gustavo KUHN ANDRIOTTI *
Franziska KLÜGL *
Guido RINDSFÜSER †

Abstract.
Combined route and mode choice models can be solved by traditional macroscopic approaches. However, the solutions are becoming far from trivial. Agent-based techniques form an interesting alternative for simulating individual decision making based on bounded rationality.

In this paper a simple agent-based model is proposed. On that, the actual decision making process can be based on established discrete choice mechanisms. But the alternative set to chose evolves according to two agents’ procedures: exploiting their mental maps and exploring the general network.

1. Introduction
Models of route choice are highly valuable in traffic simulation. Now, there are several macro- and microscopic approaches for resembling (rational) decision making about the path to follow from origin to destination locations. One major issue is the generation of alternatives between which a decision maker may select its actual route.

In real-world networks the generation of a master set containing all possible routes is neither practicable nor realistic due to the possible restricted knowledge of a traveller. And a combination of route and mode choice makes the set even bigger.

In this paper, a microscopic approach is proposed to combined route and mode choice. For that, agent concepts are applied for learning alternatives for selection.

On that account, this text continues by giving details about how a traffic network is represented. Then, it explains how the agent mental map is constructed and updated. In the following section, the agent reasoning engine is explained, which uses this

*Department of Artificial Intelligence, University of Wuerzburg, Am Hubland, DE-97074 Würzburg, Germany. {andriotti,kluegl}@informatik.uni-wuerzburg.de
†Emch+Berger AG Bern, Gartenstrasse 1, P.O. Box 6025, CH-3001 Bern, Switzerland. guido.rindsfueser@emchberger.ch

358
mental map for route and mode choice. The paper ends with a short discussion about related work and followed by conclusion, in sections 4. and 5. respectively.

2. Network Representation

The basic structure used to represent the traffic network is a graph. It represents all necessary decision points, i.e. crossings, but omits redundant information about actual link geography. Graphs are applied on two levels of representation: the representation of the actual network – the “World” – and a graph that denotes the agents current beliefs about the network – the “mental map”. The first is the whole network and is used as environment for all the agents. The later contains the specific knowledge of individuals.

A schematic view of a network and its graph representation is pointed out on figure 1(a).

![Network representation and different stages of mental map during agent learning process](image)

(a) Network to Graph. (b) Agent Mental Map.

Figure 1. Network representation and different stages of mental map during agent learning process
2.1. Environmental Model

As mentioned above, the network must first be transformed into a graph: crossings are vertex and links are arcs. This network representation is denoted in the remainder of the paper as “network map”.

Every arc is assumed directed (that is not depicted on figures 1(a) and 1(b) to keep it simple). Therefore, every link direction is represented by a directed arc.

Also, different transportation modes are not visible in figure 1(a). They are mimic by different arcs, even if they share the same physical space on real world. Vertices connect the otherwise distinct networks for allowing mode changes.

2.2. Mental Map

In figure 1(b) an example of an evolving mental map is depicted. Starting with only a few known paths, the agent learns about the network, when it explores new paths to its destination. In principle, the mental map do not supposed to cover the entire network. Of course, an agent can evolve its mental map to be a copy of it, but based on concepts of bounded rationality, this would not be realistic.

Thus, the mental map is a dynamic graph that changes according to agent’s experience. So, its structure and dynamics are highly related to the particular agent model involved. Experience is “stored” in the agent’s mental map attributes. Therefore, it really represents agents’ beliefs about the network.

As depicted in figure 1(b), edges on the mental map are not regular link representations rather a sequence of links/arcs from a map. For example in figure 1(b) a, d is a single edge on agent’s mental map, but a sequence of arcs – arc a followed by d – in the network. Then a route is a sequence of sub-routes. A sub-route is assumed to be a sequence of arcs that leads from a crossing to another. In an extreme case a sub-route will have just one arc and therefore its edge representation will contain just that arc.

The agents exploit their individual mental maps for constructing alternative routes and selecting between them. So, the set of alternatives is learnt based on the evolving mental map. This is detailed on the following section.

3. Agent Reasoning Engine

As mentioned before, the evolution of the mental map depends on the agents’ behaviour and decision making. A sketch of the agents’ reasoning engine is shown in figure 2.

In that figure the relevant blocks are Change current route? and Explore for new Routes?. There, the agent may deviate from its initial route and mode decision. The agent may access information about arcs, that are not currently integrated into its mental map. This is the prerequisite for exploring new sub-routes resulting in mental map growth.
Figure 2. Decision making process of the agents in our model.

It is assumed that the agents have an individualised fixed origin - destination pair. In most of the cases an agent won’t change its initial choice, as expected.

After selecting an initial route, an agent starts to follow this arc sequence. When it reaches the destination, it stops, evaluates and stores its experience. Every time an agent reaches a non-destination node/vertex it reasons about re-planning its route. This re-planning step is triggered by events, like a blockade on the following arc/link, or spontaneous willingness to change it, a low probability.

When an agent is willing to change its route, there are two alternatives: reformulation or exploration. Reformulation means that the agent uses its mental map for finding a new path to its destination. In agent learning, this alternative is also called “exploitation”, as the agent uses previously acquired knowledge for problem solving. Thus, the longer the agent “lives” in its environment, the more non-standard routes are available to it.

If an agent decides for exploring, it gathers information from the network map and tries to generate a new arc sequence leading from its current position to destination. On figure 1(b) that happens from 1 to 2.

After having followed a new route, the agent has to integrate the new information about the network into its mental map. It has to check whether there are common arcs/vertices with existing edges/nodes in its mental map. If so, they have to be condensed on agent’s graph to ensure consistency and avoid redundancy. For instance, it happens on figure 1(b) from states 2 to 3.

It is also important to “forget” edges, not only for efficiency reasons, but also due to realism in agent behaviour. Therefore, agents must eliminate edges from their
mental maps. A history based heuristics seems to be appropriate.

3.1. Mode Choice Validation

A route, when generated by exploration, must be “mode” valid. That means, only realistic combinations of modes should be allowed. That is important as the agents may choose multi-modal routes.

It is proposed the use of a grammar to check allowable route composition. So, a route can be only composed if it is a valid “word” on the grammar depicted in 1. The “letters” of that “word” are arcs with the transportation mode as the only significant arc attribute.

\[
\begin{align*}
W & \rightarrow \text{walk} \\
B & \rightarrow \text{bike} \\
P_1 & \rightarrow \text{public transportation where bikes are not allowed} \\
P_2 & \rightarrow \text{public transportation where bikes are allowed} \\
I & \rightarrow \text{private individual vehicle} \\
T & \rightarrow \text{multi-modal transportation route}
\end{align*}
\]

\[
\begin{align*}
B_W & \rightarrow B | WB | BW | WBW \\
P_{1W} & \rightarrow P_1 | WP_1 | P_1W \\
P_{2W} & \rightarrow P_2 | WP_2 | P_2W \\
I_W & \rightarrow I | WI | IW | WIW \\
B_A & \rightarrow B_W | B_WP_2 | P_2WB | B_AB_A \\
P_A & \rightarrow P_{3W} | P_2W | P_A | P_{3W} \\
I_A & \rightarrow I_W | I_WP_A | P_AI_W \\
T & \rightarrow W | B_A | P_A | I_A
\end{align*}
\]

This grammar just defines all possible transportation mode combinations. For instance, the rule \(P_{1W} \rightarrow P_1WP_1WP_1W\) denotes that when using public transportation without bikes, the agent is only allowed to walk to take that transit and also just to walk afterwards. It is not possible to “suddenly” use a car. Of course, some seldom real-world valid combinations are not permitted according to this grammar.

4. Related Work

This paper presents an alternative approach to standard route choice derived from econometrics ([3, 4, 7] and others). There, route choice is based on a master set – containing all possible paths – or on some arbitrary subset. The idea of using sub-path to built routes can be already found in [5]. Nevertheless, here a microscopic approach is used and it focuses on individual network knowledge.

It is not new the idea of using agent-based techniques to simulate traffic relevant decision making, examples can be found in [2, 1, 9]. However, these approaches are
only tested with very abstract networks, only with two routes. A BDI approach, like in [9], seems not apt for more complex scenarios. Hence, more simple models of decision making may be used like a strategy based algorithm from [2, 1], that are based on reinforcement learning. In [8] an even more complex strategy based algorithm is proposed for the day scheduling level.

There are some issues using meso levels, where routes are generated to allow a macroscopic analysis. In [6] some non-obvious behaviours on microscopic level related to public transportation could be observed. However, those behaviours can be better modelled using an entirely microscopic model like the one presented here.

5. Future Work and Conclusion

In this paper is presented an algorithm to reproduce individuals route and mode choice on a microscopic level. This approach focuses on the individual’s network knowledge and expertise.

The current status of the project is that a first prototype is ready and works for simple networks. The next step will be applying the agent model to a real-world network. This will involve additional tasks related to calibration and validation. Especially the latter turns out to be demanding as a comparison will be made between routes that were actually taken by subjects and routes that are selected by agents in a simulation. This problem relates to validation of traffic simulation on the microscopic level, that in general is also an interesting and highly demanding research topic.

6. Acknowledgment

Gustavo Kuhn Andriotti would like to thanks CNPq for financial support on his PhD.

References


LEVEL CROSSING RISK ASSESSMENT: ANALYSIS AND SIMULATION OF SYSTEM

Alessandro BALDASSARRA1; Stefano IMPASTATO1, Gabriele MALAVASI1

Abstract. Level crossings represent intersection with high risk and cause serious damages and injuries to many thousand of road users and railway passengers. The paper analyzes various alternative systems for level-crossing protection and problems relating to the systems adopted. The aim of the research is the evaluation of road user’s behaviour approaching a level crossing and the influence of the different typologies and characteristics of level crossing. Different phases of closing barrier are simulated using Petri net model.

1. Introduction

Road/rail grade intersections are unique in the world of transport because they represent the only case in which two different infrastructures, placed under different responsibilities and travelled by vehicles with different performances, converge and meet during their normal operation. The result is that these intersections constitute high-risk spots for all railways in the world.

Each year, accidents at level crossings not only cause deaths or serious injuries to many thousands of road users and railway passengers, but also impose a heavy financial burden in terms of interruption of railway and road services and a lot of damages to railway and road vehicles and property.

The great majority of these collisions are caused by the negligence, incompetence or inability of road vehicle driver.

Aim of the study is:

• to analyse various alternative systems for level-crossing protection;
• to present technologies and problems relating to the systems adopted for level-crossing protection;
• to investigate different level crossing typologies and the influence of the characteristics on transit time;

1 University of Rome “La Sapienza” DITS Transport Area - Via Eudossiana, 18 - 00184 Rome Italy, stefano.impastato@uniroma1.it
• to evaluate the approaches of the application of quantitative risk;
• to simulate, using methodologies based on formal methods and simulations (e.g. Petri nets), what really happens in correspondence to a level crossing during the train operation.

2. Level crossing characteristics

Level crossings represent a point of intersection between rail and road. Its can be classified depending on the road circulation (e.g. pedestrian, carriageable, etc.); depending on their position along the rail network, (e.g. station, line, etc.); depending on their closing types (normally opened and unmanned, private with own responsibility, with request opening, timing closing, remote control of barrier, automatic, etc); depending on closing system (light signs and vane, with gate, with single barrier, with half barrier, with double barrier, etc).

Studying the existing level crossing, it’s possible to classify them in four different types. The first class includes the automatic level crossing with barriers, the second class includes the level crossing with manual barriers, the third class includes level crossing without barriers, with flashing lights, sound alarms, rail and road signs, and the forth class includes the level crossing without barriers and road signs only.

With these four classes it’s possible to list most types of the Italian and European carriageable level crossings. In table 1 the four types of level crossings are represented.

<table>
<thead>
<tr>
<th>Types</th>
<th>Characteristic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type 1</td>
<td>Automatic LC with barrier and flashing light</td>
</tr>
<tr>
<td>Type 2</td>
<td>Manual LC with barrier</td>
</tr>
<tr>
<td>Type 3</td>
<td>LC without barrier with light, sound and road signs</td>
</tr>
<tr>
<td>Type 4</td>
<td>LC without barrier and only road signs</td>
</tr>
</tbody>
</table>

Table 1. Level crossings (LC) types

3. Automatic level crossing with barriers

For an automatic level crossing, with single or half barriers, the movement of the barriers is delayed respect to the road warning traffic light activation (about 20 seconds).

Generally in this time (from the beginning of activation of road warning signal to the successive movement of the barriers) the road users approach the level crossing and continue their running. Nevertheless, in some cases, road users stop their running only when the barriers are yet partially closed with a certain angle of rotation. This problem is the focus of research.

In particular the behaviour of the road users related to closure of the barriers has been investigated. In fact it’s possible that some road users are not able to cross the level crossing before the exit barrier is been closed.

The height of barrier depends on time and rotation speed. In the equation (1) the height of barriers in function of rotation speed and time is represented.
\[ h_{\varphi} = L \cdot \tan \left[ \frac{\pi}{2} - \varphi \right] = L \cdot \tan \left[ \frac{\pi}{2} - (\omega \cdot t) \right] \]  

(1)

In which:

- \( h_{\varphi} \): height of barrier
- \( L \): distance from vehicle to rotation axis of barrier
- \( \varphi \): rotation angle
- \( \omega \): rotation speed of barrier
- \( t \): rotation time

In particular for a level crossing with single barrier the height is different for entering and exiting barrier, and depends on different distance from vehicle to rotation axis of barriers.

In figure 1 a level crossing with single barrier is represented.

![Figure 1. Level crossing with single barrier](image)

It’s possible to notice the different height (at the same time) of barrier 1 (entering) and barrier 2 (exiting) for a vehicle that runs on a lane.

In figure 2 the height of barriers is represented. The height of barriers is calculated using equation (1), considering different horizontal distances between longitudinal vehicle axis and barriers rotation axis. Conventionally the position of the vehicle is represented in the middle of the lane. The speed rotation of barriers used is 0.1047 rad/s: this speed allows the closure of barriers in 15 seconds.

At the same time the height of two barriers is different. Even if the total closure time is the same for both barriers, the barrier 1 is always shorter than barrier 2. This situation represents a natural delay of closure for barrier 2.

Fixing a height limit for the transit of the vehicle under the barrier 1 it’s possible to determine the transit time through the level crossing. In case represented in figure 2 with a height limit of 5 meters the transit time limit under the barrier 1 is about 6 seconds from the beginning of the movement of barriers, and the transit time limit under the barrier 2 is about...
10 seconds. The road user has an interval time $\Delta t_{1,2}$ (in this case it is equals to 4 seconds) to pass through the level crossing. If the limit height of barrier decreases the interval time $\Delta t_{1,2}$ will reduce also, and this can generate critical situations.

![Figure 2. Height of single barrier, closure time and transit time](image)

With the adoption of double half barriers to have a $\Delta t$ value that allows transit of vehicles that approach the level crossing with partial closure of barrier 1 it’s necessary to delay the closure of barrier 2.

In this second case the total closure time is no more the closure time of the barrier 1 (15 seconds as the single barriers) but this depends on delay.

In figure 3 the closure time of double half barriers is represented. With a delay $\Delta t$ of 4 seconds the total closure time is 19 seconds.

![Figure 3. Height of double half barrier and closure time](image)

The extension of total closure time using double half barrier could lead a handicap for the train operation and for road circulation (e.g. extended waiting time for car, etc).
4. **Petri net simulation**

Using Petri nets the closings of different types of level crossings have been simulated. In particular the simulation of level crossings with single barrier closure is presented. It’s been simulated also the behaviour of car drivers, who doesn’t respect the road signals. In fact the average closing time of the barriers is 15 seconds. During the first 5 seconds incorrect car drivers continue their running, and cars pass while the barriers are in the closing-phase, even if the road traffic lights are at danger 20 seconds before the barriers closing-phase.

All devices considered and designed with the Petri nets are represented in the top level. In particular the devices simulated are: track circuits; foot switches for opening and closing barriers; devices for opening and closing verify; road and rail signals; transit of cars on road.

Normally the barriers are open and the cars can pass through the level crossing. When a train passes over a foot switch (closing), immediately the road traffic lights becomes at danger. After 20 seconds, the barriers begin to close (closing phase is 15 seconds long) and the cars stop at traffic lights, but during the first 5 seconds it’s possible that incorrect car drivers pass through the level crossing.

After this period, the closing device verifies that the barriers are closed; then the rail signal becomes at clear and the train can continue its running.

During the train operational time the cars build up at traffic lights until the train passes over the opening devices (foot switches).

Only when is verified the opening of the barrier, the cars can start one by one. If other car drivers reach the level crossing while other vehicles are yet waiting for running, they have to queue.

Using the Petri nets simulation it’s possible to know the status of the different devices step by step.

The tool used for the simulation gives numerical results, as represented in the figure 4. In the first column is represented the time; in the second the devices; in the third the action and in the fourth the status change of the devices.

![Figure 4. Output: numerical representation of the status change of devices and car queue](image)

Using the numerical data it’s possible to report the status of devices in a graphic representation, as shown in figure 5.
5. Conclusions

Aim of the research is the level crossing risk assessment. In the first part of the papers has been investigated different typologies of level crossing and the influence of different parameters on railway and road operation. In the second part, using formal methods, has been simulated the road users behaviour and the closure of level crossing barriers.

This approach permits to evaluate and to quantify the risk connected on a level crossing.

References


A TIME-SPACE NETWORK BASED APPROACH FOR INTEGRATED VEHICLE AND CREW SCHEDULING IN PUBLIC TRANSPORT

Vitali GINTNER, Ingmar STEINZEN, Leena SUHL

Abstract. This paper proposes a model for the multiple-depot integrated vehicle and crew scheduling problem which is based on a novel time-space network approach. The main contribution of this modeling technique is that possible connections between trips are not explicitly presented by arcs, but trips are implicitly connected by the flow in the network. This results in fewer arcs and a smaller MIP which reduces the computational effort. The approach presented in this paper has been evaluated with academical instances available and compared with recent references from literature.

1. Introduction

The integrated vehicle and crew scheduling problem (IVCSP) combines the (multiple-depot) vehicle scheduling problem (MDVSP) and the crew scheduling problem (CSP). Given a set of trips, it minimizes the total sum of vehicle and crew costs such that both the vehicle and the crew schedule are feasible and mutually compatible. Each trip has fixed starting and ending times and can be assigned to a vehicle and a crew member from a certain set of depots.

Traditionally, vehicle and crew scheduling problems have been approached separately, so that vehicles are first assigned to trips, and in a second phase, crews are assigned to the vehicle blocks calculated before. It is well known that the integrated treatment of vehicle and crew scheduling can lead to efficiency gains.

A completely integrated model of vehicle and crew scheduling for the single depot case was first investigated by Freling and co-authors, e.g. [3, 2]. They solve the combined model by Lagrangian relaxation and column generation. An exact algorithm for the single-depot vehicle and scheduling problem was proposed in [4]. The authors use a set partitioning formulation of the problem and solve it by combining column generation and cut generation in a branch-and-bound algorithm. A further contribution for the single-depot case can be found in [6]. They solve the crew scheduling problem while incorporating side constraints

*Decision Support & OR Lab, University of Paderborn, Warburger Str. 100, D-33100 Paderborn, Germany. {gintner,steinzen,suhl}@dsor.de
for the vehicles. This is done in such a way, that the solution of this problem guarantees an overall optimal solution being found after constructing a compatible vehicle schedule.

An extension of the integrated model [2] to the multiple-depot case is addressed by [7]. Another approach for the multiple-depot case was proposed by [1]. Their main contributions were using bundle technique instead of subgradient optimization, and fixing variables during column generation. Very recently, a branch-and-price approach which uses the formulation presented in [7] is discussed in [9].

Approaches mentioned above are based on a connection-network which contains explicit connection arcs between each pair of compatible trips. Thus, the number of arcs grows quadratically with the number of trips. Note that the complexity of the MDVSP (which is part of the IVCSP) strongly depends on the number of arcs in the network since each arc is represented by a decision variable.

The model formulation of the IVCSP presented in this paper is based on a time-space network. The time-space network modeling was successfully applied to the multiple-depot vehicle scheduling problem, see e.g. [8, 5]. The main contribution of this approach is that possible connections between trips are not explicitly presented by arcs. Instead, compatible trips are implicitly connected by the flow in the network. This results in fewer arcs and a smaller MIP which reduces the computational time needed.

The paper is organized as follows. In Section 2., we describe the time-space network model. Section 3. provides the mathematical formulation. In Section 4., we present the solution approach. Finally, computational results and a comparison with results presented in [1] are provided in Section 5..

2. Model

For the sake of simplicity, we assume that a trip is equivalent to a task (each trip has a relief point at the start and the end). Nevertheless, the model can be easily extended to the general case. We start with the case with only one depot. For each trip we create two nodes representing start and end points, and an arc in-between representing the trip.

Furthermore, we create two additional depot arcs for each trip: from-depot arcs connect a start-depot node with a trip-start node while to-depot arcs originate at a trip-end node and terminate at an end-depot node. All nodes are grouped by corresponding stations and sorted by ascending time.

We create a waiting arc between two consecutive nodes at the same station if there is not enough time to perform a round-trip to the depot. Waiting arcs represent vehicles waiting at a station. Thus, a trip arriving on its end station can be connected with each trip departing later from the same station through a flow using waiting and/or depot arcs.

Furthermore, it is possible to connect trips between different stations by creating dead-heading (dh) arcs. For each trip \(i\), we create a dh-arc from its arriving node to the first available departing nodes on every other station. Note that for each trip, there is at most one dh-arc to each station. All later trips are connected with \(i\) through the dh-arc and a sequence of waiting and/or depot arcs. Thus, all allowed connections between compatible trips are implicitly included. Let \(n\) and \(m\) be the number of trips and stations, respectively. Then the number of arcs is \(O(nm)\) instead of \(O(n^2)\) for the connection-network model,
while usually \( n \gg m \). Moreover, the number of dh-arcs can be reduced by applying the aggregation technique described in [8] and [5].

Finally, we create a circulation arc from the last to the first depot node. A path from the first to the last depot node represents a day schedule for one vehicle. We create such a network for each depot. A detailed description of the time-space network for MDVSP can be found in [8] and [5].

Figure 1 shows an example of time-space network for a problem with 6 trips and 3 stations. As one can see, a flow is possible between each pair of compatible trips (e.g. a feasible path which represents a day schedule for one vehicle could be: \( d_1, t_1, w_1, t_2, w_4, dh_1, t_4, w_8, t_6, w_{12}, d_11 \)).

Figure 1. Time-space network

3. Mathematical Formulation

Let \( T \) be the set of trips and \( D \) the set of depots, respectively. For each depot \( d \in D \) we define the vehicle scheduling network \( G^d = (V^d, A^d) \), which is an acyclic directed network with nodes \( V^d \) and arcs \( A^d \) as described in Section 2. Let \( c^d_{ij} \) be the vehicle cost of arc \((i, j) \in A^d\), which is usually some function of travel and idle time. The cost of the circulation arc is set to the fixed cost for using a vehicle. Furthermore, let \( u^d_{ij} \) be the capacity of arc \((i, j) \in A^d\), which is 1 for trip and depot arcs and \( u^d \) otherwise, where \( u^d \) is the maximum number of available vehicles in depot \( d \). Denote \( A^d_T \subset A^d \) the set of arcs representing trips, and \( A^d_E \subset A^d \) the set of "empty" arcs, i.e. arcs which do not have to be covered by a duty (waiting arcs in the depot and the circulation arc). Let \( A^d_{T}(t) \) be the arc corresponding to trip \( t \in T \) in the network \( G^d \).

Furthermore, \( K^d \) denotes the set of duties corresponding to depot \( d \) and \( f^d_k \) the crew cost of duty \( k \in K^d \), respectively. Let \( K^d(i, j) \) denote the set of duties covering the task corresponding to arc \((i, j) \in A^d \setminus A^d_E \). Note, that we assume that a trip corresponds to exactly one task.

An integer variable \( y^d_{ij} \) indicates the flow value on arc \((i, j) \in A^d \) (each flow unit represents a vehicle using the arc), while a 0-1 variable \( x^d_k \) indicates whether duty \( k \) assigned to depot \( d \) is selected in the solution or not. The IVCSP can be formulated as follows.
The objective is to minimize the sum of total vehicle and crew costs. The first two sets of constraints, (2) and (3), correspond to the formulation for the MDVSP. Constraints (2) assure that each trip is covered by exactly one vehicle, while (3) are typical flow-balance constraints. The linking constraints (4) assure that if an arc is used by one or more vehicle (corresponding value of $y$) then it will be covered by the same number of duties, i.e. a driver is available for all vehicle activities.

4. Solution Approach

The algorithm we propose to solve IVCSP1, is a combination of column generation and Lagrangian relaxation and is very similar to the algorithm proposed by [7]. In the following, an outline of the algorithm is given:

**Step 0: Initialization**
Solve MDVSP and CSP for every depot and take as initial set of columns the duties of the CSP-solution.

**Step 1: Computation of dual multipliers**
Solve a Lagrangian dual problem with the current set of columns. This gives a lower bound of the objective function for the current set of columns.

**Step 2: Deletion of columns**
If there are more columns than a certain given number, then delete columns with positive reduced cost greater than a certain threshold value.

**Step 3: Generation of columns**
Generate columns with negative reduced cost. If there are no such columns (or another termination criterion is satisfied), go to Step 4; otherwise, return to Step 1.

**Step 4: Construction of feasible solution**
Solve a second Lagrangian dual problem with the set of columns generated in Step 3. Get a feasible vehicle schedule and solve CSP for each depot.
First, we compute a feasible solution and an initial set of columns by using the sequential approach and CPLEX. As proposed by [7], we use Lagrangian relaxation instead of LP-relaxation in order to solve the master problem (Step 1). In Step 2, columns with high positive reduced costs are removed, if the master problem gets too large. Afterwards, in the pricing problem (Step 3), new columns with negative reduced costs are generated. And finally, in Step 4, we compute feasible solutions. The next sub-sections provide a detailed description of steps 1-4.

The Master Problem
Since the complete formulation IVCSP1 is very difficult to solve, we use its Lagrangian relaxation in order to solve the master problem. First, the equality signs in the constraints (2) and (4) are replaced by $\geq$-signs, which are subsequently relaxed in a Lagrangian way. In doing so, we associate non-negative Lagrangian multipliers $\lambda_t$ and $\mu^i_{dj}$ with constraints (2) and (4), respectively. Then, the remaining Lagrangian subproblem can be solved by pricing out the $x$ variables and solving $|D|$ small single-depot vehicle scheduling problems (SDVSP) for the $y$ variables. The Lagrangian dual problem is solved by applying subgradient optimization. Furthermore, we use a procedure to update Lagrangian multipliers after solving the Lagrangian dual problem proposed by [2] and [7]. This procedure prevents duties in the current master problem from being re-generated during pricing.

An alternative approach is to relax constraints (3) instead of (2) (for this purpose a slight reformulation of (3) is needed). The remaining Lagrangian subproblem for $y$ variables corresponds with solving a big SDVSP instead of $|D|$ small.

The Pricing Problem
The number of possible duties is huge, since vehicle blocks are not known in the IVCSP. In the pricing problem, we use a two phase procedure related to one proposed by [2]. In the first phase, a set of pieces of work (a continuous sequence of tasks corresponding to (a part of) one vehicle block) is generated using a special piece generation network. In the second phase, a set of duties is generated from the obtained pieces.

The piece generation network $\hat{G}$ is an extension of network $G$ for the MDVSP (see Section 2.) except the circulation arc and waiting arcs in the depot. Nodes represent relief points, while each path between two nodes corresponds to a piece of work. The cost associated with each arc $(i,j) \in \hat{A}^d$ is defined in such a way that the cost of each path equals the reduced costs of the corresponding piece. Since the feasibility of a piece is only restricted by its duration, we generate the set of pieces by solving a shortest path problem between each pair of nodes in $\hat{G}$ that satisfy the duration constraint. [2] proved that by generating only this subset of pieces, the column generation optimality condition is satisfied.

In the second phase, we define a time-space duty generation network for each duty type where each piece of work is represented by two nodes and an arc in-between. Two additional arcs for each piece connect it to a source and a sink which represent a depot. The corresponding nodes of compatible pieces are connected by break-arcs. Furthermore, we reduce the number of nodes and arcs significantly by applying a special aggregation technique.
Furthermore, we denote a set of resources for each arc (e.g., working time, duty length etc.). A path from the source to the sink corresponds to a feasible duty if it satisfies the resource restrictions. The cost of the path equals the reduced cost of the corresponding duty. The subproblem is formulated as a resource constrained shortest path problem and is solved by dynamic programming.

Feasible Solution
In Step 4, we only relax constraints (4) in a Lagrangian way. Therefore, the solution of the remaining subproblem (which is MDVSP and a NP-hard problem) gives a feasible vehicle schedule. Since we start with good multipliers (the best one from the last iteration in Step 1), we need to solve only a few iterations of the subgradient algorithm to get good solutions. Finally, for the last $M$ feasible vehicle schedules, we solve the CSP for every depot in order to get feasible crew schedules.

5. Computational Results
We have tested our approach on some randomly generated instances published in [7] and available from the author. Each problem class contains 10 instances of 80, 100, 160, and 200 trips involving 4 depots, see [7] for a detailed description. We use the same duty rules as described in [7]. All tests were performed on an Intel P4 3.4GHz/2GB personal computer running Windows XP.

Table 1 reports average solution values for each problem class. Rows ref1 and ref2 give the sum of the number of vehicles and duties as published in [7] and [1], respectively. Row time2 gives the computational time published in [1] (executed on a dual Intel Xeon 3.0GHz/4GB running SuSE Linux 9.0).

<table>
<thead>
<tr>
<th>trips</th>
<th>80</th>
<th>100</th>
<th>160</th>
<th>200</th>
</tr>
</thead>
<tbody>
<tr>
<td>vehicles</td>
<td>9.2</td>
<td>11.0</td>
<td>14.8</td>
<td>18.4</td>
</tr>
<tr>
<td>duties</td>
<td>20.1</td>
<td>23.0</td>
<td>32.6</td>
<td>39.9</td>
</tr>
<tr>
<td>total</td>
<td>29.3</td>
<td>35.0</td>
<td>47.4</td>
<td>58.3</td>
</tr>
<tr>
<td>time</td>
<td>00:08</td>
<td>00:11</td>
<td>00:27</td>
<td>00:55</td>
</tr>
<tr>
<td>ref1</td>
<td>29.6</td>
<td>36.2</td>
<td>49.5</td>
<td>60.4</td>
</tr>
<tr>
<td>ref2</td>
<td>29.6</td>
<td>35.7</td>
<td>47.7</td>
<td>59.0</td>
</tr>
<tr>
<td>time2</td>
<td>0:13</td>
<td>00:21</td>
<td>00:44</td>
<td>01:46</td>
</tr>
</tbody>
</table>

Table 1. Results for ECOPT-Instances with 4 Depots Variant A

Our approach is faster than the best known algorithm presented in [1]. The problem complexity is reduced by using time-space network formulation and, therefore, solution times of the master problem were sped up. On the other hand, the pricing problem is solved faster by using the two-phase approach and a new time-space formulation for the duty generation network. Due to faster computation times, we could afford to weaken some termination criteria. Thus, better results were produced for each problem class. Note that
results and computational times shown in Table 1 were obtained after tuning our approach in order to realize a good solution quality. However, the computational time can even be improved by weakening quality requirements.

References


ON THE POTENTIAL OF SOCIAL-PSYCHOLOGICAL ASPECTS IN MODELLING TRAVELLERS’ CHANGE OF BEHAVIOUR

Yos SUNTITIOSO1, Erel AVINERI1, Kiron CHATTERJEE1

Abstract. This study attempts to investigate the potential of incorporating social-psychological aspects, such as social interaction, social learning/imitation and social influence, in modeling travellers’ change of behaviour. A laboratory experiment with real individuals is conducted to provide empirical evidence about the existence of these aspects in individuals’ and group behaviour. The experiment gives some indicative results, which will be used to develop an agent-based simulation model that is also highlighted in this paper.

1. Introduction

Economical aspects (e.g. car ownership, travel time, monetary cost) and individual-psychological aspects (e.g. attitude, habit, preference) have become common considerations on modelling travellers’ behaviour. Social-psychological aspects, such as social interaction, social learning/imitation and social influence, have often been left unconsidered. Given the fact that behavioural change does not take place in a social vacuum, broader society and its social values have important roles to play. Therefore these social-psychological aspects may influence travellers’ decision making and, furthermore, affect change of behaviour.

Change of behaviour is a dynamic process that occurs over time, which may involve a learning process. The concept of individual learning suggests that individuals learn from their past experience and utilise an adaptive decision making process to cope with uncertainty. There is another form of learning, the social learning, where individuals learn from others’ experiences or observed behaviours. In transport field, the individual learning concept has been studied (for review, see [1]), while social learning has not been investigated intensively although evidences from other disciplines [5] have shown that this kind of learning process is influential and important.

1 Centre for Transport and Society, University of the West of England, Coldharbour Lane, Bristol, United Kingdom BS16 1QY, {yos.sunitiyoso, erel.avineri, kiron.chatterjee} @uwe.ac.uk
2. Objectives of study

This study emphasizes on developing a model which considers social-psychological aspects, including social interaction, social learning/imitation and social influence. The situational context is a dynamic decision making process with characteristics including the availability of feedback and interdependent decisions between individuals. Interdependence can be found in a public goods (social) dilemma situation where personal decision of an individual will not only affect himself/herself, but also other members of group.

3. Methodology

The development of the dynamic model, representing a social dilemma of mode-choice, is based on two stages. First, a laboratory experiment with real individuals as the subjects has been conducted in order to provide empirical evidence about individuals’ and group behaviour. Second, an agent-based simulation is to be built to simulate and analyse behaviours of individuals in larger environments, including a greater number of individuals and groups, and in various situational settings. As this research is still on-going, in this paper we focus on the first part with an overview about the agent-based simulation model as the further step of study.

4. Laboratory experiment

A laboratory experiment with subjects having a role as employees of a company is used to provide empirical evidence about individuals’ behaviour whenever they have to face a situation of whether or not to contribute to an employer-based demand management measure to reduce employees’ car-use. The employer-based measure asks each employee to contribute by using bus, as an alternative to car, for a number of days in every month. A reward (bonus) is given by the employer to all employees. The size of the reward depends on the total contribution (collective bus-use) of the employees. The reward is then distributed equally to each of them regardless of the amount of their contribution. The experiment, which utilizes computer interface developed on Z-tree [2], simulates a repeated decision making environment. A time period of one month is represented in the experiment by one round. Each experimental session contains 10 rounds. There are three sessions in the experiment. Session 1 is a baseline treatment, where each subject knows only her own experience. In Sessions 2 and 3, subjects may obtain social information about the decisions or behaviours of close colleagues and of all colleagues respectively. Several hypotheses, such as the existence of individual and social learning process, equilibrium situation, and also the influence of majority and minority decision on individuals’ behaviour, underlie the setting of these experimental treatments.
5. Results and discussions

A preliminary run of the laboratory experiment with a group of subjects (N=9) suggest several findings about the influence of social-psychological aspects on travellers’ decision making.

Figure 1 shows the dynamics of total contribution (collective bus-use) in each round of decision making for all three sessions. Since all subjects in the group participate in all sessions, in addition to the effects of experimental treatments, their behaviours in a succeeding session might also be influenced by their experiences in the preceding session. Based on that, in the process of analysis we may consider the rounds in all sessions as 30 continuous rounds with a different treatment in each 10-round session. It can be seen from the graph that providing an option to access social information, as in Sessions 2 and 3, may influence subjects to behave differently by contributing more or less than in Session 1. In these sessions, forms of social learning may exist when subjects observe the behaviours of others.

Table 1 presents descriptive statistics (mean, standard deviation, and standard error of mean) of experimental results in all three sessions. Session 1 has a higher mean of total contributions than Session 2, since in its first four rounds (Round 1-4) the total contribution is high (see Figure 1). But then it decreases to around 60 days of collective bus-use in the last six rounds (Round 5-10). It is likely that some subjects might have learnt the strategy of contributing less and gaining more benefit from others’ contribution which results in a low level of total contribution. This situation continues to early rounds in Session 2 (Round 11-13), but then a higher level of total contribution can be achieved in Round 14-19, before it drops significantly at Round 20. Session 3 has the highest mean of total contributions. Although the total contribution fluctuates from round to round (Round 21-30), its level is relatively high (total contribution of at least 90 days or on average at least 10 days per subject) in four rounds of Session 3.

Figure 1. Dynamics of total contribution in Sessions 1, 2 and 3.
In order to know the significance of the difference between sessions, t-Test Paired Two Samples is conducted on the data of total contribution in these three sessions (Table 2). Although the means of total contributions in the three sessions give an indication that there are differences between results of all three sessions, the t-Test results show that only the difference between Sessions 2 and 3 is proved to be statistically significant with level of significance $\alpha=0.05$. While the difference between Sessions 1 and 2 ($t=1.340$) is only significant at $\alpha=0.15$.

### Table 1. Descriptive statistics

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Std. dev.</th>
<th>Std. error of mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Session 1</td>
<td>79.60</td>
<td>23.833</td>
<td>7.536</td>
</tr>
<tr>
<td>Session 2</td>
<td>63.20</td>
<td>22.175</td>
<td>7.012</td>
</tr>
<tr>
<td>Session 3</td>
<td>88.60</td>
<td>21.823</td>
<td>6.901</td>
</tr>
</tbody>
</table>

### Table 2. t-Test Paired Two Samples

<table>
<thead>
<tr>
<th>Pair</th>
<th>T</th>
<th>Df</th>
<th>Sig. (2-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sessions 1 &amp; 2</td>
<td>1.340$^a$</td>
<td>9</td>
<td>0.213</td>
</tr>
<tr>
<td>Sessions 1 &amp; 3</td>
<td>-0.884</td>
<td>9</td>
<td>0.400</td>
</tr>
<tr>
<td>Sessions 2 &amp; 3</td>
<td>-2.189$^b$</td>
<td>9</td>
<td>0.056</td>
</tr>
</tbody>
</table>

$^a$significant at $\alpha = 0.15$ (t-critical = 1.100); $^b$significant at $\alpha = 0.05$ (t-critical = 1.833)

Non-cooperative game theory predicts that no one will contribute anything since it is the most dominant strategy to just enjoy benefits from others’ contribution without giving own contribution. However, Session 1 of the experiment gives an indication that some degree of altruism, which can be defined as costly acts that confer benefit on other individuals, may exist since a significant proportion of individuals contribute to collective action until the end of the session.

In addition to an individual learning process within each individual, the results clearly show that social interaction between individuals influences their decision making. As we can see in Sessions 2 and 3 of Figure 1, different patterns of behaviour and higher levels of contribution are produced. This shows that providing an option to access social information may give a positive impact to increase collective contribution. A form of social imitation/learning may exist whenever an individual observes others’ behaviour. Both of social influences: the majority influence (the majority’s efforts to produce conformity on the part of a minority) and minority influence (a minority’s effort to convert the majority to its own way of thinking) [6], are expected to exist during the experiment.

When individuals use only their own experiences to make decision, an equilibrium situation may exist when the total contribution is relatively steady round-by-round. Instead of a polymorphic equilibrium, a monomorphic equilibrium is more likely to exist. A monomorphic equilibrium is an equilibrium situation where all members play an identical strategy, with the indication that their behaviours are close to the average behaviour (in this case, average contribution), while a polymorphic equilibrium exists where a fraction of subjects consistently gives high contribution and the other fraction gives less contribution.
Individual contributions in Session 1 provide information that round-to-round most of subjects contribute close to the average contribution, which means that they may have similar strategy. However, there are two subjects who consistently behave in two extremes, giving high and zero contribution respectively. An ANOVA test is conducted on all subjects’ individual contributions, resulting in rejection of the existence of monomorphic equilibrium hypothesis (F-value = 9.012; F-critical\(\alpha=0.1\) = 1.747). However, by excluding these two “extreme” subjects, different result is produced and the hypothesis is accepted (F-value = 1.767; F critical\(\alpha=0.1\) = 1.869). This means that a monomorphic equilibrium is more likely to occur than a polymorphic equilibrium.

Unlike in Session 1 where subjects could only know their own decisions and outputs, in Session 2, social information about those of close colleague(s) is an option for every subject. The experimental results indicate that individuals may access this social information before they make their own decisions. It is revealed that on average 6.9 subjects accessed social information in each round. The fact that social information is provided for free, as it often happens in real life, can be the main reason why most of subjects access the social information.

In Session 3, different form of social information is provided. Each subject knows the decisions of all group members. By knowing all members’ decision, each subject would understand the majority decision in current round which later might affect his or her decision in the following round. In the experiment, the decision that is being the majority changes round-to-round and it creates fluctuation in each individual contribution as well as the system’s total contribution (see Figure 1, Session 3).

A way of measuring the influence of majority (conformity) is by fitting the output of experiment into a model of social influence. Based on Gordon et al. [3]’s model of social influence, we develop a simple model which considers both personal experience and influence of others. The decision of individual \(i\) at round \(t+1\) can be formulated as:

\[
x_{i,t+1} = \beta x_{i,t} + \frac{1}{|\vartheta_i|} \sum_{k \in \vartheta_i} w_{ik} x_{k,t} + H_i + e_i
\]

where \(\beta\) is parameter of the influence of previous experience; \(x_{i,t}\) is choice of individual \(i\) at time \(t\); \(\vartheta_i\) is the set of \(i\)’s neighbours; \(w_{ik}\) is weight given by individual \(i\) to the choice of individual \(k\); \(H_i\) is idiosyncratic term of each individual; and \(e_i\) is error term.

This model is used for analyzing the results of Sessions 2 and 3 in the pilot experiment. Considering difficulties on measuring some variables, assumptions are used to simplify the analysis. It is assumed that there is no idiosyncratic term that would differ any individual from the others (\(H_i = 0\)) and error term is also assumed to zero (\(e_i = 0\)). Similar weight is given by each individual to choices of other individuals (\(w_{ik} = w_i, \forall k \in \vartheta_i\)). The number of neighbours for each individual is the members in her group, excluding herself (\(|\vartheta_i| = N-1, \forall i\)). Based on these assumptions, equation (1) becomes:

\[
x_{i,t+1} = \beta x_{i,t} + \frac{1}{N-1} \sum_{k \in \vartheta_i} w_{ik} x_{k,t}
\]

The contribution of each subject at round \(t+1\) in the experiment (\(x'_{i,t+1}\)) is then being matched with the expected contribution (\(x_{i,t+1}\)) calculated from equation (2). If both are in the same category, whether low contribution (bus-use less than 10 days/month) or high
contribution (bus-use at least 10 days/month), then it is considered as a ‘hit’. The hit ratio will be used to measure the fitness of the model to the experimental data.

Data in Session 3 is used to measure the influence of majority. Parameter $\beta_i$ is derived from the score given each subject to the variable “influence of own experiences” in the post-experiment questionnaire, while parameter $w_i$ is derived from variable “influence of majority”. Expected contribution is calculated using equation (2). Fitting actual contributions in Session 3 with expected contributions results in 67.9% hit ratio (55 out of 81 decisions).

To measure the influence of close colleague(s) to an individual, results obtained from Session 2 of the experiment are fitted with a model of social learning. As the number of other individuals being observed in each round is only one, equation (2) becomes:

$$x_{it+1} = \beta_i x_{it} + w_i x_{kt}; i \neq k$$  (3)

Parameter $\beta_i$ is derived from the score given each subject to the variable “influence of own experiences”, while parameter $w_i$ is from variable “influence of close colleague(s)”. Fitting actual contribution in Session 2 and expected contribution calculated using equation (3) gives hit ratio 69.1% (56 out of 81 decisions).

Minority decision by an altruistic individual, who always gives high contribution even if it is costly and generates low earning for her, may also produce some level of influence to other individuals who are observing her decision. For example, in Session 2, there is an individual who seemingly behaves like that. She has been observed the most often by other subjects. It is likely that her behaviour has had influence on other subjects since the total contribution of the group is relatively higher than total contribution in early rounds of Session 2 as well as in the end of Session 1 (see Figure 1). If it is true, then a positive influence should be expected since she always gives high level of contribution in most of the rounds. This phenomenon gives a sign that minority influence may exist. However, further analysis is needed to confirm the truth of this hypothesis.

6. Further study: agent-based simulation model

An agent-based simulation approach is used with the consideration that it is able to simulate the behaviour of each person individually. The applications of this approach in travel behaviour modelling have shown its promising potential as a new kind of analytical tool (e.g. [4]; [7]). A multi-agent simulation model provides an analytical tool to predict and analyse behaviour in larger and more complex environments that can not be studied in the laboratory. It is able to handle more individuals and more repetitions, and to test the robustness of the laboratory experiment results in different kinds of situation.

The simulation model consists of two layers: an individual layer where each traveller, who is modelled as an agent, collects information (individually and socially), updates its decision rules, and then makes a decision; and an interaction layer where the decisions of all agents are accumulated and used to calculate the system’s outcomes. Decision rules and parameters used by agents are based on the results of the laboratory experiment. Each agent has an opportunity to access social information about the behaviours of other agent in addition to own experience. The information is used to update its strategy, where the use may depend on whether the decision threshold has been exceeded or not. If the threshold is
exceeded then the agent relies on individually learned information. If otherwise, then individual information is disregarded and the agent relies solely in socially learned information. Different decision threshold is assigned to each agent and the value evolves during the simulation following a reinforcement process in which positive outcome would decrease the threshold so that it favours individual learning and vice versa.

It is expected that the model will be useful for predicting the changes in travellers’ behaviour during the implementation of a policy measure. Behavioural changes in a wider scope - a society, are expected to differ from those of a small group, since interdependence between individuals becomes complicated whenever the number of interacting individuals increases. An agent-based simulation model is not only able to show the dynamics of the system but also behavioural change of every single individual. The way in which the compliance to a policy measure spreads from an individual to other individuals and how individual behaviour is changed (or can be changed), can be investigated using the model.

References


Abstract. To make our cities sustainable is one of today’s major challenges. The complexity of this task requires suitable planning tools. Therefore, the analysis and assessment of numerous alternatives and scenarios is necessary. At the same time, with the objective to implement policies on different scales as well as to estimate and understand their impacts, a hierarchical modelling approach is suggested. In this paper, firstly the effectiveness of strategic models as instruments to support decision-making is evaluated. Secondly, the integration with local transport models to evaluate effects in greater detail is investigated.

1. Introduction

Sustainability is one of today's major challenges. Numerous studies provide evidence that cities worldwide do not fulfill the requirements of sustainability. According to May et al. [6], sustainability can be defined as equity between today’s and future generations. But unfortunately, the current patterns of development and travel are not sustainable.

Mobility has increased dramatically over the last years and consequently the average trip lengths. Some of this growth can be attributed to the car ownership, but it also reflects the urban sprawl. In this case, the integration of land use and transport is central to the achievement of sustainable development.

Urban planning can help, for example, to reduce the need to use the car and can foster alternative modes of transport. In conjunction with transport measures (such as pricing, new public transport services, etc.) the most effective overall strategy can be adopted.

---

1 Institute for Transport Planning and Traffic Engineering, Vienna University of Technology, Austria
2 TRANSyT – UPM – Transport Research Center, Technique University of Madrid, Spain
Considerations have to be given to how development decisions can in fact achieve sustainable development objectives.

Among factors that help to establish a suitable framework for the integration of policies and strategies, can be highlighted [3]: (a) to establish a strategic framework with respect to the policy that will be implemented, aids to assure that the individual policies are consistent with the objectives and priorities of superior levels; (b) the decision-makers need more and more better tools that help in the strategies analysis based on indicators that have to be clear, robust and, of course, comprehensive, (c) the mechanism of anticipation, detection and resolution of conflicts in early stages of the political development process helps to identify weakness and to reduce incoherence in its definition and implementation; (d) the decision-making process must be organized to achieve a balance between the policies priorities and limitations of budgets and resources.

Therefore, it is necessary to study those transport policies that better contribute to induce the long term desirables effects on the land use and vice versa. The work presented here investigates whether transport policies can contribute to the high level objective of sustainability, and also how to construct instruments and tools to help in the decision-making process. In this way, we can try to solve two questions:

 vestibule
If in the stage of planning and development of regional mobility policies, is considered the necessity to have instruments like the strategic models that fulfill the following conditions [10]: (a) the expenditure to apply a strategic model is lower as for common four-stage models; (b) the strategic model concept is able to cope with high level of uncertainties (equal or even better than a more detailed model); (c) the very short run time of strategic models allows the simulation of high numbers of different potential economic, demographic, etc. scenarios; (d) the strategic model is able to deal with competing objectives in a multiple player environment. Therefore, strategic models can be a useful support tool for the development of plans or regional mobility policies to meet the challenge of urban sustainable development.

 vestibule
If the strategic models, given their limitations, could complement with another type of instruments or models (for example local models, national models, European models, etc.) in a hierarchical way that allows to interrelate the policies and their effects in each one of vertical levels of action.

Under this perspective, on the one hand, two actions in the Madrid Region (Comunidad de Madrid) were analyzed to answer the first question. On the other hand, to answer the second one, a hierarchical approach was developed linking regional models to local models.

2. The case study area and transport projects

The study area Comunidad de Madrid covers an area of about 8,000 km² with 5.4 million inhabitants, 2.9 million of them in the central core (city of Madrid). As regards current trends, today there is a strong tendency to move activities such as housing, industry and trade to the suburbs. Nevertheless, in spite the current trend, the central area still constitutes the main destination for most daily trips. As a result a high share of people commutes between the outskirts and the core city. Empirical data show that currently as well land use as transport do not fulfill the requirements of sustainability.
As an example, the municipalities Rivas-Vaciamadrid and Arganda del Rey are situated in the A3 corridor in the South-East of Madrid. Due to the relative proximity many of the residents commute into the city of Madrid. As public transport was solely bus based the share of public transport trips was only about 21% in 1996 [1]. The traffic situation was furthermore worsened by a rapid development of housing, service sector business and industry. The municipality of Madrid reacted with the decision to extend the metro line number 9 from its former end station Puerta de Arganda until Arganda del Rey. This extension consists of four new metro stations. A more detailed description can be found in Vieira [11].

Another potential policy instrument is the extension of bus and high occupancy vehicle (HOV) facilities to all radial highways. This suggestion is based on the experience with the existing bus/HOV lane of the highway A6 (a 16 kilometers long stretch from Las Rozas to Moncloa interchange). Prior to the implementation of the bus/HOV facility, the A6 corridor was chronically congested. The opening of the bus/HOV lane in 1995 improved the situation. Peak travel times for bus and HOV lane users have decreased substantially. Bus patronage has increased significantly, from 24% in 1991 to 36% in 2001 [8]. Recently the Spanish minister Magdalena Álvarez presented plans to construct more than 100 kilometers two-way bus lanes on all radial highways [5].

3. Assessment framework

A framework to assess the contribution of these two instruments (extension of metro and bus/HOV lanes) to the objective of sustainability is suggested. It consists of an indicator based approach and a modified cost benefit analysis [7]. The strategic, dynamic land use and transport interaction model MARS –Metropolitan Activity Relocation Simulator– [9] is the core of the assessment framework. Effects on land use, regional travel patterns and transport emissions can be predicted with MARS.

3.1. MARS Model

MARS is an integrated strategic and dynamic land-use and transport model. The basic underlying hypothesis is that settlements and the activities within them are self-organizing systems. Therefore it is sensible to use the principles of system dynamics, synergetics and chaos theory to describe collective behavior [4].

MARS assumes that land-use is not a constant. It is rather part of a dynamic system that is influenced by transport infrastructure. Therefore at the highest level of aggregation MARS can be divided into two main sub-models: the land-use model and the transport model. The interaction process is implemented through time-lagged feedback loops over a period of 30 years.

MARS estimates the effects of several demand and supply-sided instruments whose results can be measured against targets of sustainability. These instruments range from demand-sided measures, such as with public transport fare (increases or decreases), parking or road pricing charges to supply-sided measures such as increased transit service or
capacity changes for road or non-motorized transport. These measures, furthermore, could be applied to various spatial levels and/or to time-of-day periods (peak or off-peak).

To date the model MARS was applied to seven European case study cities: Edinburgh, Helsinki, Leeds, Madrid, Oslo, Stockholm and Vienna. Within the ongoing EU funded research project SPARKLE (Sustainability Planning for Asian cities: making use of Research, Know-how and Lessons from Europe) MARS is adopted and applied to Asian cities [2]. A full description of MARS is given in Pfaffenbichler, 2003 [9].

3.2. Hierarchical approach

The integration could be considered on the basis of two main vectors: a horizontal integration of public policies related to transport, territorial planning, social and economic development, as well as the environmental one; and a vertical vector in which the policies from different administrative and geographical levels (national, regional, local) could be linked. Related to vertical integration, an hierarchical approach is proposed and developed by Vieira [11]. Strategic models do not provide suitable information for local and detailed strategies, or just to evaluate the effects of any measure at a detail level, but they do evaluate transport or urban policies at a regional scale. So, in order to adapt decision support instruments to urban and mobility planning process with different scales, a two levels or action plans must be contemplated: (a) Strategic Level –what actions are undertaken in coherence with the urban or regional mobility models; (b) Detail level –in order to evaluate the effects and impacts of certain policies in a detailed area in coherence with the scenarios of strategic level. Under this hierarchical approach, linking strategic model (MARS) to models covering other levels of dis-aggregation, is proposed. The MARS outputs, settlement patterns and origin-destination matrices, are disaggregated to smaller study areas. Detailed transport models are applied to these areas.

4. Results

A modified cost benefit analysis as laid out by Minken, et al. [7] was used to assess the effects of the transport projects towards the objective of sustainability.

Table 1 summarizes the results of the cost benefit analysis. The effects are assessed for the period 1996 to 2026. The interest rate was estimated with 6%. The result of the scenario “extension of the metro line number 9” is slightly positive. The group, which receives the highest benefits, are public transport users. They gain time savings worth about 220 million Euros which is about 2/3 of all benefits generated. The group, which pays for the strategy, are public transport operators. They bear about 60% of total costs while receiving just about 7% of total benefits. Concerning land use property owners gain additional profit while property users have to bear higher costs. The environmental benefits account for about 5% of the total benefits.

Both bus lane scenarios result in a welfare surplus. But nevertheless there are differences. Both have in common that the positive result is driven by a highly positive value for public transport user time savings. Car user time savings are positive in the scenario “New lanes” and negative in the scenario “Replace Car Lanes”. The same is true
for car user costs. In both scenarios public transport operators create about the same revenues from additional fares. The government finances the investments in both scenarios. The total external costs are negative for the scenario “New Lanes” and positive for the scenario “Replace Car Lanes”.

<table>
<thead>
<tr>
<th>Source of costs and benefits</th>
<th>Line 9 “New Lanes”</th>
<th>“Replace Lanes”</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>User</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Public transport</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time savings</td>
<td>221.5</td>
<td>1,181.8</td>
</tr>
<tr>
<td>Car</td>
<td>-3.5</td>
<td>112.7</td>
</tr>
<tr>
<td>Time savings</td>
<td>4.0</td>
<td>33.3</td>
</tr>
<tr>
<td>Money</td>
<td></td>
<td>-81.2</td>
</tr>
<tr>
<td>Residences</td>
<td>-123.5</td>
<td>-148.0</td>
</tr>
<tr>
<td><strong>Operator</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Public transport</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Investment</td>
<td>-113.3</td>
<td>-722.0</td>
</tr>
<tr>
<td>Operating costs</td>
<td>-77.6</td>
<td>-5.9</td>
</tr>
<tr>
<td>Revenues</td>
<td>22.4</td>
<td>141.4</td>
</tr>
<tr>
<td>Road</td>
<td>-0.4</td>
<td>0.0</td>
</tr>
<tr>
<td>Maintenance</td>
<td></td>
<td>0.0</td>
</tr>
<tr>
<td>Residences</td>
<td>71.1</td>
<td>151.1</td>
</tr>
<tr>
<td>Government</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fuel tax, Parking</td>
<td>-8.5</td>
<td>1.3</td>
</tr>
<tr>
<td>Society (external costs)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accidents, local emissions</td>
<td>10.2</td>
<td>-36.0</td>
</tr>
<tr>
<td>Greenhouse gas emissions</td>
<td>5.2</td>
<td>6.9</td>
</tr>
<tr>
<td><strong>Total Objective Function</strong></td>
<td>16.6</td>
<td>716.6</td>
</tr>
<tr>
<td><strong>Present Value of finance (PVF)</strong></td>
<td>-177.4</td>
<td>-585.2</td>
</tr>
</tbody>
</table>

Table 1: Results of the cost benefit analysis (million Euros)

Finally it can be concluded that all three investments tested create a surplus in welfare. Nevertheless they require higher public spending than in “Do Minimum”. The positive results are in any case driven by the time savings of public transport users. Without time savings all results would be negative. Furthermore the contribution to the overall objective of sustainability is rather small.

5. Conclusions

The results obtained from the evaluation of two recent transport initiatives demonstrate that no single project alone will be able to achieve the goal of a sustainable urban region. A comprehensive strategy including other complementary instruments like pricing is necessary to achieve the objective of sustainability.

From a methodological point of view, it has been shown that the robustness of strategic integrated land use and transport models is a useful and essential tool to support the definition and development of integrated mobility policies and strategies. Nevertheless, associated to the obtained results, it is possible to conclude that, from a sustainability point of view, to evaluate and to choose a certain policy requires to contemplate the effects beyond the economic ones including environmental, social and health effects. This way leads to a horizontal improvement of the models, incorporating new aspects into the assessment. Finally, the hierarchic models seem to be helpful instruments to guarantee coherence in the integration of policies in different spatial scales. In this communication an experience of linking regional and local models was formulated, but others can be
incorporated to this integration. These studies are continued by the research centres: Vienna University of Technology, Institute for Traffic Planning & Traffic Engineering, the Institute for Transport Studies, Leeds and TRANSyT (UPM - Technical University of Madrid) in order to test the robustness of this approach for the transport policy planning process.

References


A CONTRIBUTION TO PREDICTING THE MODAL SPLIT IN URBAN PASSENGER TRANSPORT UNDER INCOMPLETE DATA

Jan BRŮHA* & Hana FOLTÝNOVÁ†

Abstract. In this paper, we propose a novel approach to estimation and prediction of the modal split of passenger urban transport under a form of data incompleteness. Researchers that have an access to micro data in the form of various household expenditures statistics can usually observe expenditures on or demand of various transport modes, including motor fuels. But the demand or expenditures data often do not inform where motor fuels are used. Yet, for some policy analyzes (such as health impacts of transport in cities), the estimation of the split of household usage of motor fuels between urban and non-urban areas is needed. We propose an approach, which can – based on certain assumptions – estimate this split even under the data incompleteness.

1. Introduction

Since emissions and their impacts on the welfare and climate change are considered to be one of the main sources of external costs from transport, various policy measures to decrease the amount of emissions are currently being discussed in the European Union. There is therefore a need for evaluation of impacts of various policy measures on transport behavior and related external costs. To do that, price responsiveness of passenger transport behavior with respect to the relevant measures has to be estimated. Nowadays, researches can use detailed micro databases, such as the household expenditures statistics, and can estimate a detailed demand system over a suitable defined groups of commodities. For example, Brännlund and Nordström [1] estimate a demand system which include petrol expenditures, public and other transport and other commodities on Swedish data. Similarly Brůha and Ščasný [2] estimate a demand system of transport demand for the Czech Republic. It consists of four commodities: motor fuels, urban public transport, rail and inter-city buses plus the composite of the rest of commodities for various household groups (distinguished

*Charles University Environment Center, U kříže 8, Praha 150 00, Czech Republic, jan_bruha@yahoo.co.uk
†Charles University Environment Center, U kříže 8, Praha 150 00, Czech Republic, hana.foltynova@czp.cuni.cz
by size, the living location and social status). Estimated demand systems can be then used to predict changes in the transport demand when relative prices change.

However, such an approach is not sufficient for some purposes. The issue is that health impacts (and therefore the related external costs as well) of fossil motor-fuel usage differ significantly between urban and rural areas. Thus to assess effectiveness of policy measures, one has to estimate separate motor-fuels demand functions for driving in urban and rural areas. This is a difficult task, since the usually available databases contain data on the total expenditures on motor fuels, but they directly do not provide information where the motor fuels are used, whether in cities or in rural areas. Clearly, it is unsatisfactory to assume that the urban and rural price elasticities of the motor-fuel usage are similar: substitution possibilities as well as social and economic status of ‘typical’ households differ in cities and in rural areas.

The purpose of this paper is to propose an econometrically sound method to estimate area-differentiated price responsiveness under the condition that the usage of motor fuels in the two areas is unobservable for the researcher. We show that if the researcher is willing to make some assumptions it will be possible to do that.

2. The proposed methodology

As discussed above, the researcher typically does not observe the shares of motor fuels used in urban versus rural areas. Thus to identify the price responsiveness of commuting in the two areas separately, one has to rely on identifying assumptions.

We propose a demand system, which accommodates such assumptions. It is constructed as a two-stage budgeting process. In the first stage, the households decide shares of expenditures on urban commuting, on transport outside cities (inter-city transport and rural commuting) and on the other goods. This is done using the almost ideal demand system by Deaton and Muellbauer [4]. In the second stage, two other almost ideal demand systems are used again: the first one to determine shares of public urban transport and motor-fuels expenditures on the total urban commuting expenditures and the other one to determine shares of public transport and motor-fuels expenditures on the total non-urban commuting expenditures. The level-budgeting used is precisely the instrument, which enables us to estimate the division of motor-fuels used for city and inter-city transport without having it in data. The presented demand system is our methodological contribution.

First, we assume that each household allocates desired expenditures among expenditures on urban transport, inter-urban transport and other goods. The share of expenditures on urban transport (which include public urban transport and motor-fuels expenditures purposed to urban commuting) shall be denoted by $\omega_M$, and the share of expenditures on inter-urban transport (which include expenditures on rail and buses and motor-fuels expenditures purposed to non-urban commuting) by $\omega_V$.

Although these two shares are unobservable to the researcher, we assume that the shares satisfy the Almost Ideal Demand (henceforth AID) system restrictions:

$$\omega_M = \alpha_M(h, \epsilon) + \gamma_M \log \left( \frac{P_M}{P_0} \right) + \gamma_{MV} \log \left( \frac{P_V}{P_0} \right) + \beta_M \log \left( \frac{X}{P} \right),$$ (1)
\[
\omega_V = \alpha_V(h, \epsilon) + \gamma_{MV} \log \left( \frac{P_M}{P_0} \right) + \gamma_V \log \left( \frac{P_V}{P_0} \right) + \beta_V \log \left( \frac{X}{P} \right),
\]  

(2)

where \( P_M \) is the Stone price index for urban transport (defined in more details below), \( P_V \) is the analogous Stone price index for non-urban transport (also defined in more details below), \( P_0 \) is the price index for other goods approximated by the CPI, \( X \) is the total expenditure, \( \alpha_M(h, \epsilon) \), \( \alpha_V(h, \epsilon) \) are intercepts, which depend on household characteristics \( h \) and a random-effect term \( \epsilon; \gamma_{MM}, \gamma_{MV}, \gamma_{VV}, \beta_M, \beta_V \) are the rest of parameters, and \( P \) is the Stone price index rewritten as:

\[
\log P = \alpha_0 + \alpha_M(h, \epsilon) \log \left( \frac{P_M}{P_0} \right) + \alpha_V(h, \epsilon) \log \left( \frac{P_V}{P_0} \right) + \log P_0 + \ldots
\]  

(3)

\[
+ \frac{1}{2} \gamma_{MM} \log^2 \left( \frac{P_M}{P_0} \right) + \gamma_{MV} \log \left( \frac{P_M}{P_0} \right) \log \left( \frac{P_V}{P_0} \right) + \frac{1}{2} \log^2 \left( \frac{P_V}{P_0} \right).
\]

We explore parametric restriction of the AID system, and writes the system in a compact form (1) – (2), deleting the redundant third equation for the other-goods demand. Also Equation (3) already accommodates the parametric restrictions.

After deciding expenditures on urban, non-urban transport and other goods, the household decides the modal split of the two composite transport categories.

The expenditures on urban transport are split on urban public transport, share of which is denoted as \( \omega_H \), and motor-fuel expenditures for urban commuting with the share \( \omega_{MF} \equiv 1 - \omega_H \). These shares satisfy another AID system:

\[
\omega_H = \alpha_H(h, \epsilon) + \gamma_H \log \left( \frac{P_H}{P_F} \right) + \beta_H \log \left( \frac{\omega_M X}{P_M} \right),
\]  

(4)

where \( P_H \) is the price of public urban transport, \( P_F \) is the consumer price of fuel, \( \alpha_H(h, \epsilon), \gamma_H, \beta_H \) are parameters (again the intercept depends on household characteristics and a random-effect term) and \( P_M \) is the Stone price index for urban transport, derived from the AID specification:

\[
\log P_M = \alpha_{0M} + \alpha_H(h, \epsilon) \log \left( \frac{P_H}{P_F} \right) + \log P_F + \frac{1}{2} \gamma_H \log^2 \left( \frac{P_H}{P_F} \right).
\]  

(5)

This index enters System (1) – (2).

Similarly, the expenditures on non-urban transport are split on inter-urban transport using buses and rail, share of which is denoted as \( \omega_B \), and motor-fuel expenditures for non-urban commuting with the share \( \omega_{VF} \equiv 1 - \omega_B \). These shares satisfy yet another AID system:

\[
\omega_B = \alpha_B(h, \epsilon) + \gamma_B \log \left( \frac{P_B}{P_F} \right) + \beta_B \log \left( \frac{\omega_V X}{P_V} \right),
\]  

(6)

where \( P_B \) is the price index of bus and rail inter-urban transport, \( P_F \) is the consumer price of fuel, \( \alpha_B(h, \epsilon), \gamma_B, \beta_B \) are parameters (yet again the intercept depends on household characteristics and a random-effect term) and \( P_V \) is the Stone price index for inter-urban transport, derived from the AID specification:

\[
\log P_V = \alpha_{0V} + \alpha_B(h, \epsilon) \log \left( \frac{P_B}{P_F} \right) + \log P_F + \frac{1}{2} \gamma_B \log^2 \left( \frac{P_B}{P_F} \right),
\]  

(7)
which again enters System (1) – (2).

Thus our system is characterized by the stage budgeting, where e.g. expenditures on public urban transport on the urban-transport expenditures $\omega_H$ are independent on say $\omega_B$, once the shares of greater commodities $\omega_M, \omega_V$ are determined. Precisely, this ‘independence’ assumption identifies the model.

The researcher is assumed to observe prices indices (the CPI, the price index of inter-urban transport and the price of motor fuels), household characteristics $h$ and expenditures $X$, and three combinations of expenditures shares on total expenditures, namely the share of public urban transport expenditures $\omega_H\omega_M$, the share of public non-urban transport $\omega_B\omega_V$, and the motor-fuel share $\omega_{VF} + \omega_{MF}$. We do not observe $\omega_{VF}, \omega_{MF}$ individually, nor do we observe $\omega_M, \omega_V$.

Denote observable shares as follows: the share of public urban transport expenditures as $\sigma_H$, the share of public non-urban transport $\sigma_B$ and the motor-fuel share $\sigma_V$. If the relevant Stone price indexes are observable (or if we approximate them as in [1, 6]), it will be possible to write down a linear Seemingly Unrelated Regression (henceforth SUR) system, with the three observable shares as left-hand side variables and cross-products of relative prices and real expenditures as right-hand side variables. The AID system imposes a number of parametric restrictions on the SUR. This will be a reduced form for the structural model (1), (2), (4), (6), which is actually over-identified by the parametric restrictions on the SUR.

If one computes the Stone indexes in a model-consistent way, then one will obtain a non-linear SUR as a reduced form for the structural model (1), (2), (4), (6). The constraints (3), (5), (7) then increase the over-identification of the structural model.

2.1. Econometric Implementation

Since demand functions derived from an almost ideal demand system can be aggregated, the proposed methodology can be used both on micro- as well as macro-data. The usage of microdata (such as an household expenditures survey) increases efficiency, but if these data are not available, the researcher can use aggregated data in the form of expenditures shares aggregated over suitable classes of households. Naturally, these classes can be defined on a geographical basis. Nevertheless intercepts of regression equations used in the proposed model may include also public-transport characteristics related to residents’ cities, especially various measures of density. Thus it is possible to test the well-known Mohring effect, [5].

There are two possible approaches to estimate the model: either to use a Gibbs sampling, which is known to be a suitable model with unobservable states in the Bayesian paradigm, or to use a ‘classical’ minimum-distance estimator. We use the second option, leaving the former option for an ongoing research.

We applied the proposed model for a selection of Czech cities with available data on public transport. The data span is 2000-2004. We impose parametric restrictions on the reduced form model and apply the minimum-distance estimator. The standard errors are approximated by bootstrapping the sample.

Our main finding are following: the own-price motor fuel elasticities significantly differ according to the usage. The point estimates of the elasticity for urban commuting is -1.04
(up to 2 decimal points), while the elasticity for the non-urban commuting is estimated to -0.40. These two figures are significantly different. This suggests that estimation, which ignores the difference between commuting purposes, would likely underestimate the price effect in cities but overestimate it in the rural areas. This may bias efficiency conclusions of policy instruments assessment.

3. Conclusion

This note introduces an econometric model of estimation and prediction of motor-fuels demand under a form of data incompleteness. It is showed that the proposed model is identified, that it is consistent with the neoclassical microeconomic theory and we propose a consistent estimator, which is applied on Czech data. The results of the estimation serve for calibration of the demand side of a Czech micro-simulation model of transport regulation.

The plans for future research is following: first, to explore a possibility of computationally intensive Bayesian approaches (such as Monte Carlo Markov Chain algorithms) to estimate the model. Second, we plan to investigate whether the stage budgeting can identify the model even for semi-nonparametric demand systems (such as models based on Bernstein polynomials [3]), thus relaxing the parametric assumptions.

References


COST-BENEFIT ANALYSIS OF CYCLING INFRASTRUCTURE: A CASE STUDY OF PILSEN

Hana FOLTÝNOVÁ1 and Markéta BRAUN KOHLOVÁ2

Abstract. The paper analyses impacts of improved cycling infrastructure on demand for this means of transport. We use a stated preferences design for the elicitation of willingness to use the bicycle in the event of various improvements to the cycling environment in the city (in strict and tolerant level). In the CBA applied to the planned cycling infrastructure network in Pilsen we include the following benefits: i) improvements in health by regular physical activity of new cyclists (quantification of impacts is based on costs of illness); ii) changes in number and severity of accidents (based on accident costs); iii) changes in atmospheric pollution (using the ExternE data). When the demand change is calculated according to the strict level, the social benefits do not cover social costs of building the new cycling infrastructure.

1. Introduction

The share of cycling in the total modal split is relatively low in most medium-sized and large cities in the Czech Republic. It is about 0.5% in Prague and 0.3% in Pilsen compared to e.g. Munich (13% in 2002, see [12]) or Vienna (4.5% in 2001, [12]). To make cycling more attractive, sums invested in cycling infrastructure gradually rise in the Czech Republic, even if they still only make up a tiny part of the public funds expenditures (about 0.2% of total expenditures from the State Transport Infrastructure Fund). Logically, the question of social benefits of these investments and their economic efficiency appears. Extensive literature exists dealing with CBA of transportation projects, but substantially less deals with that of cycling projects. As Elvik [5] summarizes, CBA of measures for pedestrians and cyclists should apply the same methodology that is used for transport projects in general. However, the specific impacts to be considered will not be the same as in projects that mainly benefit motorised travel. The essential question to be answered before starting CBA is: what is the potential demand for a new cycle-way network in the Czech cities and what increase in the number of cyclists and share of kilometres and time ridden by bicycle can be expected in case cycling facilities are improved and expanded. We aim to answer the question on the case of the Czech city of Pilsen using individual data from a transport behaviour survey.

1 Charles University Environment Centre in Prague, U Krize 8, 158 00 Prague 5, hana.foltynova@czp.cuni.cz
2 Charles University Environment Centre in Prague, U Krize 8, 158 00 Prague 5, marketa.braun.kohlova@czp.cuni.cz
There are only a few studies focused on estimation of demand for cycling facilities from individual data, e.g. [8] or [6]; in addition, they usually neglect the variability in purposes of individual journeys. Moreover, the studies of transport demand are often limited to a short segment of cycling infrastructure [6]. The focus of our study is the whole cycling network in a particular city, similarly [8].

The paper is structured in the following way: CBA methodology and literature review (Chapter 2); demand estimate for the city of Pilsen (Chapter 3); cost-benefit analysis for the planned infrastructure in Pilsen (Chapter 4); sensitivity analysis (Chapter 5). The final chapter concludes.

2. Demand estimate

Our approach to the demand estimate is based on the current transportation behaviour of the target population (18+). Using a questionnaire, we asked respondents about trips made during the previous working day (Tuesday, Wednesday, and Thursday). We focused on each trip on the given working day separately.

The data used come from a standardized questionnaire survey carried out in June and July 2005 in the Czech city of Pilsen (N=763). A quota sample (residential area, age, gender, education level) was used.

The daily journeys including purpose, distance, duration and the means of transport as described in the interviews were taken for reference values for the state of demand before the change.

The demand change itself is estimated from stated preferences of the following wording: “How significantly would the following changes in transport situation in Pilsen influence your willingness to use the BICYCLE more than currently?” Stated preferences are confronted with mutual fulfilment of current preconditions for bicycle use such as bicycle ownership/accessibility, at least experimental or recreational use of the bicycle, perception of bicycle use as an alternative (revealed preference) corresponding to the stages in the process of travel behaviour change as identified in [11].

The stated willingness to use the bicycle more often, measured on a 5-point scale, for a certain purpose is used for the potential demand estimate in two different levels. The tolerant level (Level 1) derives estimated demand from the positive stated willingness (two most positive answers on the 5-point scale) only. The strict level (Level 2) requires other preconditions as well.

Consequently, two different demand scenarios are estimated for each level and each purpose. The change in demand on a given working day is estimated as (1) a change in the number and/or share of travellers using the bicycle; (2) a change in the number and/or share of kilometres ridden by bicycle; and (3) a change in the amount of time spent on the bicycle and/or its share. The switch from car use to the bicycle is estimated separately.

For the estimate of potential demand, the residence, the purposes of the trips and their locations are fixed. We suppose that the time spent travelling by those who do not switch stays constant. The changes in lengths and durations of the trips are estimated using the CUBE model.
The CBA results are calculated for the following demand change (level 2 – strict) regarding cycling as a means of transport: 1) increase in the number of cyclists commuting from 6% to 8.3% and 2) increase in the number of kilometres ridden by bicycle from 8% to 9.9%.

3. Cost-benefit analysis

The demand estimates reflect demand for the same adult population (18+) on a working day in the same period of the year in which the infrastructure is improved. To calculate the change in demand for the entire year, we assume that people in Central European geographical conditions cycle only six months in a year (April to September). We assess benefits connected to two different scenarios of demand intensity: (1) the willingness to cycle on every convenient working day (neutral scenario); (2) the willingness to cycle on every second convenient day (conservative scenario). For the sake of the simplicity of the presentation, the benefits for the strict level (both the neutral and conservative scenarios) are displayed first. The results for the tolerant level are shown later in Chapter 4.

The costs side of the CBA includes infrastructure investments and maintenance costs. The table below shows the existing and planned networks of cycling infrastructure and estimated costs of their construction. The costs of infrastructure construction vary between 1,000 and 2,000 CZK per square metre [10]. Such a difference in costs per square metre is caused by a broad range of material being used, the terrain conditions, etc. For the further analysis we use the average value of 1,500 CZK. The results of the CBA for the lower (1,000 CZK) and the higher (2,000 CZK) values of the infrastructure construction costs are reported in the sensitivity analysis.

<table>
<thead>
<tr>
<th></th>
<th>Cycle path</th>
<th>Cycle lane</th>
<th>Combined walking + cycling path</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Planned network for cycling (km)</td>
<td>34.2</td>
<td>39.2</td>
<td>52.2</td>
<td>125.5*</td>
</tr>
<tr>
<td>Remaining parts of the network (km)</td>
<td>26</td>
<td>19</td>
<td>34</td>
<td>78</td>
</tr>
<tr>
<td>Cost estimates for completing the network (mil. CZK / mil. Euro**)</td>
<td>51.8</td>
<td>18.6</td>
<td>50.4</td>
<td>120.8 / 4.2</td>
</tr>
<tr>
<td>Maintenance costs per year (mil. CZK / mil. Euro)</td>
<td></td>
<td></td>
<td>21.6</td>
<td>0.7</td>
</tr>
<tr>
<td><strong>Total costs</strong> (mil. CZK / mil. Euro)</td>
<td>111.9</td>
<td>76.4</td>
<td>136.2</td>
<td>345.9 / 11.9</td>
</tr>
</tbody>
</table>

* calculated construction costs 1,500 CZK/m²
** the exchange rate is 29 CZK/Euro

In the CBA applied to the planned cycling infrastructure network in Pilsen we include the following benefits:

i) improvements in health by regular physical activity of new cyclists (quantification of impacts is based on costs of illness);
ii) changes in number and severity of accidents (based on accident costs);

iii) changes in atmospheric pollution (using the ExternE methodology for impact quantification);

iv) benefits from reduced insecurity;

v) changes in travel times.

**Improvements in health by regular physical activity of new cyclists.** Because there are no reliable figures for the Czech Republic concerning impacts of regular physical activity on mortality, we assume a 9% decrease in mortality by cardiovascular diseases as [3] did. The value of a statistical life (VSL) for the Czech Republic used is 18.52 mil. CZK [1].

To calculate benefits of improved health from regular cycling (morbidity), we estimate the cost of illness using the prevalence approach (costs connected to an existing case during the assigned period). The benefits from improved health are calculated only for new cyclists regularly cycling to work. The reason is a higher probability of regular everyday trips by bicycle.

First, we focus on the coronary heart disease. We use the 50% reduction in the risk of coronary heart disease [13]. The costs are calculated separately for in-patient and out-patient treatment. The value of social costs includes treatment, drugs and technical treatment, and the loss of productivity.

According to [13], there is a strong evidence of a relationship between physical activity and colon cancer (an average risk reduction of 40-50%). The social costs are again estimated separately for in-patient and out-patient treatment.

Physical inactivity is a major risk factor for the development of type 2 diabetes and increases the risk of its development by 33-50% [3]. Because of unavailability of data specifying the loss of productivity, this item is not included.

**Reduction in the number of accidents involving cyclists.** According to [5], a review of evaluation studies of impacts of separated crossings indicates that the number of pedestrian accidents is reduced by about 80% and the number of accidents involving motor vehicles only is reduced by about 10%.

No comparable figures for accident reduction connected to infrastructure improvement are available for the Czech Republic. Even if there is no proven evidence for the risk of accidents related to cycling, we assume a 10% reduction in accidents involving cyclists and no change in accidents involving motor vehicles only. To calculate impacts of the injuries, we use the value of 200 thousand CZK per light injury (as suggested by [7]) and the accident statistics of 2005.

**Reduced external costs of motorized road transport connected to air pollution.** Atmospheric pollution (caused also by emissions from transport) has an inauspicious impact on human health (e.g. on respiratory diseases, cancer, and premature deaths). It is also a cause of material damages on buildings and plants. At the regional level, pollution causes acidification and globally it is a contribution to the greenhouse effect.

We apply the ExternE data (for more see [4]). The value of external costs of atmospheric pollution includes emissions of NOx, SO2, carbohydrates, particular matters (PM10) and their impacts on human health and early death and CO2. The structure of the vehicle fleet in Pilsen was derived from [2] summarizing data from a vehicle census done in Pilsen in 2001.

**Costs of travel time.** We assume that cycling on a cycle track could reduce travel times by comparison with cycling on an ordinary sidewalk only negligibly. We - similarly to [9] -
assume that travel times for the already cycling will stay unchanged. Because congestion is not a significant problem in Pilsen, we assume that even travel times for car drivers who do not substitute cycling for driving stay unchanged too.

It should be pointed out that travel times increase for those who make the shift from the car to the bicycle (by about 21 minutes for an average trip made by bicycle instead of by car). If an individual declares the willingness to change the mode of transport even if it would lead to a longer travel time, however, we can assume that the individual’s benefit outweighs the private costs connected to this choice (for example as an increase in travel time). That is why this private negative benefit does not represent a social cost.

**Benefits from reduced insecurity.** These benefits are included in the ‘ideally designed’ CBA by [5] and also in the CBA done by [9]. Both the authors distinguished between the reduced insecurity of those who already cycle and those who do not. Nevertheless, we do not include these benefits in the CBA. Similarly to the costs of travel time, this benefit is already internalized in the personal benefits of each cyclist.

### 4. Results

The CBA results are calculated for the following demand change regarding cycling as a means of transport:
- increase in the number of cyclists commuting from 6% to 8.3%, and
- increase in the number of kilometres ridden by bicycle from 8% to 9.9%

The following adjustments are made:
- the present values of benefits are calculated using a discount rate of 7%,
- and a 25-year lifetime of the project (as for example in [9]).

The costs and benefits of the partial CBA analysis for level 2 (strict willingness to change the current means of transport) are summarized in the following table.

<table>
<thead>
<tr>
<th>Benefit and cost components</th>
<th>Impacts per year</th>
<th>Neutral scenario</th>
<th>Conservative scenario</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benefits of cycling infrastructure (present value)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Changes in health</td>
<td>20 persons</td>
<td>8,596.62</td>
<td>Assumed as</td>
</tr>
<tr>
<td>Accidents</td>
<td>4 accidents</td>
<td>9,533.10</td>
<td>14,299.65</td>
</tr>
<tr>
<td>Mortality</td>
<td>0.57 persons</td>
<td>122,201.78</td>
<td>61,100.89</td>
</tr>
<tr>
<td>Emissions</td>
<td>122,000 km / day</td>
<td>22.93</td>
<td>9.94</td>
</tr>
<tr>
<td>TOTAL BENEFITS</td>
<td>140,354.43</td>
<td>75,410.47</td>
<td></td>
</tr>
<tr>
<td>Costs of cycling infrastructure (present value)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Capital costs</td>
<td>78 km</td>
<td>181,200.00</td>
<td>181,200.00</td>
</tr>
<tr>
<td>Maintenance costs</td>
<td></td>
<td>8,053.33</td>
<td>8,053.33</td>
</tr>
<tr>
<td>Tax-cost factor (20%)</td>
<td>78 km</td>
<td>37,850.67</td>
<td>3,785.67</td>
</tr>
<tr>
<td>TOTAL COSTS</td>
<td></td>
<td>227,104.00</td>
<td>227,104.00</td>
</tr>
<tr>
<td>Net benefit/costs ratio</td>
<td></td>
<td>-0.62</td>
<td>-0.33</td>
</tr>
</tbody>
</table>
It can be said that when the demand change is calculated according to the strict level, the social benefits do not cover social costs of building the new cycling infrastructure. The net benefit/cost ratio is -0.62 for the neutral scenario (i.e., 168 cycling days a year), and -0.33 for the conservative scenario (i.e., 84 cycling days a year).

5. Sensitivity analysis

There are many factors which influence the results of the cost-benefit analysis substantially. The sensitivity analysis includes the effects of uncertainties present in the applied procedure: (1) the construction costs; (2) the estimate of change in cycling demand (strict vs. tolerant levels); (3) the change in the number of accidents; and (4) the discount rate and lifetime of the project.

Firstly, the costs of construction vary between 1,000 CZK and 2,000 CZK. Using the lower costs of construction (1,000 CZK per m²), the net benefit/costs ratio is -0.91, while the higher costs of construction (2,000 CZK) only yield -0.47 for the neutral scenario.

Secondly, the level of change in demand influences the results of the CBA substantially. When using the tolerant level of demand, the net benefit/costs ratio is 3.08 for the neutral scenario of demand and the lower costs of construction, and 1.59 for this demand and the higher construction costs (2,000 CZK per m²). Using the tolerant level of demand, the present value of benefits always overweighs the costs with the exception of the conservative scenario for the higher costs.

Thirdly, the impact of the improved cycling infrastructure on the safety of cyclists (the number and severity of accidents) also influences the results of the CBA. When the number of accidents decreases by 25%, the costs equal benefits for the low construction costs level.
and strict level of demand. Nevertheless, the benefits never exceed the costs when assuming the average rate of construction costs (1,500 CZK) for the strict level of demand. Fourthly, other important factors influencing the results of the CBA are the discount rate. The change in distribution of benefits during the lifetime of the project can make the project socially profitable. For example when using a discount rate of 5%, the net benefit/costs ratio reaches 1.1 for the lower construction costs estimate for the strict demand level. When assuming a very low discount rate (nearly zero), the project is profitable even for the higher construction costs estimate.

6. Conclusions

This paper presents the first systematic attempt to calculate the costs and benefits connected to the construction and improvement of cycling infrastructure in the Czech Republic. For this purpose, we analyzed the cycling network construction in Pilsen. The estimated change in demand is relatively small: an increase in persons cycling from 11.6% to 14.2% (strict level) and to 20.9% (tolerant level) for all the regular trips, and from 6.0% to 8.3% and to 14.3%, respectively, for commuting. The following benefits were included: improvements in health by regular physical activity of the new cyclists; changes in the number and severity of accidents; and changes in atmospheric pollution. The impacts on health (mortality and morbidity) hold the major share.

When the demand change is calculated according to the strict level, the social benefits do not cover social costs of building new cycling infrastructure. The net benefit/cost ratio is -0.97 for the neutral scenario and -0.52 for the conservative scenario. It should be mentioned that other possible benefits such as noise reduction and further health impacts are not included.

Still, the results are very sensitive to a range of factors, which can influence the results of the CBA substantially. Above all, the demand estimate plays an important role. Even if the net benefit/costs ratio is -0.91 for the strict level, it is already 3.08 for the tolerant level. These results differ when the higher and average costs of construction are used. Using the higher construction costs (2,000 CZK per m²), the costs do not overweigh the benefits even when the tolerant level of demand is applied. The discount rate is another important factor. When the future benefits are given high priority (using a discount rate close to zero), the project is profitable even for the strict level of demand and higher construction costs.

References


**Acknowledgements.** The research has been supported by the grant of the Ministry of Transport No. 1F43E/045/210 „CYCLE 21 – Analysis of Demand for building of cycling infrastructure” and Operation Program Infrastructure. The support is gratefully acknowledged.
Predicting a Suitable Capacity Estimation Method for Single-Lane Roundabouts in İzmir, Turkey

Serhan TANYEL¹, Mustafa ÖZUYSAL², Pelin ÇALIŞKANELLİ³, Türkay BARAN⁴

Abstract. The capacity of a roundabout can be calculated by using two different methods: gap-acceptance theory and regression analysis. In this study, the capacity of roundabouts in İzmir, Turkey is tried to be determined by using regression analysis. For this purpose, the data are obtained from six approaches of four roundabouts in İzmir, Turkey. The results have shown that, entry width, entry radius and conflict angle have a strong effect on roundabout entry capacity. On the other hand it is found that inscribed diameter has no effect on capacity, which is quite unexpected.

1. Introduction

Today, the capacity of roundabouts can be determined by using two different calculation methods:

1. The theoretical approach or gap-acceptance theory
2. Empirical models mostly depend on regression analysis.

In gap acceptance method, it is assumed that drivers in the minor-stream approach of a roundabout decide to join the major (circular stream) stream according to the size of gaps between successive vehicles in the major stream [1] [2] [3].

On the other hand, some scientists stated that gap acceptance theory couldn’t be easily applied to roundabout entries. The reasons can be listed as follows: firstly in some cases move-up times turned out to be greater or equal to the critical gap and secondly it was nearly impossible to define the correct major streams at multi-lane roundabouts with multi-lane entries [4]. Kimber [5] was one of the leading researchers who used regression analysis in prediction of roundabout capacity and performance. This model is known as the TRL method.

¹ Faculty of Engineering, Dokuz Eylül University, İzmir, 35160, Turkey, serhan.tanyel@deu.edu.tr
² Faculty of Engineering, Dokuz Eylül University, İzmir, 35160, Turkey, mustafa.ozuysal@deu.edu.tr
³ Faculty of Engineering, Dokuz Eylül University, İzmir, 35160, Turkey, pelin.caliskanelli@deu.edu.tr
⁴ Faculty of Engineering, Dokuz Eylül University, İzmir, 35160, Turkey, turkay.baran@deu.edu.tr
The following objections for gap-acceptance theory have been stated by Kimber [3]:

a) The follow-up headway (and critical gap) should be dependent on circulating flows (not a constant value),

b) Gap acceptance parameters would need to be related to geometric layout,

c) The gap acceptance process, do not apply at high circulating flows due to priority sharing and gap reversal.

Troutbeck [6] has stated that, the best way to estimate the capacity is the direct measurement of the average of the maximum throughput. Flows for a number of 1 min. and 5 min. periods should be used to calibrate the linear regression equations. Thus this cannot be possible in most cases. In order to obtain a linear regression equation, there should be a constant queue of vehicles which’s drivers are waiting a suitable gap to enter the roundabout. It is quite difficult to find a roundabout like this, especially in Turkey. For the reasons stated above gap-acceptance theory is preferred by many researchers.

In this study, applicability of regression analysis methods to roundabout capacity analysis is investigated. The results are based on the data obtained from six single-lane roundabouts in İzmir, Turkey.

2. Study Sites and Data Collection

In the study, the required data was obtained from the observations made at six approaches of four roundabouts in İzmir, Turkey. Two of the roundabouts are located in Karşıyaka (Bostanlı and Soğukkanlı), one is located in Bornova (Ege Rektörlük) and one is located at Alsancak (Gündoğdu).

Required traffic data is collected by using video cameras. Cameras are placed at points like nearby houses or workplaces at which the roundabout approach/approaches can be seen clearly. The video recordings have been made at evening peak hours; on clear days under dry road surface condition for 3 hours in order to have at least 30 minute periods at which there is a queue at the approach entries. Thus, a successive 238 minute recording can be obtained for all intersections as a total, which are used in regression analysis.

3. A New Capacity Model for Roundabouts in İzmir, Turkey

As stated before, nearly in all of the studies on roundabouts in Turkey, methods which depend on gap acceptance theory are used [7] [8] [9] [10] [11] [12] [13]. In this study applicability of regression analysis techniques in capacity analysis is investigated. For this purpose, the relation between circulating flow ($Q_c$), entry capacity ($Q_e$) and geometric parameters is analyzed.

As a first step, the relation between observed entry capacity and circulating flow is investigated (Figure 1). As can be seen in Figure 1, the result of linear regression is acceptable but by using geometric parameters a better solution may be obtained.
For the second step, all of the geometric parameters are included in multiple regression analysis. The results are given in Table 1. When Table 1 is investigated carefully it can be seen that the “r” and “R^2” values are quite high. However t-statistic value of inscribed diameter is lower and and P-value of inscribed diameter is higher than other parameters’ t-statistic and P-values. This means that, inscribed diameter has no effect on capacity of single-lane roundabouts. This is quite unexpected. Inscribed diameter is an important factor in TRL method and it is also used in Troutbeck [2] and SIDRA [3] methods for predicting follow-up times and critical gap acceptance values.

On the other hand it is also clear that, the entry capacity of a roundabout is highly related with circulating flow. It can be also said that entry capacity, entry radius and conflict angle have also a great effect.

\[
y = -1,1093x + 1417,1
\]
\[
R^2 = 0,6196
\]

![Relation between observed capacity and circulating flow](image)

**Figure 1. Relation between observed capacity and circulating flow**

<table>
<thead>
<tr>
<th>Coefficients</th>
<th>Standard Error</th>
<th>t Stat</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>567.05</td>
<td>210.56</td>
<td>2.69</td>
</tr>
<tr>
<td>(D_i)</td>
<td>5.52</td>
<td>6.46</td>
<td>0.86</td>
</tr>
<tr>
<td>(w_e)</td>
<td>60.81</td>
<td>14.44</td>
<td>4.21</td>
</tr>
<tr>
<td>(r_e)</td>
<td>-4.59</td>
<td>0.65</td>
<td>-7.10</td>
</tr>
<tr>
<td>(\phi)</td>
<td>21.55</td>
<td>2.65</td>
<td>8.13</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>R^2</th>
<th>0.906</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adjusted R^2</td>
<td>0.904</td>
</tr>
<tr>
<td>Standard Error</td>
<td>81.71</td>
</tr>
<tr>
<td>Observations</td>
<td>238</td>
</tr>
</tbody>
</table>

**Table 1. Results of first multiple regression analysis**
As the third step, multiple regression analysis is performed without including inscribed diameter. The results are given Table 2. When Table 1 and Table 2 are compared it is clear that although inscribed diameter is not included in multiple regression analysis, no change in “r” and “R²” values is observed moreover standard error is decreased. The intercept is increased but the coefficients of entry width and conflict angle are decreased. The coefficient of entry radius is also decreased slightly but there is no change in the coefficient of circulating flow.

![Table 2. Results of second multiple regression analysis](image)

The results of the new model obtained from multiple regression analysis are compared with the observed capacity values (Figure 2). In order to have a better idea on the accuracy of the model, a linear regression analysis is performed and the intercept is assumed as zero. From the figure it is clear that model gives close results to observed capacity values.

![Figure 2. Comparison of observed and predicted capacity values](image)
4. Conclusions and Recommendations

In this study, applicability of regression analysis methods in calculation of entry capacity of single-lane roundabouts in İzmir, Turkey is investigated. The results have shown that entry width, entry radius and conflict angle have a strong effect on roundabout entry capacity. On the other hand it found that inscribed diameter has no effect on capacity, which is quite unexpected.

The results of this study are limited with the data obtained from six approaches of four roundabouts. The model needs to be improved by increasing the number of observed roundabouts from all around Turkey. Especially the effect of inscribed diameter should be examined in detail.

References


A METHODOLOGICAL APPROACH TO OPTIMIZE TEMPORAL INTEGRATION OF REGIONAL PUBLIC TRANSIT LINES

Domenico GATTUSO¹, Antonella POLIMENI²

Abstract. Given the importance of reducing users’ disutilities at an interchange node, in this work both a new methodological approach and an analytic formulation for generating synchronized transit timetables in an extra-urban context (low frequency) are proposed.

1. Introduction

One of the most evident differences between private and public transport is the temporal and space availability of service. In fact, while in the first case, after deciding the departure time of their trip, users can enjoy a “door-to-door” service, in the second case, not only is the service available only at specific scheduled times, but it is also very expensive, especially over large areas, to connect directly each origin-destination pair with a single bus line. For this reason, passengers often have to use an intermodal transit service, consisting of main and feeder lines connecting at different interchange nodes and, therefore, they may need one or more transfers to complete their journey.

Obviously, the longer the transfer time, the higher the public transport disutility value. This involves a more limited use of public transport, in favour of individual transport.

For this reason, the integration process of different transit services is considered one of the most useful operational tools in order to contain the undisputed predominance of individual transport and to promote a modal redistribution in favour of public transport.

In fact, synchronizing the timetables of one or more transit services can contribute to reduce the waste of time connected with transfer operations between services operating at the same interchange point (connection node), thus controlling the disutility values perceived by users.

¹ Mediterranea University of Reggio Calabria (IT), gattuso@ing.unirc.it
² Mediterranea University of Reggio Calabria (IT), antonella.polimeni@ing.unirc.it
2. **State of the art**

Nowadays, the issues related to the temporal integration process are treated in a marginal way, and the methodological and modelling literary contributions are quite meagre (Ceder, 1986, 2001; Ceder and Tal, 2001; Chowdhury, 2000; Chowdhury and Chien, 2001; Coppola, 2005; Dessouky et alii, 2002; Fleurent et alii, 2004; Nuzzolo et alii, 2003; Russo, 1998; Wong and Leung, 2004).

Some of these studies analyze the problem from a static point of view, by considering that the schedule depends on transport demand. Others coordinate the timetables in relation to the real traffic conditions (ITS system); in this case, it is possible to predict vehicle arrival times and update the demand in real time, so that the dispatching of vehicles at transfer stations can be determined dynamically in order to improve transfer efficiency.

All these studies are very different because of their application context: some of these pay attention to urban and/or extra-urban contexts, others study the problem from the operator’s or the user’s perspective or focus on the importance of designing synchronized timetables that can satisfy demand requirements as much as possible; again, some studies define the frequency, and thus the timetable, of runs along the same line in connection with demand, in order to guarantee the desired occupation on board; finally, other studies optimize the slack times of coordinated routes and study dynamic vehicle dispatching.

3. **Proposed methodological approach**

The aim of this work is to elaborate a methodological approach based on an iterative modeling procedure, for generating optimal transit timetables; in other words, the objective is to supply a modeling procedure for identifying the optimal departure time from the origin terminal of a given number of runs belonging to different lines, in order to synchronize their arrival/departure times at the transfer nodes of the transit network and to assure a functional temporal coordination of the supplied services.

It is important to underline that, though the advantages of a timetable synchronization can have some positive consequences on the operator’s perspective in terms of supplied service effectiveness and efficiency, this work focuses on the importance of generating synchronized timetables, matching vehicle schedule times with user target times, by analyzing the integration process solely from the user’s perspective.

This work focuses on the case of extra-urban transport services as, because of their medium/low frequency, these are the services in which possible long traveler waiting times and transfer times at the stops can contribute to reduce the probability of the users choosing the public transport system.

Another important aspect is the typology of the system; in particular, two different types of system are possible: the comb and the network system.

In the case of the comb system, it is quite easy to distinguish between two different kinds of lines, “main” and “feeder” lines, operating on the network; the former are the most important lines, usually characterized by high demand values and good performance characteristics in terms of commercial speed, frequency of service, hourly capacity, etc. On the other hand, the “feeder” lines serve the main ones, contributing to the enlargement of their attraction basin by increasing the accessibility of the main lines themselves.
While the difference between main and feeder lines is immediate in the case of a comb system, it is much more complex in the case of a network system; in fact, in this case all the lines operating on the network might have the same level of importance and it could be useful to define some priority rules, in order to synchronize timetables starting with the most important runs.

In both cases, three different vehicle holding strategies can be defined:
- no hold for connecting vehicles;
- hold for all connecting vehicles;
- hold for connecting vehicles, but for no longer than a maximum holding time.

The decision for a connecting vehicle to hold or depart depends on the arrival/departure times of the two connecting runs and on the number of transferring passengers on connecting vehicles. Obviously, these different strategies need to be simulated in different ways and there is no best strategy; the optimal one depends on the type of service.

Moreover, no capacity constraints are taken into account.

On the basis of all these hypotheses, the proposed methodological approach is summarized in figure 1.

The procedure starts with the definition of transport supply and demand systems, both specified temporally; in particular, as far as supply is concerned, a run-based supply model (diachronic graph) has been used in order to explicitly take into account the service timetable.
The methodology is based firstly on the definition of the minimum cost paths, i.e. the paths for which the following disutility function (weighted sum of different temporal shares) is minimum:

\[ V = (\beta_{a/e} \cdot t_{a/e}) + (\beta_{ear} \cdot t_{ear}) + (\beta_{lat} \cdot t_{lat}) + (\beta_{b} \cdot t_{b}) + (\beta_{tr} \cdot t_{tr}) \]

in which \( t_i \) are temporal shares and \( \beta_i \) are model parameters, whose values in disutility/hour, for home-work trips, are summarized in table 1 (Nuzzolo et alii, 2003).

<table>
<thead>
<tr>
<th>Access/Egress</th>
<th>( t_{a/e} )</th>
<th>( \beta_{a/e} )</th>
<th>7.39</th>
</tr>
</thead>
<tbody>
<tr>
<td>Early penalty</td>
<td>( t_{ear} )</td>
<td>( \beta_{ear} )</td>
<td>3.96</td>
</tr>
<tr>
<td>Lateness penalty</td>
<td>( t_{lat} )</td>
<td>( \beta_{lat} )</td>
<td>4.12</td>
</tr>
<tr>
<td>On board</td>
<td>( t_{b} )</td>
<td>( \beta_{b} )</td>
<td>2.42</td>
</tr>
<tr>
<td>Transfer</td>
<td>( t_{tr} )</td>
<td>( \beta_{tr} )</td>
<td>2.42</td>
</tr>
</tbody>
</table>

Table 1. Attributes and parameters.

To find the minimum paths, a recursive algorithm based on a limited-depth-tree, subject to certain constraints has been used; in particular, the algorithm finds out the \( x \) best solutions, which satisfy the following criteria:
- the maximum number of transfers \( n_{tr,max} \) is fixed, in order to avoid that during the same journey the user can choose a path with more than \( n_{tr,max} \) transfers;
- the total travel time of the selected minimum paths has to be lower than \( m \) times total travel time spent by users who leave from the same point of departure and reach the same destination by car, following the minimum path existing between the same O/D pair by using an individual transport mode;
- the differences between scheduled and desired user departure or arrival times have to be lower than a fixed threshold because, for great differences the user will not derive any advantage from the use of public transport because of great early/late penalties.

After defining the set of minimum disutility paths, a dynamic assignment model is used which, unlike a static assignment model, assigns transport demand by distinguishing the flows for each run of each line. In this way, it is possible to know the transferring flow from one run to another, and to decide the priority order for schedule synchronization. Obviously, it is worth defining some criteria in order to determine priority rules; in particular, in the proposed procedure, the synchronization process starts from the interchange nodes in which the transferring flows and the number of lines and runs are maximum. The former criteria takes into account the user’s perspective, because it gives priority to the transfer points in which the number of users transferring from one run to another is maximum; on the contrary, the latter criteria takes into consideration the operator’s perspective, because it gives priority to the transfer points in which the number of interconnected runs is maximum. Once the criteria is fixed, the ordered nodes list is obtained by means of a multicriteria analysis, in which the two criteria can be weighted differently.

When the ordered list is defined, a temporal coordination procedure is applied to the first node of the list in order to define the departure time of each run making a stop at that node. To achieve this, a “what to” approach is adopted: when the transferring flows for each interconnected runs pair are known, the timetable is obtained by minimizing the following objective function:
\[ \omega = \min_{\omega} \sum_{i} \sum_{j \neq i} (\Delta T_{ij} \cdot f_{ij}) \]

in which \( \omega \) is the departure time from the origin terminal vector; \( i \) and \( j \) are two generic runs making the stop at the transfer node \( n \); \( \Delta T_{ij} \) is the transfer time spent by the single user to transfer from run \( i \) to run \( j \), and \( f_{ij} \) is the transferring flow from run \( i \) to run \( j \), obtained at the previous step. Of course, transfer times \( \Delta T_{ij} \) depends on departure times and, therefore, by minimizing the above-mentioned function, a synchronized timetable can be obtained:

\[ \Delta T_{ij} = (\omega_i + t_{i}^{o-n}) - (\omega_j + t_{j}^{o-n}) \]

In the previous expression, \( \omega_i \) and \( \omega_j \) are, respectively, the departure time from the origin terminal of run \( i \) and run \( j \), while \( t_{i}^{o-n} \) and \( t_{j}^{o-n} \) are, respectively, the travel time from the origin terminal to the transfer node by runs \( i \) and \( j \).

Obviously, the objective function is subject to certain constraints; for example:
- all the runs can not depart before a certain schedule (for example 6.00 a.m.);
- transfer times have to be included in a given range \([\Delta T_{tr,\min}, \Delta T_{tr,\max}]\), in order to avoid transferring users missing the connection (\( \Delta T_{tr,\min} \)) or having to wait a long time for the connection (\( \Delta T_{tr,\max} \)). In order to achieve this, it can be useful to introduce a “slack time”, which can help reduce the probability of missed connections. So, if transfer time is not too long, the run arriving at the transfer node first waits for the interconnected run before departing, to allow the alighting users to catch their connection.

Minimum and maximum headway can also be fixed and “cadenced” timetables can also be introduced.

After calculating the new departure times of all the runs making a stop at the considered transfer node, a new indicator is calculated; it is the sum of all the transfer times spent by all the users at the \( N \) interchange nodes in the network:

\[ \sum_{n=1}^{N} \sum_{i} \sum_{j \neq i} (\Delta T_{ij} \cdot f_{ij}) \]

If the value of this indicator at the generic iteration \( k \) is greater than its value at the iteration \( k+1 \), the procedure updates the supply model and it is reiterated once again. Without varying the obtained schedules, it optimizes the departure time of the other runs, starting from the second node of the ordered list. The procedure ends if the objective function value does not change after two successive iterations.

The proposed procedure has been developed in an auto-produced computational program, that needs further refinement, in order to simplify and speed up schedule optimization.

4. Application to a test network

Let us consider a transport network, with two origin and two destination centroids, served by three different public transport lines interconnected at three different transfer nodes (fig.2).
The network is composed of access/egress links, infrastructural links and board links, whose characteristics are reported in table 2.

![Graph of the test network](image)

**Figure 2. Graph of the test network.**

<table>
<thead>
<tr>
<th>Arc</th>
<th>L (km)</th>
<th>v (km/h)</th>
<th>$T_{bus}$ (min)</th>
<th>$T_{car}$ (min)</th>
<th>$T_{foot}$ (min)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-5</td>
<td>0.1</td>
<td>3.6</td>
<td>-</td>
<td>-</td>
<td>1.67</td>
</tr>
<tr>
<td>2-8</td>
<td>0.1</td>
<td>3.6</td>
<td>-</td>
<td>-</td>
<td>1.67</td>
</tr>
<tr>
<td>3-11</td>
<td>0.1</td>
<td>3.6</td>
<td>-</td>
<td>-</td>
<td>1.67</td>
</tr>
<tr>
<td>4-12</td>
<td>0.1</td>
<td>3.6</td>
<td>-</td>
<td>-</td>
<td>1.67</td>
</tr>
<tr>
<td>5-6</td>
<td>7.5</td>
<td>60</td>
<td>15.00</td>
<td>7.50</td>
<td>-</td>
</tr>
<tr>
<td>6-7</td>
<td>10.0</td>
<td>60</td>
<td>20.00</td>
<td>10.00</td>
<td>-</td>
</tr>
<tr>
<td>6-9</td>
<td>3.5</td>
<td>60</td>
<td>7.00</td>
<td>3.50</td>
<td>-</td>
</tr>
<tr>
<td>7-10</td>
<td>5.0</td>
<td>60</td>
<td>10.00</td>
<td>5.00</td>
<td>-</td>
</tr>
<tr>
<td>8-9</td>
<td>5.5</td>
<td>60</td>
<td>11.00</td>
<td>5.50</td>
<td>-</td>
</tr>
<tr>
<td>9-10</td>
<td>8.0</td>
<td>60</td>
<td>16.00</td>
<td>8.00</td>
<td>-</td>
</tr>
<tr>
<td>9-11</td>
<td>4.0</td>
<td>60</td>
<td>8.00</td>
<td>4.00</td>
<td>-</td>
</tr>
<tr>
<td>10-12</td>
<td>6.5</td>
<td>60</td>
<td>13.00</td>
<td>6.50</td>
<td>-</td>
</tr>
<tr>
<td>11-12</td>
<td>6.0</td>
<td>60</td>
<td>12.00</td>
<td>6.00</td>
<td>-</td>
</tr>
</tbody>
</table>

**Table 2. Characteristics of the test network.**

Once user desired departure and arrival times ($DDT$ and $DAT$) (table 3) are known, the O/D matrix (table 4) and the present scheduled times of the runs (table 5), the set of runs departure time can be determined in order to minimize the temporal cost spent by users at interchange nodes and, therefore, to optimize the objective function.

<table>
<thead>
<tr>
<th>$DDT$</th>
<th>$DAT$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>2</td>
<td>-</td>
</tr>
<tr>
<td>3</td>
<td>8.30</td>
</tr>
<tr>
<td>4</td>
<td>8.45</td>
</tr>
</tbody>
</table>

**Table 3. DDT/DAT.**

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>Tot.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>30</td>
<td>50</td>
<td>80</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>40</td>
<td>60</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Tot.</td>
<td>0</td>
<td>70</td>
<td>110</td>
<td>180</td>
<td></td>
</tr>
</tbody>
</table>

**Table 4. O/D matrix.**
By applying the procedure described in the previous chapter, the objective function value varies from 2171 to 581 users-min and the following schedule set is obtained (table 5).

<table>
<thead>
<tr>
<th>Line</th>
<th>Run</th>
<th>Present schedule (hh:mm)</th>
<th>Derived schedule (hh:mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>1</td>
<td>7.00</td>
<td>7.04</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>8.00</td>
<td>8.21</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>9.00</td>
<td>9.00</td>
</tr>
<tr>
<td>b</td>
<td>1</td>
<td>7.20</td>
<td>7.20</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>8.00</td>
<td>7.48</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>8.40</td>
<td>8.37</td>
</tr>
<tr>
<td>c</td>
<td>1</td>
<td>7.30</td>
<td>7.43</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>8.30</td>
<td>8.32</td>
</tr>
</tbody>
</table>

Table 5. Present and derived scheduled times.

The trend of the objective function is reported in figure 3. Note that the procedure ends after only three iterations, because of the test network’s small dimensions.

![Figure 3. Trend of the objective function.](image)

In short, it is worth underlining that the procedure defines the optimal timetable that minimizes the time spent by users at the transfer nodes, by means of a “what to” process. At the same time, it optimizes the service from the operator’s perspective, in terms of network design, because the procedure identifies the runs whose “on board” flow is so low that a run could be unnecessary.

5. Conclusions

For larger networks, the manual application of the proposed procedure is quite complex from a computational point of view; for this reason, an auto-produced computational program has been developed, in order to simplify and to speed up schedule optimization.
The efficiency of the computational program and of the methodological approach itself will be tested on some real local contexts, such as the areas of Reggio Calabria, Crotone and Locride (Italy), which need an integrated bus and train operating plan. To achieve this, transit networks have already been modelled and demand surveys have already been carried out.

References


EFFECT OF HEAVY VEHICLES AT SIGNALIZED INTERSECTIONS IN TURKEY

Serhan TANYEL1, Gülden (CANSEVEN) KALFAOĞLU2, Burak ŞENGÖZ3,
Pelin ÇALIŞKANELLİ4, Mustafa ÖZUYSAL5

Abstract. Heavy vehicle effect has an important role on determining capacity of signalized intersections. Although this problem is very important for intersections in Turkey, only very limited number of studies has been made. In this study, PCE value for heavy vehicles, especially minibuses, at signalized intersections is tried to be determined by using the observations made at two signalized intersections from different cities. Observation results are evaluated by using “headway ratio” and “delay based passenger car equivalent” methods. The results have shown that, PCE value for minibuses increases if the leading vehicle in a queue of a signalized intersection is a minibus which is similar with the results obtained by Ramsay et al. [1].

1. Introduction

Heavy vehicles, due to their size and lower acceleration/deceleration capabilities, may adversely affect traffic flow performance at intersections. The passenger car equivalents concept has been used to account for these effects of heavy vehicles on traffic operations. Heavy vehicles maintain longer headway than cars when crossing intersections; thus

---

1 Faculty of Engineering, Dokuz Eylül University, Izmir, 35160, Turkey, serhan.tanyel@deu.edu.tr
2 Proset Software Inc., İzmir, 35100, Turkey, gcanseven@proset.com.tr
3 Faculty of Engineering, Dokuz Eylül University, Izmir, 35160, Turkey, burak.sengoz@deu.edu.tr
4 Faculty of Engineering, Dokuz Eylül University, İzmir, 35160, Turkey, pelin.caliskanelli@deu.edu.tr
5 Faculty of Engineering, Dokuz Eylül University, İzmir, 35160, Turkey, mustafa.ozuysal@deu.edu.tr
causing delay to the vehicles behind [2].

In Turkey, minibuses are used as a paratransit mode and have an important role in daily life of Turkish citizens. Since Turkish driver’s behavior shows great difference with the other countries’ drivers, these differences should be reflected on the PCE values. Dispersion of vehicles in a queue is another important factor that affects operations of signalized intersections.

In this study, the effect of minibuses on signalized intersection capacity is tried to be determined by using the observations made at one intersection in İzmir and one intersection in Antalya. Results are evaluated by using “headway ratio” and “delay based passenger car equivalent” methods.

2. **Methods for Computing Passenger Car Equivalents (PCE)**

PCE indicate the number of passenger cars that would have the same effects on traffic operations under prevailing road and traffic conditions as one heavy vehicle. The methods used for determining PCE at signalized intersections can be listed as follows:

1. Webster’s Method [3]
3. Headway Ratio Method (HR Method) [6]
5. D_PCE (Delay Based Passenger Car Equivalent) Method [8]

In this study, “Headway Ratio Method” and “D_PCE Method” are used to determine PCE, and according to these methods, determined PCE values are compared.

3. **Observed Intersections**

In this study, headway data was collected at ESBAŞ (İzmir) and Güllük (Antalya) intersections by using video camera. Headways between vehicles have been observed as they cross the stop line of intersection starting from the first vehicle. First headway can be described as the time between the beginning of the green time and the back of the first vehicle passing from the stop line. Second headway can be described as the time between the back of the first and the second vehicle passing from the stop line. Following headways can be measured as the same. For each intersection, it is assumed that after six vehicles have passed, headway for each vehicle would be the same.

4. **Determination of PCE Values by Using HR Method**

In order to compute PCE values by using HR Method, average headways are determined from the video recordings. Average headways and PCE values of minibuses for ESBAŞ and GÜLLÜK intersections are shown in Table 1.
The PCE value of minibuses at Güllük intersection is higher than the PCE value observed at ESBAŞ intersection. The main reason for this situation can be explained by defining the minibus drivers’ behavior. At ESBAŞ intersection, minibus drivers try to find clients that may come from the ESBAŞ Free Zone and they block the intersection for a longer period. However, at Güllük intersection, because of the geometry, minibus drivers cannot stop to wait for new passengers.

5. Determination of PCE Values by Using D_PCE Method

At signalized intersections, vehicles approaching to intersection at red phase will be forced to stop at stop line and join the queue. If there are heavy vehicles in this queue, adverse effect of these vehicles on queue discharge will be important. Increase in queue discharge time and additional delay of vehicles, causes great decrease in capacity and performance of intersection.

Two different queue types are considered in D_PCE method: the base queue in which is composed of all passenger cars and mixed traffic queue in which there is at least one minibus. According to these definitions, a base and various mixed traffic queues were composed using average headways for ESBAŞ and Güllük intersections. Each mixed traffic queue was numbered from 1 to 35. Heavy vehicle percentage and their positions were changed between queues 1 to 35.

According to D_PCE method, for each queue composed to determine delay based passenger car equivalents for minibuses, total and average delay values were determined from field data. In addition, D_PCE values, which are changing according to percentage and position of minibuses in the queue, were plotted as seen in Figure 1, for ESBAŞ intersection and Figure 2 for Güllük intersection.

From the figures, it can be seen that, when the queue leader is a heavy vehicle, D_PCE values is usually greater than the values obtained when the leader was a passenger car. This is because of low acceleration capability of heavy vehicles having a greater average headway value compared with the value of passenger cars.

D_PCE values obtained for minibuses at ESBAŞ intersection does show important differences with the values obtained for passenger cars (Figure 1). Even when percentage of minibus is 50%, D_PCE values is the same (1.37). Hence, headways between passenger cars and minibuses are close to each other.
Especially for Güllük intersection when the queue leader is a passenger car, delay of vehicles waiting behind the leading vehicle increases because of the passengers who get on/off the minibuses (the same reason explained for HR method). Thus, greater D_PCE values were obtained.
6. Conclusions and Suggestions

In this paper, passenger car equivalent for minibuses were investigated and some values were suggested.

From the study it is seen that, vehicles which travel behind the heavy vehicles in the queue, meet additional delay [1]. For ESBAŞ and Güllük intersections, when the percentage of minibus is 14%, the ratio of average base delay to average mixed delay is 96-93%. If minibus percentage is 50%, this value changes as 96-94%. Additional delays caused by minibuses increase intersection discharge time and therefore decrease capacity.

Results have also shown that, the PCE values vary for different percentages of minibuses at each observed intersection. Thus, a constant PCE value was not recommended in D_PCE method [8]. PCE value can vary according to intersection characteristics, positions of minibuses, volume of turning vehicles, driver characteristics, pedestrian flow and vehicle performance. Because of these reasons, special interest should be given in designing signalization facilities.

References


A SEQUENTIAL MODEL BASED ON THRESHOLDS OF LOG-NORMAL ACCEPTABILITY OF THE ATTRIBUTES

Massimiliano Gastaldi, Riccardo Rossi, Romeo Vescovi

Abstract. An analysis of behaviour linked to mobility choices must be considered as multidisciplinary because it can be seen according to various social perspectives [4]: only the convergence of these different perspectives can lead to the definition of valid scientific criteria that can evaluate the effects caused by changes to the system (in the more general sense) on human behaviour. In particular, an analysis of travel behaviour must have valid economic and psychological bases. The better-known disaggregated simulation models of the demand for mobility belong to the family of random utility models. Other simulation tools exist of behaviour of choice, but some of these modelling structures do not acknowledge compensation of the effects and for this reason are called “non-compensatory”. Various scientific studies have demonstrated the need to consider the possibility that choices of mobility are not made according to criteria of a compensatory nature. Starting from these considerations, in order to identify an alternative analysis tool to the classical methods linked to the theory of random utility that can represent the behaviour of transport system users in different situations of choice, and taking inspiration from the behavioural psychology literature, the authors of this paper have developed a model of choice that is partially non-compensatory, sequential to thresholds of acceptability, designated SSA.

1. Logical development of the model and its evolution

The SSA model is a model based on the concept of phased decision strategy, in which it is hypothesised that the sequence of evaluation of the attributes is not known and/or certain: taking inspiration from the EBA model [16], a probability of occurrence is associated to each sequence of attributes, given by the product of the probability of selection of the single attributes in the different steps of the process. Some points are of particular interest:

University of Padova – Department of Structural and Transportation Engineering, Via Marzolo 9 I35131 – Padova (Italy), riccardo.rossi@unipd.it
every sequence of attributes can be interpreted as a particular mental state of the individual decision-maker;
the probabilistic structure of the model can interpret the variability of human preferences over time;
the developed model shows, in relation to the instant choice process adopted by the individual, a strong non-compensatory connotation in places, although being, in the same way as the EBA model, globally compensatory with reference to the mathematical formulation of the probability of choice.

The considered model schematises the real choice process as follows:
- initially the individual takes into consideration the attribute which he feels is the most important at that moment;
- he then identifies the best value of that attribute between the available alternatives;
- he then compares the relative shift between the value of the attribute of each alternative and the best value identified previously with his own threshold of acceptability;
- the alternatives that satisfy this threshold are retained, while the others are eliminated from any subsequent evaluation; in this way, the set of available alternatives is redefined at each step of the sequential choice process;
- only the retained alternatives are evaluated on the attributes not yet considered, according to the methods described in the first step;
- the evaluation process ends when just one alternative is retained, independently of the step of the sequence at which the individual finds himself (non-compensatory sequential process, fig. 1a);
- if at the end of the non-compensatory sequential process more than one alternative is acceptable, the next evaluation phase will be of the compensatory type (fig.1b).

In reality, individuals may alter their order of preference on the attributes over time (this order of preference also usually differs from individual to individual). It is therefore also necessary to express those elements characterising the decisional process in probabilistic terms. In analogy with Tversky’s EBA model, there is a function \( u \) that associates a representative value of its importance to each attribute. At the generic step of the process of choice, this function allows to associate a value of probability of being taken into consideration to each individual attribute: this probability is directly proportional to the importance of the attribute in question and inversely proportional to the sum of the values of importance associated to the available attributes. To each sequence of attributes it will therefore be possible to associate a probability of occurrence, given by the product of the probability of selection of each attribute at every step of the process.

Given the impossibility of knowing and expressing all the significant attributes in a specific situation of choice, in analogy with what Fader and McAlister [4] proposed, it is possible to introduce another variable, specific to the alternative, called the “attractiveness index” of that alternative, which summarises the preference accorded by the individual decision-maker to each one, independently of the aspects explicitly considered.

In the case where more than one alternative passes the entire sequence of evaluation, as mentioned above, the choice between the retained alternatives can be made on the basis of a compensation of the effects of the attributes comprising the sequence of evaluation. This behaviour hypothesis highlights a randomness linked both to the differing sensitivity of the decision-makers to the single attributes (variability of the sequential order of evaluation of the attributes), and directly associable to the compensatory phase of the process.

In reality, individuals may alter their order of preference on the attributes over time (this order of preference also usually differs from individual to individual). It is therefore also necessary to express those elements characterising the decisional process in probabilistic terms. In analogy with Tversky’s EBA model, there is a function \( u \) that associates a representative value of its importance to each attribute. At the generic step of the process of choice, this function allows to associate a value of probability of being taken into consideration to each individual attribute: this probability is directly proportional to the importance of the attribute in question and inversely proportional to the sum of the values of importance associated to the available attributes. To each sequence of attributes it will therefore be possible to associate a probability of occurrence, given by the product of the probability of selection of each attribute at every step of the process.

Given the impossibility of knowing and expressing all the significant attributes in a specific situation of choice, in analogy with what Fader and McAlister [4] proposed, it is possible to introduce another variable, specific to the alternative, called the “attractiveness index” of that alternative, which summarises the preference accorded by the individual decision-maker to each one, independently of the aspects explicitly considered.

In the case where more than one alternative passes the entire sequence of evaluation, as mentioned above, the choice between the retained alternatives can be made on the basis of a compensation of the effects of the attributes comprising the sequence of evaluation. This behaviour hypothesis highlights a randomness linked both to the differing sensitivity of the decision-makers to the single attributes (variability of the sequential order of evaluation of the attributes), and directly associable to the compensatory phase of the process.
The unknown parameters of the model are represented by weights associated to the single attributes of choice and by the respective thresholds of acceptability. In the first version of SSA, reference was made to a single threshold value for each individual attribute (central tendency value). However, it is clear that in the application phase of the model this simplification might not be very effective depending on how wide the scattering is of the real individual thresholds of tolerance compared to the aggregate value estimated in the calibration phase.

After taking this scattering into consideration, also on the basis of what Kurauchi and Morikawa [11] observed, the single thresholds of tolerance have been considered as random variables distributed according to a function of log-normal probability. This choice is consistent with the implicit non-negativity of the acceptability thresholds of the model’s structure. In relation to the specific sequence of evaluation, the probability of choice of an alternative is given by the sum of the probabilities that at each step of the process only this one is retained. At every step of the process, it is also necessary to consider all the possible sets of alternatives still available: the possibility that an alternative is retained at every step of the process being defined in probabilistic terms, so necessarily will the different sets of retained alternatives have to be. This probability will be expressed by the product of the probabilities that, up to the step in question, the alternatives belonging to that set are retained and those which do not belong to it are eliminated. It follows that at every step of the process the probability that a given alternative remains the only one in play will be given by the sum of the products of the probabilities of retention of the possible sets of choice up to the preceding step, and of the probability that, among the alternatives in these sets, only the alternative in hand “survives” to the step in question. The value thus obtained represents the probability of selecting an alternative in a totally non-compensatory way; to
have a measurement of the probability of choice of an alternative, for a given sequence $s$, it is therefore necessary to sum up the probability that this is chosen in a compensatory way at the end of the sequence of elimination on all the possible subsets of “surviving” alternatives. This probability is given by the product of the probabilities of the individual subsets for the probabilistic component of the compensatory type linked to the alternative in question (expressed using the Logit model as suggested by Kurauchi and Morikawa).

Every sequence of attributes has its own probability of occurrence (see what was said about function $\omega$), which must be multiplied for the probability of choice of the alternative once the sequence is known. The final probability is therefore given by the sum, extended to all the possible sequences, of these products.

2. Application of the SSA Log-Normal Model

The proposed model was used to simulate the choices declared during the survey on the choices of goods transport held on the occasion of a national research project².

The model was developed considering a set of choices composed of two alternatives (current and hypothetical), just the attributes cost and time and only two evaluation sequences: firstly a non-compensatory evaluation on just one attribute, then, in the case where both alternatives pass the first phase, a compensatory evaluation (formula of the Logit type).

The results of the application of the SSA Log-normal model (SSA_Ln), the first version SSA model and a Logit model are shown below:

<table>
<thead>
<tr>
<th>Model</th>
<th>Attributes</th>
<th>$lnL(\theta)$</th>
<th>$ln L(\theta)$</th>
<th>$\rho^2 = 1 - \frac{ln L(\theta)}{ln L(\theta)}$</th>
<th>% Right</th>
<th>APCA³</th>
</tr>
</thead>
<tbody>
<tr>
<td>SSA</td>
<td>T, C</td>
<td>-124.77</td>
<td>-63.64</td>
<td>0.490</td>
<td>88%</td>
<td>79.60%</td>
</tr>
<tr>
<td>Logit_01</td>
<td>T, C</td>
<td>-124.77</td>
<td>-73.79</td>
<td>0.410</td>
<td>88%</td>
<td>71.80%</td>
</tr>
<tr>
<td>SSA_Ln</td>
<td>T, C</td>
<td>-124.77</td>
<td>-63.46</td>
<td>0.495</td>
<td>88%</td>
<td>71.80%</td>
</tr>
</tbody>
</table>

Tab 1–Vicenza mechanical industries, distribution of machinery of varying dimensions in Europe (excluding Italy): explanatory variables, diagnostic statistics and validation tests for the calibrated models.

Taking the results from the calibration of the models, the following observations can be made:
- both models of the SSA type have produced a scale of preference of the attributes coherent with that stated by the interviewees;
- the average values of the thresholds for the SSA Log-normal model are consistent with the average values of the levels of variation of the attributes considered unacceptable by the interviewees (table 2);

² “Logistics and transportation for local production systems: methods and models applied to the manufacturing situation in the north-eastern Po Valley”, 2002 [3].
³ Average Probability of Chosen Alternatives
The attribute cost presents lower values (obtained from the calibration of SSA Log-normal) than the average and standard deviation compared to the attribute time (Fig. 2). This implies greater severity when the alternatives are evaluated on that attribute.

A comparison follows of the trends of the probabilities of choice of the hypothetical alternative with the varying of the two attributes cost and time, produced by the SSA Log-normal and Logit_01 models. Figure 3 shows the trend of the probability of choosing the hypothetical alternative with the varying of its cost in relation to the cost of the current service (the value of time of the hypothetical alternative was increased with respect to the value of the same attribute in the current option by an amount close to, but slightly below, the average value of the threshold of tolerance (36%)).

Instead, figure 4 represents the trend of the probability of choice of the hypothetical alternative with the varying of the value of its time, again with respect to the value of time in the current option (a value of cost of the hypothetical alternative 11% higher than that of the current option was taken, which is slightly less than the average threshold of tolerance of cost).
Fig. 3 - Logit_01 and SSA Ln Models. Probability of choice of the hypothetical alternative as a function of the percentage variations in cost compared to the transport option being used (hypothetical alternative time 36% higher than the time of the current option).

The above-described comparative analysis demonstrates the following:

- as regards the probability trend as a function of the variation in cost, the two models give a similar interpretation. The greater weight of this attribute makes the difference in terms of time between the two alternatives not very significant. However, it can be noted that the increase in probability produced by the SSA model is considerable in correspondence to the threshold value, unlike that indicated by Logit, which proposes a significant increase for similar values of cost for the two alternatives (differences close to zero). The SSA trend is more realistic because it is able to show the presence of an inertia effect in the real behaviour towards the current option, which is abandoned only when the advantage of the hypothetical alternative is significant in terms of cost (just noticeable difference);

Fig. 4 - Logit_01 and SSA Ln models. Probability of choice of the hypothetical alternative as a function of the percentage variations in time compared to the transport option being used (hypothetical alternative cost 11% higher than the cost of the current option).
− when the cost threshold is passed, the value of probability indicated by SSA Log-normal increases, if very little, tending towards an asymptotic value much lower than that proposed by Logit. This SSA Log-normal trend appears to be closer to reality (figure 3). This is also true reasoning in aggregate terms and therefore interpreting the probability as a quota of individuals who select the hypothetical alternative: almost all the individuals potentially willing to change from the old to the new alternative, in this case following a significant variation in cost, the threshold of acceptability having been passed, make the change and therefore for further increases in the cost difference the individuals willing to change will always be fewer. SSA Log-normal, inducing an asymptotic value of less than one, admits the existence of a quota of individuals not willing to change because they are tied to their current choice (inertia effect) or not willing to accept a 36% increase in time (non-compensatory behaviour in relation to the attribute time);

− the trends of the probability curves as a function of the variation in time (figure 4) show that, in the area close to nil differences in terms of time between the two alternatives, the probability indicated by SSA Log-normal is markedly lower than that indicated by the Logit model, this is more realistic for the following reasons:
  (a) inertia in favour of the current option;
  (b) the differences in terms of time in this area are not enough to justify perceptible changes of choice (just noticeable difference); a net improvement in terms of time is needed to induce the individual to change. The step effect, due to the passing of the threshold of acceptability of time, is visible but not preponderant; having passed this threshold, there is once again a significant increase in the probability with the reducing of time, with a strong compensatory connotation of the behaviour of choice (the 11% increase in cost compared to the current service is accepted only when there are consistent improvements in terms of time);

− the probability curve as a function of the variation in time produced by SSA Log-normal tends towards an asymptotic value clearly lower than one (the value to which the Logit model tends) and also towards the asymptotic value obtained for variations in cost. This highlights the much more marked non-compensatory aspect of human behaviour the greater the sensitivity is towards the analysed attribute (in this specific case there are a significant number of individuals who are unwilling to accept an 11% increase in cost, whatever the proposed improvement in time);

− the difference between the minimum and maximum value of probability is much more accentuated for the trend of the probability curve relative to the variation in cost (greater sensitivity to variations of this attribute);

3. Conclusions

The SSA Log-normal model certainly offers interesting ideas, but the results obtained require being confirmed by further verifications.

The model requires the same data as those requested by the classical compensatory models such as Logit, allowing a very simple and direct comparison to be made between the two. Regarding this, there is a need to understand the conditions in which one model would be more suitable than the other.
Various studies have demonstrated that ignoring the presence of preferences of the lexicographical type can lead, during the calibration phase of compensatory type models, to an incorrect estimate of the coefficients of substitution between the attributes included in the utility of the alternatives of choice.

Non-compensatory models can represent a valid contrast to the classical compensatory models linked to the paradigm of random utility (in particular the Logit models), as they can take advantage of decisional strategies in stages and thus better represent the behavioural heterogeneity of decision-makers.

The interest for the issues dealt in the present paper also is confirmed from recent studies lead from other research groups [2].

References


AN ONLINE LEARNING SOLUTION FOR DATA DRIVEN FREEWAY TRAVEL TIME PREDICTION MODELS

Hans van Lint, MSc, PhD

Abstract. Unlike predicting traffic speeds and or flows on links in a traffic network, travel time prediction on routes is not a one-step prediction procedure, since travel times are available (and can hence be measured or estimated) only once trips are realized. This implies that standard online learning algorithms, such as the Extended Kalman Filter (EKF), can not be directly applied to train data driven travel time prediction models such as time series or regression models (e.g. neural nets). In this extended abstract we propose an online learning algorithm which can be used for travel time prediction. This algorithm is based on the fact that for non-realized trips a censored observation of travel time (in fact a lowerbound value) is available, which can be used to update the parameters online in a stepwise fashion. A test with real data from a densely used freeway data shows that the algorithm works and yields improvements in comparison to using realized travel times only.

1. Introduction: the travel time prediction problem

The travel time $TT_{ik}$ for a vehicle $i$ departing during period $k$ on some route $r$ in a traffic network is the result of the traffic conditions (speeds, flows, densities) on route $r$ at time periods $p \in \{k, ..., k+TT_{ik}\}$. These traffic conditions off course, may be influenced by many internal or external factors affecting both traffic demand and route capacity along the route during these periods, some of which are clearly beyond the ability of the analyst to predict (e.g. incidents, accidents). Similarly, the expected travel time $TT_{ik} = \langle TT_{ik} \rangle$ for vehicles departing at $k$ on $r$ is a result of traffic conditions $p \in \{k, ..., k+TT_{ik}\}$. Travel time prediction hence implicitly requires predicting - to the degree this is possible – those future traffic conditions along $r$. This poses a “chicken-and-egg” type of problem because the length of the prediction horizon equals the travel time which we wanted to predict in the first place. In [1] a comprehensive overview is given on how different strands of travel time

---

1 Delft University of Technology, Faculty of Civil Engineering and Geosciences, Transport & Planning, h.vanLint@citg.tudelft.nl, http://www.transport.citg.tudelft.nl
prediction approaches tackle this problem. Instantaneous approaches ignore the time dynamics altogether, whereas (simulation) model approaches predict speeds or flows for as long ahead as is required to derive a travel time estimate on the route of interest. Data driven approaches (time series, regression, neural nets, etc) either seek to deduce travel times from past data directly or predict speeds in closed loop for as many periods ahead as is required to deduce travel times. This extended abstract focuses on data driven approaches only.

2. Batch versus incremental learning of data-driven travel time prediction models

In general terms data driven travel time prediction models regress the expected travel time $TT_{kr}$ from a vector of input data $u_{kr}$ (e.g. speeds, flows or even realized travel times), where $u_{kr}$ may include also inputs at time periods $< k$. In the ensuing the subscript $r$ (route) is dropped for ease of reading. The input-output mapping of data driven travel time prediction models can be expressed as

$$\mathbb{R}^m \to \mathbb{R} : y_k = G(u_k, w)$$

(1)

where $y_k$ denotes the travel time predicted by model $G$ which is governed by a vector of adjustable parameters $w$ and a specific choice of model $G$ (e.g. linear, nonlinear, etc). Examples include support vector regression approaches [2], generalized linear regression [3, 4], nonlinear time-series [5], feed forward neural networks [6, 7], and recurrent neural networks [1] to name a few. Training (calibrating) such a model (that is, fit its parameters) usually entails minimization of some cost function (e.g. the expected error) on a training dataset of $N_k$ example input, output records. In batch training approaches, parameters are adapted after presenting the entire batch of training patterns to the model, where this process is iterated for as much epochs as required to attain the desired model accuracy. Naturally, appropriate precautions need to be taken to avoid overfitting the model to the training data (e.g. cross-validation, Bayesian methods, see [1]). In contrast, in incremental training algorithms parameters are adapted after observing each single input-output data pair. As demonstrated in several chapters of [8] the (extended) Kalman Filter algorithm provides an efficient second order approach to this end. The underlying idea is that the parameters $w_k$ are assumed to correspond to a stationary process (a random walk) over time

$$\hat{w}_k = \hat{w}_{k-1} + r_{k-1}$$

(2)

where $r_k \sim \mathcal{N}(0, R_k)$ depicts a process noise vector with covariance $R_k = E[r_k r_k^T]$. With in mind that the travel time prediction model $G$ makes a (non)linear observation (1) on $w_k$, the well-known (E)KF equations can be applied to recursively solve these state-space equations. In this case the state refers to the parameters which are to updated in an incremental fashion as to minimize the sum of squared errors cost function

$$C = \sum_{n_k} e_k^2, e_k = TT_k - y_k$$

(3)
At the same time the algorithm maintains an estimation error covariance matrix $\Sigma_k$ of the parameters. For details on the EKF algorithm, its assumptions and their consequences we refer to for example chapters 2 and 5 of [8]. Here we restrict ourselves to the actual parameter update equations of the algorithm, which read

$$\hat{w}_k = \hat{w}_{k-1} + K_k \varepsilon_k$$

(4)

where

$$K_k = \frac{\Sigma_{kk-1} J_k^T}{\Sigma_{kk-1} J_k^T \Sigma_{kk-1} + R_k^T}$$

(5)

and $\Sigma_{kk-1} = E[\varepsilon_k \varepsilon_k^T]$, $J_k = \frac{dy_k}{d\varepsilon}$ depicts the so-called Kalman gain which can be interpreted as follows

$$K_k = \frac{\text{variance parameters}}{\text{variance model outputs}} \times \text{sensitivity model to parameters}$$

Thus, what the algorithm effectively does is that it recursively updates $w_k$ with a factor $K_k$ which balances the uncertainty in the models’ parameters with the total uncertainty (noise) in the measurement equation, which is also a function of $\Sigma_{kk-1}$. For example, large model uncertainty and small output uncertainty imply large weight updates.

3. An EKF algorithm for online travel time prediction

Recall that an incremental learning algorithm as the EKF can be applied online (in real time) only in case the model performs a one step prediction task, which is not the case for travel time prediction. Nonetheless, consider that at some time period $p$ the last realized travel time $TT_m$ is available from vehicles departing at period $m$, where $m = p - TT_m$. Although for periods $k, m < k < p$, no realized travel times are available yet, a censored observation (in fact a lowerbound value) is, namely

$$TT_k > TT_k^*(p) = p - k$$

(6)

Given the models parameters $w_k$ and inputs $u_k$ the travel time prediction model can calculate a predicted travel time $y_k$ according to (1). Although the true prediction error $\varepsilon_k = TT_k - y_k$ is not available, again, a censored observation of this error is, which reads

$$\varepsilon_k^*(p) = TT_k^*(p) - y_k$$

(7)

Due to (6) this represents a monotonically increasing lower bound of the true error:

$$\varepsilon_k^*(p) < \varepsilon_k, m < k < p + m + TT_k$$

(8)

At each time period $p > k$ for which no realized travel time of vehicles departing at $k$ is available, the censored error (8) provides an incremental estimate of the model prediction error. Letting

$$\xi_k(p) = \varepsilon_k^*(p) - \varepsilon_k^*(p-1) > 0$$

(9)

in which
implies that for a particular departure time $k$ for which no realized travel time is available, the parameters $w_k$ can be stepwise updated at each $p > k$ by substituting (9) into (4), given that this update indeed improves model performance, that is, if

$$TT_i'(p) - y_i |_{n_i} > TT_i'(p) - y_i |_{n_i}$$  \hspace{1cm} (10)

In turn, this is the case if and only if

$$G(u_i, w_i^*) > G(u_i, w_i)$$  \hspace{1cm} (11)

At any particular time period $p$, there will be a number of past time periods $k$ for which no realized travel times are available yet. This means that per time period $p$ possibly more than one weight can be applied with censored errors. In this research this is done sequentially, while at each update (11) is evaluated with respect to the last weight update, which could also have been applied during $p$.

Note that (11) implies that if the parameter-update results in a larger predicted travel time than before, it is retained, otherwise it is discarded, in which case $\varepsilon_k(p)$ must be reset to zero. Intuitively this procedure makes sense. For example, in case travel times (of e.g. 10 minutes) are of an order larger then the unit of discrete time $k$ (of e.g. one minute), the lower bound of eqn (6) will initially (as $p$ is only a few time steps away from $k$) be much smaller than free-flow travel times. Adapting the parameters to these clearly underestimated travel times would not improve learning at all. In situations of congestion build-up, during which travel times tend to increase, it is clear that (11) retains updates only if these contribute to the increasing trend. In case of congestion dissolve, during which travel times tend to decrease, (11) has no effect, since in those cases realized travel times will become available increasingly faster.

4. Results

4.1. Experimental setup

For demonstration purposes we test the algorithm on a state-space neural network (SSNN) as used in [1], where $G$ of eqn (1) represents the SSNN mapping. In this study, a SSNN travel time prediction model is built for the 7 km 3-lane A13 Southbound freeway stretch between The Hague and Delft (The Netherlands). The data for training and testing reflect all Tuesday and Thursday afternoon periods (between 12:00 and 22:00) in 2004 and consist of speeds and flows from dual inductive loops installed on average every 500 meters along this freeway stretch and come from the Regiolab-Delft traffic data server [9]. The dataset hence contains over 60,000 {input, output} records. As inputs, spot mean speeds and flow per minute from each detector along the main carriageway are used. For targets, travel times are estimated (offline) with the so-called piece-wise linear speed-based (PLSB) trajectory method [10]. Travel times range from 6 (free-flow) to 20 minutes (congestion) on this freeway stretch. The ratio between congested and freeflow periods is about 50/50 in the dataset used.
Given the used inputs – spot mean speeds and flows (over the entire carriage way) from in total 14 consecutive dual loops on the A13 southbound mentioned above, the SSNN model structure used below is straightforwardly derived (see [1]). It consists of 13 hidden units, each receiving input signals associated with the thirteen consecutive freeway sections each enclosed by up- and downstream detectors. The context layer consists also of 13 units and is fully connected to the hidden layer.

3 SSNN models were trained on the available data with (a) the incremental offline EKF algorithm; (b) the censored online algorithm described above and (c) with a similar algorithm, which solely uses full errors (as travel times are available). We call this method the online delayed method. For evaluation purposes, four performance measures listed below were used, where $y_k$ denotes the model prediction and $t_k$ the target (realized) travel time.

<table>
<thead>
<tr>
<th>Performance measure</th>
<th>learning method</th>
<th>MAPE [%]</th>
<th>RMSE [s]</th>
<th>Bias [s]</th>
<th>RRE [s]</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMSE$^*$ (Root Mean Squared Error)</td>
<td>Offline</td>
<td>5.2</td>
<td>44</td>
<td>3.2</td>
<td>44</td>
</tr>
<tr>
<td>RRE (Root Residual Error)</td>
<td>Online (censored)</td>
<td>7.9</td>
<td>85</td>
<td>-9.6</td>
<td>85</td>
</tr>
<tr>
<td>Bias (Bias)</td>
<td>Online (delayed)</td>
<td>9.5</td>
<td>95</td>
<td>-4.9</td>
<td>95</td>
</tr>
</tbody>
</table>

* Note that $\text{RMSE}^2 = \text{Bias}^2 + \text{RRE}^2$

### 4.2. Performance

TABLE 1 shows the four performance measures for the SSNN model trained with the three different learning methods described earlier. Clearly, offline incremental method performs best, but due to the fact that targets (realized travel times) are not available in the one-step correction scheme, this result is misleading. The censored approach performs better than the delayed approach in terms of the overall error (MAPE and RMSE). Interestingly this reduction due to a decrease in residual variance and at the cost of a slightly increased structural error (bias). Tentatively, this may be attributed to the fact that with censored approach, the model tracks congestion onset more smoothly, which decreases error variance but may induce some extra bias.

**TABLE 1:** Performance of SSNN travel time prediction model under different learning algorithms
5. Conclusions and discussion

The results indicate that an online extended Kalman filter (EKF) algorithm based on censored errors for online learning parameterized travel time prediction models provides a small but significant advantage over one which uses realized travel times alone, which is most likely due to improved tracking of congestion build-up. Although the standard incremental learning algorithm is superior, this algorithm can never be applied online, since the targets it needs (realized travel times at the next discrete time step) are simply not available.

References

SOLVING A DYNAMIC TRAFFIC ASSIGNMENT PROBLEM WITH SPILLBACK CONGESTION BY MEANS OF THE IGSM ALGORITHM

Lorenzo MESCHINI¹, Guido GENTILE²

Abstract. The aim of this work is to solve a Dynamic Traffic Assignment problem with spillback congestion, by means of a matrix free algorithm aimed at solving general large scale fixed point problems and systems of nonlinear equations.

1. The Dynamic Traffic Assignment model with spillback congestion

The purpose of this work is to solve the Dynamic Traffic Assignment (DTA) with spillback congestion presented in [4] by means of the iGSM algorithm presented in [2], which is a matrix free algorithm aimed at solving general large scale fixed point problems and systems of nonlinear equations.

The model we consider here is a continuous-time formulation of the DTA, where a user equilibrium is expressed as a fixed point problem in terms of arc inflow temporal profiles. In [4], it is shown that, by extending to the dynamic case the concept of Network Loading Map (NLM), stated in [3] for the static case, it is no more needed to introduce the CDNL as a sub-problem of the DTA, because the coherence through the arc performance model between the travel times and the flows loaded on the network consistently to given path choices will be attained jointly with the equilibrium. With specific reference to a Logit route choice model where all the efficient paths are implicitly considered, a formulation of the NLM is devised exploiting the concepts of arc conditional probability and node satisfaction. On this basis, a specific network loading procedure is also obtained as an extension to the dynamic case of Dial’s algorithm. With reference to a Probit route choice model, a formulation of the NLM is devised in [5], where a new all-or-nothing assignment procedure to dynamic shortest paths is proposed and used within a Montecarlo simulation to evaluate the NLM in the Probit case.

¹ University “La Sapienza”, Via Eudossiana 18, 00185 Roma, Italy
lorenzo.meschini@uniroma1.it

² University “La Sapienza”, Via Eudossiana 18, 00185 Roma, Italy
guido.gentile@uniroma1.it
Again in [4], spillback congestion is represented translating any interaction among the flows on adjacent arcs in terms of time-varying arc exit capacities; then, the spillback phenomenon is reproduced as a hypercritical flow state, either propagating backwards from the final section of an arc and reaching its initial section, or originating on the latter, that reduces the capacities of the arcs belonging to its backward star and eventually influences their flow states, consistently with the definition provided by [1]. The key idea is to introduce the spillback representation directly in the arc performance function, without affecting the network flow propagation model internal to the NLM. In particular, the arc performance function is given there by the Network Performance Model (NPM), which is the system of spatially non-separable macroscopic flow models proposed here to simulate both the congestion effect on arc travel times and costs, and the propagation of congestion among adjacent arcs due to queue spillovers.

On this basis, DTA can be formulated as a fixed point problem, which in turn is formalized by combining the NLM, yielding the arc inflow temporal profiles corresponding to given arc travel time and cost temporal profiles, and an arc performance function, yielding the arc travel time and cost temporal profiles corresponding to given arc inflow temporal profiles. The temporal profiles of the maneuver flows at nodes are considered as current variables of the resulting fixed point problem (see Figure 1). Indeed, these play a role in modeling the spillback phenomenon when the available capacity at a node is split among its upstream arcs. However, note that the maneuver flows coincide with the arc inflows and outflows if mergings and diversions are represented separately at the graph level.

\[
t = t(\Phi) \quad \text{; (1)}
\]

\[
c = c(\Phi) \quad \text{; (2)}
\]

Figure 1. Scheme of the fixed point formulation for the DTA with spillback congestion.

Following [4], the arc performance function is expressed by the following functionals:
where $\varphi$, $t$ and $c$ express respectively the vectors of maneuver flow, arc exit times and arc cost temporal profiles; in its turn, the NLM with implicit path enumeration is expressed by the following functional:

$$\varphi = \omega(c, t ; D)$$

where $D$ is the vector of demand flow temporal profiles.

Finally, DTA can be formalized as a fixed-point problem in terms of maneuver flow temporal profiles by substituting into the NLM (3) the arc performance function (1)-(2):

$$\varphi = \omega(c(\varphi), t(\varphi); D) = \Phi(\varphi).$$

On this basis, it is possible to devise efficient assignment algorithms, whose complexity is equal to the one resulting in the static case multiplied by the number of time intervals introduced.

Another important feature of the above model lays in the fact that it does not exploit the acyclic graph characterizing the corresponding discrete-time version of the problem. Specifically, the hypothesis that the longest time interval must be shorter than the smallest free flow arc travel time is avoided. This way, it is possible to define “long time intervals” (5-10 min), which allows overcoming the difficulty of solving DTA instances on large networks.

2. The iGSM algorithm

In [4] the fixed point problem introduced in the previous section is solved through a simple MSA algorithm; thus the possibility to device more efficient solution algorithms for this problem is still an open research issue. It is worth noticing that the problem to be solved can be very large if real size networks are considered; consequently, even if the first order derivatives of the NLM and NPM functions were available (which presently is not the case), their evaluation would be really expensive.

In [2] an algorithm is presented for the solution of systems of nonlinear equations, of the type $F(x) = 0$ with $x \in \mathbb{R}^n$ and $F(x) \in \mathbb{R}^n$, that presents the following properties:

- no analytical form of $F$ is needed;
- the computational costs of evaluating $F$ may be very high;
- the size of the problem can be extremely large;
- $F$ may be subject to stochasticity.

To be noticed that a fixed point $x = T(x)$ is the solution of a system $F(x) = 0$ where $F(x) = T(x) - x$.

The solution method proposed in [2], called iGSM, inherits features of both quasi-Newton and inexact Newton methods, which are methods devised to solve systems of nonlinear equations where the first-order derivative is either too expensive to compute, or is not available at all. The proposed approach use a population of previous approximate solutions of the problem to generate an approximation of the inverse of the Jacobian. The key idea is to identify the approximate tangent hyperplane to the function at a given point through a least square estimation based on a given number of previous approximate solutions. This approach allowed the authors to design a matrix-free algorithm, a crucial property for large-
scale systems, leading to a very competitive algorithm to solve large scale nonlinear systems of equations.

For its characteristics, this solution method appears to be highly suitable to improve the efficiency in finding the solution of the proposed DTA problem; in order to prove this statement, we will present and discuss a number on numerical experiments.

References


CONTROLLING CONGESTION ON PLATFORMS
-A CASE STUDY WITH LONDON’S UNDERGROUND NETWORK-

Hiroshi SHIMAMOTO1
Jan-Dirk SCHMÖCKER2
Fumitaka KURAUCHI3
Michael G H BELL4

Abstract. The problem of full transit does not only cause discomfort on the trains but also safety risks on the platforms. It is therefore required to manage the level of congestion by controlling the number of passengers on the platform. This study proposes to control the flow of passengers by changing the flow rate (capacity) of the ticket gate. By enforcing capacities, the number of passengers on the platform can be managed to be less than the acceptable level. Transit assignment model considering the risk of failing-to-access as well as failing-to-board at the platform is used to explain the passenger flow. Passengers decide which line to take by considering the travel time, waiting time and risk of failing-to-access/board. By controlling the number of passengers, the effect of congestion mitigation on the platform is discussed as well as increase in travel cost.

1. Introduction

The problem of full transit vehicle is expected daily by passengers in most of the big cities all over the world. For example, many commuters during the morning peak-hours experience failing to board the transits due to the over crowding in London underground.

1 Graduate School of Engineering, Kyoto University, Yoshidahonmachi, Sakyō-ku, 606-8501, Japan, shimamoto@trans.kuciv.kyoto-u.ac.jp
2 Centre for Transport Studies, Department of Civil and Environmental Engineering, Imperial College London, SW7 2AZ, UK, j-d.schmoecker@imperial.ac.uk
3 Graduate School of Engineering, Kyoto University, Yoshidahonmachi, Sakyō-ku, 606-8501, Japan, kurauchi@trans.kuciv.kyoto-u.ac.jp
4 Centre for Transport Studies, Department of Civil and Environmental Engineering, Imperial College London, SW7 2AZ, UK, m.g.h.bell@imperial.ac.uk
This situation causes not only complaints from commuters but also safety problem on the platforms. The U.K.’s Rail Safety and Standards Board (RSSB) for example pointed out that overcrowding at platforms and access links to platforms are one of the biggest hazards for accidents ([1]). The RSSB further recommended in the same report that the throughput of ticket gates should be adjusted to control the crowding of platforms.

Based on above background, this paper aims to evaluate the effect of introducing the flow control at ticket gates in the London underground network, using the transit assignment model which can represent the change in passenger flow in response to a flow-control measure. This can be regarded as applying ramp metering measures to congested transit networks.

2. Existing research

Similar to traffic assignment, the capacity of transit lines should be considered in public transit assignment. Many researchers have proposed a transit assignment model to consider this matter.

According to queuing theory, waiting time at the station becomes infinite when passengers’ arrival rate reaches vehicles’ capacity. De Cea and Fernandez [2] showed a model which approximates waiting time at stations as a BPR-type function. The problem of this approximation is that waiting time remains finite even when passengers’ arrival rate reaches vehicles’ capacity. Then, Cominetti and Correa [3] presented a framework for congested transit assignment that can incorporate congestion functions obtained from queuing models. Since those two models are based on queuing theory, it is assumed that the arrival rates of passengers are stable and less than the capacity during the modelled period. Therefore, it is impossible to represent a passenger arrival rate exceeding the network’s service capacity in a short period, which could happen during peak hours.

Bell et al [4] introduced a different approach based on absorbing Markov Chains to handle capacity constraints in transit networks. Central to the approach is the introduction of a probability that passengers are not able to board the first vehicle arriving if this vehicle does not have sufficient available space. This “fail-to-board probability” is set in such a way that all the available space is used but all demand exceeding the available capacity remains on the platform. The detailed idea is shown in the next section. Bell and Schmöcker [5] then consider that overcrowding and the resulting fail-to-board probability deter passengers from attempting to board at overcrowded platforms if they have feasible alternative routes and hence influence their route choice. Instead of focusing on the shortest route, passengers might consider taking longer but less congested routes. The method of successive averages is used to set the fail-to-board probability in an iterative process to find the user equilibrium under consideration of passengers’ risk-averseness regarding failure to board. Kurauchi et al [6] further improve the path choice algorithm by replacing it with a search for the optimal hyperpath reflecting the fact that in transit networks the common lines problem is often encountered. Then, using this model, Schmöcker et al [7] evaluate the impact of controlling the access to platforms. They applied the proposed method to a small toy network to verify the model, but an application to a larger network has not been carried out. In this study, the model proposed by Schmöcker et al [7] is applied to London’s underground network, and the practicability of the proposed model is discussed.
3. Capacity constraint transit assignment model

3.1. Fail-to-board and fail-to-access probability

In the public transit, the number of passengers on board can not exceed the line capacity and passengers already on board have an absolute priority to passengers who want to board. To reflect it, Bell et al. [4] introduced “fail-to-board probability”, $q_{kl}^B$, which is the probability that passengers who want to board can not get the vehicle due to overcrowding. Then, the number of passengers travelling between stations can be represented as below:

$$x_{line_{kl}} = x_{on-board_{kl}} + \left(1 - q_{kl}^B\right)x_{boarding-demand_{kl}}, \forall k \in U, l \in L$$  (1)

where $x_a$ denotes the flow of arc $a$ and $line_{kl}$, $on-board_{kl}$, $boarding-demand_{kl}$ denote line arc, on-board arc, boarding arc of platform $k$ of line $l$.

The adjustment is done with

$$q_{kl}^B = 1 - \max\left\{0, \min\left(\frac{cap_{line}(l) - x_{on-board_{kl}}}{x_{boarding-demand_{kl}}}, 1\right)\right\}, \forall k \in U, l \in L$$  (2)

where $cap_{line}(l)$ denotes the capacity of line $l$.

In the transit assignment model this is realised through the introduction of a Fail-to-Board node in the network which is located between the platform and the boarding node (Figure 1). Passengers wishing to board a service must travel via this node at which they might not be able to progress but instead are hypothetically moved to their destination by the probability $q^B$. Figure 2 further shows that the on-board passengers are directly transferred to the boarding node which ensures their priority over those wishing to-board.

![Figure 1. Network description](image)
The “fail-to-access probability”, $q_k^A$ can be defined in a similar way. In Figure 1, the number of passengers who can access to the platform can be represented as

$$x_{access_k} = (1 - q_k^A)(x_{access-demand_k}), \forall k \in U, l \in L$$

(3)

where,

$$q_k^A := 1 - \min\left(\frac{cap_{platform}(k)}{x_{access-demand_k}}\right), \forall k \in U, l \in L$$

(4)

access$_k$ and access-demand$_k$ denotes access arc, access-demand arc of platform $k$, respectively. cap$_{platform}(k)$ denotes the capacity of the platform $k$, which depends on the number of ticket gates and the escalator speeds.

3.2. The cost of hyperpaths

Passengers would consider not only the travel time but also how crowded the vehicle or platform is when they deter travel routes: i.e. they might take longer but less congested routes instead of taking the shortest but heavily crowded routes. To reflect it, the cost of hyperpath $p$ is assumed as below.

$$g_p = \sum_{a \in A_p} \xi_a \cdot \alpha_a \cdot c_a + \sum_{i \in i_p} q_i^a \cdot WT_{ip} - \theta_A \ln\left(\prod_{i \in i_p}(1 - q_i^a)^{\xi_a}\right) - \theta_A \ln\left(\prod_{i \in i_p}(1 - q_i^B)^{\xi_a}\right)$$

(5)

$$\xi_a = \begin{cases} 1 & (a \in LA) \\ \phi & (a \in WA) \\ \nu & (a \in BA) \\ 0 & (otherwise) \end{cases}$$

(6)

The first term and the second term represent the travel time for the arcs and the expected waiting time at stop nodes respectively. $\alpha_a$ and $\beta_a$ denotes the probability that traffic traverses arc $a$ and the probability of traversing node $i$ respectively. $\phi$ and $\eta$ denotes ratio of the time value of waiting and walking, respectively to the value of boarding. $\nu$ denotes the boarding resistance. $WT_{ip}$ is the expected waiting time at stop node $i$ of hyperpath $p$, which can be calculated as follows:

$$WT_{ip} = \frac{1}{|F_{ip}|}$$

(7)

$$F_{ip} = \sum_{a \in OUT(i)} f_{a(p)}$$

(8)

The third and forth term represents the cost associated with the risk of failing to board and failing to access. The parameter for risk of failing to board, $\theta_A$ or $\theta_B$, denotes risk averseness. If $\theta \rightarrow \infty$, then passengers are absolutely risk averse, and they are not interested in travel time and expected waiting time; when $\theta = 0$ passengers do not care about the risk of failing to board or failing to access the platform due to overcrowding. Note that because equation (5) can be separated into costs from each node, Bellman’s principle can be applied to find the minimum-cost hyperpath (Kurauchi et al [6]).
3.3. Mathematical formulation

Let us assume that passengers use a hyperpath of minimum cost in Equation (5). The cost of a hyperpath is a function of the fail-to-board probability for each transit line on each platform and the fail-to-access probability for each platform. By contrast, these probabilities depend on boarding demand, passengers already on board, and transit line capacity, which in turn depends on these probabilities. Therefore, the transit assignment model can be formulated as a fixed-point problem, which defines the equilibrium:

\[
\begin{align*}
\mathbf{y}^* \cdot \mathbf{u}(\mathbf{y}^*, \mathbf{q}^A, \mathbf{q}^B) &= 0, \quad \mathbf{u}(\mathbf{y}^*, \mathbf{q}^A, \mathbf{q}^B) \geq 0, \quad \mathbf{y} \in \Omega \\
\mathbf{q}^B \cdot \mathbf{v}(\mathbf{y}^*, \mathbf{q}^B) &= 0, \quad \mathbf{v}(\mathbf{y}^*, \mathbf{q}^B) \geq 0, \quad \forall 0 \leq q^B \leq 1 \\
\mathbf{q}^A \cdot \mathbf{w}(\mathbf{y}^*, \mathbf{q}^A) &= 0, \quad \mathbf{w}(\mathbf{y}^*, \mathbf{q}^A) \geq 0, \quad \forall 0 \leq q^A \leq 1
\end{align*}
\]

where

\[
\begin{align*}
\mathbf{u}_p(\mathbf{y}^*, \mathbf{q}^A, \mathbf{q}^B) &= g_p(\mathbf{y}^*, \mathbf{q}^A, \mathbf{q}^B) - m^*_r \\
v_k(\mathbf{y}^*, \mathbf{q}^B) &= cap_{line}(l) - x_{on-board} - (1-q^B_k)x_{board}, \forall k \in U, l \in L \\
w_k(\mathbf{y}^*, \mathbf{q}^A) &= cap_{platform}(k) - (1-q^A_k)x_{demand}, \forall k \in U, l \in L
\end{align*}
\]

As shown in equation (12), \( \mathbf{u} \) denotes a vector of the cost difference between \( g_p(\mathbf{y}^*, \mathbf{q}^A, \mathbf{q}^B) \) and the minimum cost from the origin \( r \) of hyperpath \( p \) to the destination \( s \). Therefore, equation (9) represents the user equilibrium condition. Moreover, as shown in equation (13), \( \mathbf{v} \) denotes the vector of vacancies on the line arc on line \( l \) from platform \( k \). Therefore, equation (10) represents the capacity constraint condition of the lines. Similarly, equation (11) represents the capacity constraints of the platforms. This fixed-point problem can be solved by combining the method of successive averages and absorbing Markov chains (Kurauchi et al. [6]).

4. London case study

4.1. Research area and data collection

The proposed method is applied to the central part of London’s Underground network as shown in Figure 2, which consists of 56 stations and 12 lines. Those data listed below are available from London Underground: travel times between stations, estimated waking time between different platforms, service frequency and the capacity of each line. Since London underground announces that they cancelled about 5% trains due to longer dwell times, it is assumed that the service capacity and service headway on all lines is reduced to 90%. We use the OD matrix of the whole network from 7:45 am to 8:45 am for passengers entering the inner zone estimated by Schmöcker [8]. As the capacity data of the platforms is not available, it is defined as the number of passengers per minutes who pass through that platform during the busiest period (from 8:15 am to 8:30 am).
4.2. Parameter setting

Parameters to be estimated in this study is $\phi$, $\eta$, $\nu$, and $\theta_A$, $\theta_B$. As $\phi$, $\eta$ and $\nu$ are chosen from the value listed in Table 1 to minimise RMS errors from observed passengers' flow of each line, $\eta=2$, $\phi=1.5$, $\nu=3$ are selected. However, it is impossible to estimate risk averseness parameter $\theta_A$, $\theta_B$ appropriately from the observed passengers' flow. Therefore, it is assumed that $\theta_A=150$ and $\theta_B=50$.

Table 1. Nomination values of parameters

| $\phi$, $\eta$ (1/boarding time value) | 1 | 2 | 3 | 4 | 5 | 6 |
| $\nu$ (minutes/number of boarding) | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
Figure 3. The comparison of observed and assignment passenger flow

Figure 3 shows the comparison of the observed passengers’ flow and assignment result using the parameters listed above. It is said from Figure 3 to be more or less over-estimated, but as RMS errors are 1374.3, it can be concluded that the reproducibility is not bad.

4.3. Evaluation

To evaluate the impact of controlling the in-flow rate, the capacity of access arcs to crowded platforms is assumed to be controlled by the operator (see Figure 1) in order to remain the number of passengers on the platform is less than a certain threshold. This could be for example through controlling the number of ticket gates letting passengers in or through controlling the speed of the escalators leading down to the platforms. Then, passengers’ flow would change because some passengers might avoid congested arcs with a higher fail-probability and a new equilibrium condition will occur. The effect of the in-flow rate control is evaluated by its impact on the total travel time and the total number of passengers who fail to board and fail to access. The results will be shown at the conference.

References


BOUNDARY RATIONALITY AND TRANSPORTATION BEHAVIOR:
LESSONS FOR PUBLIC POLICY

Jonathan GIFFORD, Cristina CHECHERITA

Abstract: This paper examines bounded rationality and non-rational travel behavior and their implications for transportation policy. Its primary aim is to assess how well policy and planning models address such behavior and whether such behavior matters in a public policy sense. The paper concludes that it is necessary to incorporate bounded rationality more extensively in travel demand modeling at small-scale levels. This could help understanding patterns that do not usually correspond to the utility maximization paradigm.

1. Bounded rationality in social sciences

Why do you travel to shopping center X instead of Y? How many alternatives do you evaluate when setting up a holiday trip? Have you ever driven only for the sake of driving? What would cause you to change the habitual way you travel to work? What would you say if an “intelligent agent” calculated for you the most efficient route in terms of real-driving time?

These are questions related to the way people choose in real life. A rational process of choice, in a stricter (or looser) form, implies consideration of all (or as many as possible) alternatives, in a panoramic fashion; evaluation of their consequences; and choosing the alternative that maximizes a preferred value criterion, usually utility. In addition, the economic theory of rational choice is based on the fundamental assumption of consumer

1 An earlier version of this paper was submitted for presentation to the 4th International Scientific Conference "New Challenges in Transport and Communications", to be held in Pardubice, Czech Republic, on September 14-15, 2006.
2 Jonathan Gifford, Professor and Director, Master’s in Transportation Policy, Operations and Logistics, George Mason University, School of Public Policy, 3401 Fairfax Drive, MS 3B1, Arlington, Virginia 22201 U.S.A., Tel. +1-703-993-2275 / fax +1-801-749-9198, Email: jgifford@gmu.edu.
3 Cristina Checherita, Graduate Research Assistant, Ph.D. Student, George Mason University, School of Public Policy, 3401 Fairfax Drive, MS 3B1, Arlington, Virginia 22201 U.S.A., Tel. +1-703-993-2275 / fax +1-801-749-9198, Email: cchecher@gmu.edu.
sovereignty – that is, that consumer preferences are innate, pre-determined in any choice situation and do not depend on the alternatives available for choice.

“Bounded rationality” seems to come closer to the way we make most choices in real life. It refers to the cognitive limitations facing a human decision maker due to the finite amount of attention and knowledge available for consideration of all existing alternatives and their consequences in solving a particular problem. The term was introduced by Herbert Simon in the 1950s. In his first edition of Administrative Behavior published in 1948, Simon grouped an individual’s limits or boundaries to rationality into three categories: (i) the individual’s unconscious skills, habits and reflexes; (ii) his values and conception of purpose; and (iii) the extent of his knowledge and information [14]. Theories of bounded rationality are defined as those that “incorporate constraints on the information capacities of the actor” [16].

In the third edition of Administrative Behavior (1976), Simon altered two assumptions to transmute the utility-maximizing “economic man” into a man of bounded rationality, “the administrative man,” recognizable from everyday life [15]:

(i) The boundedly-rational man satisfices – that is, looks for a course of action that is satisfactory or “good enough” and not necessarily optimal. A choice can be made without analyzing all possible (or all known) alternatives.

(ii) The boundedly-rational man drastically simplifies the model of the world – individuals ignore those aspects of reality that appear irrelevant at a given time. This means that decisions can be made with relatively simple rules of thumb that allow people to economize on their cognitive capacities.

Another departure from full-blown rationality is non-rational or “affective” behavior. While bounded rationality implies a reasoning process in finding the satisfactory alternatives, non-rational behavior is defined by psychologists as an impulsive response to affective mechanisms without an adequate intervention of thought [16]. Applied to choice theory, this means that the determinants of choice are based on emotional grounds, rather than objective criteria.

Social science has long accepted that human agents are subject to “bounded rationality.” Apart from the field of psychology, where the concepts are extensively studied, boundedly-rational and non-rational behavior have been incorporated in various domains of social science, including organizational theory and the theory of public policy, finance and economics, software design.

So-called “behavioral finance” has made great advances in the study of these behaviors and many findings of this literature can be extrapolated to other fields, including transportation. Of particular interest are the idea of limited attention and the role of “glittering” news in catching individuals’ attention [2], as well as the potentially adverse impact of regulations requiring greater disclosure on observers’ belief accuracies [10], which raises important policy issues and recommendations for a variety of domains.
2. Bounded rationality in transportation policy

2.1. Evidence of boundedly and non-rational choice in transportation literature

In choices related to travel, many authors emphasize the importance of economizing on one’s cognitive abilities, like habitual travel, and the role affective factors play in travel decisions. The issue of habitual travel and the possibility of inducing change is not fully clarified: some researchers find that travel behavior is neither totally repetitious nor totally variable (Schlich and Axhausen [13]), some find that habits can be broken (Fujii and Kitamura [5]), and others find habits more difficult to break (Garvill, Marell et al. [6]).

Many authors (Cao and Mokhtarian [4], Handy, Weston et. al. [9], Steg [17]) emphasize the role of affective factors in travel decisions and the difficulty of quantifying and incorporating such factors in travel demand management. They find that choice of travel-related strategies is affected not only by objective factors (e.g., amount of travel that individuals utilize, safety, cost), but also by travelers’ subjective assessments with respect to travel: desire and affinities (travel liking), travel attitudes, personality and life-style. The findings imply in many cases that policies designed to alleviate congestion may be less effective than expected, because these qualitative factors are seldom measured and incorporated into demand models.

Short distance travel is usually hypothesized to be more extensively subject to lack of proper planning, affective rules (mood, emotions) or habit. Personal vehicle use in the United States is overwhelmingly local, whether measured in terms of trips or travel (almost 90 percent of personal vehicle trips are less than 20 miles, comprising 55 percent of all personal vehicle travel) [7]). Walking especially could be more prone to emotions. American architect David Rockwell thinks that people’s movement can be “choreographed” by the design of the building or the public space they walk in [18].

Passini studied decision making related to wayfinding as a psychological process. He considers that behavioral actions are linked to the original task over a set of intermediary decisions and subtasks. The original complex task of wayfinding is broken down into subtasks, more manageable, which are undertaken sequentially, in semi-isolation, while still taking into account the problem as a whole. With respect to the dynamics of decision-making, Passini holds that it would be erroneous to assume that for an unfamiliar task, a person could work out a total plan and then execute it. Instead, the wayfinding protocol assumes people to have only a global and vague plan, consisting merely of a few general ideas, and gradually incorporating new information and dealing with unforeseen problems as they occur [12].

In a study of shopping travel behavior, based on a simple statistical analysis, Burnett identifies situations where utility maximizing rules are supplemented by simple heuristics and advocates for a shift beyond rational choice and metropolitan scale analysis [3].
2.2. Public policy approaches to boundedly and non-rational choice in transportation

The investigation of the ways bounded rationality is treated in the transportation literature suggests three public policy approaches: (i) ignore or marginally treat the “problem”; (ii) remedy the “problem”; or (iii) accept the “problem” as a norm and incorporate it in modeling.

(i) Ignore the problem or treat it marginally. This approach assumes that rational behavior is the norm in people’s choice processes and that individuals can meet the logical requirements and search continuously for optimal solutions. Or alternatively, it acknowledges bounded rationality and non-rational behavior but ignores them in modeling due to difficulties, inaccuracies and potential distortions that would result from incorporating highly qualitative variables. Since 1970, when McFadden and Domenich estimated a disaggregated urban travel demand using McFadden’s discrete choice model parameterized as conditional logit, the development of discrete choice models has been framed within the Random Utility Maximization (RUM) framework [11]. RUM assumes that individuals act as rational decision makers, optimizing their direct and indirect utility, with randomness in utility coming from heterogeneity in preferences, due to both inter- and intra-personal variation.

(ii) Remedy the problem. This view acknowledges bounded rationality and non-rational behavior, while treating rational behavior and optimization as the normative standard. Hence, policy-makers must identify interventions that improve the efficiency of individuals’ choice processes and their long-term interests. Examples of such strategies include increasing awareness of travel mode choices and real-time information about traffic. Such interventions must be demonstrably superior to the status quo, and in that sense meet Williamson’s “remediableness” criterion [20]. This approach would also broadly correspond to the recently revived notion of state paternalism, this time as “soft” or “libertarian” paternalism as opposed to either “hard paternalism” or the “nanny state.” Thaler, Sunstein, Beshears, Laibson, Choi and Madrian are among the principal supporters of state soft intervention for “behavioral” reasons. They argue for guidance, or even acting on behalf of individuals, but leaving the choice to them [19].

(iii) Accept bounded rationality as a norm and incorporate it into modeling. This approach accepts that using heuristics (i.e., “reasoning the fast and frugal way” [8]) may not always be sub-optimal. Gigerenzer and Goldstein (1996) find that satisficing rules can match or even outperform various rational inference procedures. Alternatively, even if the results are less than optimal, real life conditions require models based on rules or heuristic simulations of households’ decision-making process [8]. These techniques have been used to develop operational activity-based travel model systems, such as the Albatross system, developed by Arentze and Timmermans in 2004. The choice approach used in these models is decision heuristics, or a rule-based approach, modeled most frequently as a decision tree and consistent with cognitive theories of learning and problem solving. These theories assume that rules are formed and continuously adapted through learning while the individual is interacting with the environment or communicating with others [1].
3. Conclusions and Policy Recommendations

The transportation literature appears to be divided on the impact of bounded rationality on public policy. Mainstream modeling seems to ignore bounded rationality or acknowledge it marginally when formulating policy recommendation. Some studies point out that given the difficulty of influencing non-deliberative choices and incorporating qualitative, non-rational elements in travel demand models, transportation policy initiatives (e.g., to alleviate congestion) may be less effective than expected. Other studies indicate that it is possible to increase people’s awareness about their transportation choices and influence behavior like habitual traveling. The literature is slowly beginning to accept that bounded rationality is the norm in people’s real life-decisions regarding transportation and, as such, modeling and policy should incorporate it.

From a public policy standpoint, the divergence between actual behavior and the representation of such behavior in policy analytical models is important for three reasons. First, we concur with Burnett [3] that truth seeking in model development commands more effort to represent actual behavior in transportation policy models.

Second, the prevalence of short-distance travel, combined with the contention that short trips are more subject to boundedly and non-rational travel behavior, suggests the need for incorporating heuristics into demand modeling. We therefore think that it is necessary to incorporate boundedly and non-rational behavior more extensively in travel demand modeling at small-scale levels. This could help understanding patterns that do not usually correspond to the utility maximization paradigm.

Third, there may be other departures from the assumption of random utility maximization that significantly affect travel behavior. Understanding such departures could help policy makers make wiser decisions about policy interventions to achieve public policy goals.

One could imagine that the advances in technology and the generalization of the internet would help individuals to make travel choices increasingly more based on “artificial” IT-supported utility maximization rules. Internet searching engines allow us to rank a multitude of products in terms of price, quality or other attributes without significant costs. The process of trading off attributes between different alternatives, characterizing the utility maximization rule, will still need to be performed by the travel decision-maker, but it would necessitate less cognitive effort once the ranking is automatically performed. Online mapping, minimizing travel distance between two points, is useful in recommending how to save on travel time, again without much cognitive effort. Electronic passes, with universal use (metro, bus, parking, toll-payment etc.), may help travelers to accept more easily public transportation means or facilities, due to convenience in terms of economizing on their cognitive efforts (e.g., no need to worry about changing passes, having the exact fare or paying each time of boarding).

However, individuals need to be adequately informed about such resources, to have access to them and the willingness to use them. In principle, improving awareness of available choices combined with choice simplification by an external administrative unit would respond to the combined problem of habitual travel and limited attention to information. For example, the state could provide real-time information on congestion, plus a recommendation for a single (optimal) choice in terms of travel time. The principle of Simplify-Inform-Recommend (SIR) should govern traffic guidance in congested roads. By
registering and analyzing the speed of driving on alternative routes, real time information can be provided to drivers during rush hours. The information supplied should be very short, attention catching (e.g., ranked on a scale or showing estimated time between two points in a flashing format) and a recommendation about the best route available in terms of travel time could be provided to drivers. This would save travelers the cognitive effort involved in choosing a route and could guide traffic more adequately. Yet, such a system must be thoroughly tested before implementation in order to be credible to travelers. Such an example of soft paternalism can be more easily accepted in situations involving individual non-deliberative and random, guess-type choice, given that it takes the form of a recommended alternative.

More research is needed to differentiate the rationales and characteristics of the choice process in short versus long-distance travel and test empirically in a large sample the hypothesis that short-distance travel is more extensively subject to heuristics and affective rules than long-distance travel.

References


ANALYSIS OF DEMAND FLUCTUATION ON URBAN EXPRESSWAY USING DYNAMIC PATH FLOW ESTIMATION

Jin-Young KIM, Fumitaka KURAUCHI, Nobuhiro UNO and Yasunori IIDA

Abstract. For more reliable transport network, grasping the variation of traffic demand is a key issue. It is possible to identify the variations of link flow easily but is not possible to observe the demand variations directly. In this study, the demand fluctuation is analysed using the method of dynamic path flow estimation. Path flows are estimated by using one-year data obtained on Hanshin Expressway, and the demand variations are analysed by the estimated results. The relationship between demand variations and surrounding conditions such as weather and commercial custom are discussed. It is expected that the results are valuable for carrying out more detailed traffic control and/or measures.

1. Introduction

The value of time has increased in recent years due to enhanced economic activity and improvements in the quality of life. The reliability of transportation networks has become an increasingly important issue. Consequently, alternative routes to destinations should be provided to avoid substantial loss to trip-makers caused by the unexpected accidents, disasters or traffic demand variations. Indices of transportation network reliability should include the influences of demand and supply dynamics. For more reliable transport network, lots of works have been carried out to consider the demand variations in the framework of the network assignment. For example, Sumalee et al. [1] has proposed the methodology of designing the network with consideration of stochastic demand. The actual variation of demands in the real network, however, has not been fully understood yet.

Recent improvements in data-collecting technology allow the traffic conditions on a road to be assessed almost instantly and can be stored on a huge database. On Hanshin Expressway, traffic detectors observing the traffic volumes, time occupancies and velocities for every 5 minutes are equipped for every 500 metres. These data are stored in the intranet data warehouse and all observations together with other reports such as incident and maintenance reports [2]. Within-day as well as day-to-day fluctuations of traffic flow can

1 Department of Urban Management, Kyoto University, Katsura Campus, Nishikyo, Kyoto, 615-8540, Japan, jeanjin@trans.kuciv.kyoto-u.ac.jp
be evaluated by these data, but the true demand fluctuations can not be evaluated since the link flow is an outcome of the aggregations of OD demand.

Based on above background, this study aims at exploring the fluctuations of the traffic demand on the real expressway network. For this purpose, we shall utilise the dynamic path flow estimation model. By the demand variations obtained from the estimations, within-day / weather / commercial custom characteristics of path flows are discussed.

2. Dynamic Path Flow Estimation Model

The authors had proposed a methodology of estimating dynamic path flow from the observed traffic data [3]. We will use this model for obtaining the demand variation. The brief overview is explained here and refer to Kim et al.[3] for detailed explanation. The proposed model called DCLS-TGV (Dynamic Combined Least Square with Trip Generation as dependent Variables) utilises observed in-flow and link traffic volume and link travel time to obtain time-dependent path flows. DCLS-TGV applies the method of sequential estimation; the estimation is carried out every time interval and the estimates in the previous time intervals are treated as the given values at the subsequent time intervals. We have implemented the following assumptions for estimation.

(A.1) vehicles that enter a link are uniformly distributed for a period of one time interval, which is short enough that the vehicles do not traverse the entire link,
(A.2) vehicles are running with an observed average velocity at each link, and this velocity is maintained for a period of one time interval.

The formulation of DCLS-TGV can be summarised as follows:

\[
\begin{array}{l}
\min \\
\sum_{i \in A_{\text{out}}} \left[ \sum_{p=1}^{n_p} \sum_{l=1}^{l} \left( \delta_{lap} \cdot q_{lap} \cdot y_{lap} + \left( 1 - \delta_{lap} \right) \cdot q_{lap} \cdot \hat{y}_{lap} - y_{v_i}^* \right)^2 \right] \\
+ \sum_{i \in A_{\text{in}}} \sum_{s=1}^{l} \left( O_{ia} - O_{ia}^* \right)^2 \\
\end{array}
\]

subject to

\[
\sum_{a \in A_{\text{out}}} y_{lap} = 0 \quad \text{for all } i \in I, s(1 \leq s \leq t)
\]

\[
\sum_{a \in A_{\text{out}}} y_{lap} = \sum_{a \in A_{\text{out}}} y_{lap} \quad \text{for all } i \in I, n \in \left\{ N - N^{\text{off}} \right\}, p(1 \leq p \leq P_{ia}), s = \{ s \mid 1 \leq s \leq t \}
\]

\[
\sum_{a \in A_{\text{in}}} y_{lap} \geq \sum_{a \in A_{\text{in}}} y_{lap} \quad \text{for all } i \in I, n \in N^{\text{off}}, p(1 \leq p \leq P_{ia}), s = \{ s \mid 1 \leq s \leq t \}
\]

\[
y_{lap} \geq 0 \quad \text{for all } i \in I, a \in A, p(1 \leq p \leq P_{ia}), s = \{ s \mid 1 \leq s \leq t \}
\]
\[ O_y \geq 0 \quad \text{for all } \ i \in I, \ s = \{s \mid 1 \leq s \leq t\} \]  

where,

- \( A_a \): set of links on which link traffic volume is observed during time interval \( t \),
- \( P_{ia} \): the number of branches generated at origin \( i \) on link \( a \) according to \textit{branch enumeration}*,
- \( I \): set of origins,
- \( y_{iaps} \): traffic volume on the \( p \)th branch that uses link \( a \) generated at origin \( i \) during time interval \( s \) (origin-path-specific link traffic volume),
- \( q_{iaps} \): proportion of the traffic that is generated at origin \( i \) during time interval \( s \) and on link \( a \) during time interval \( t \), corresponding to \( p \)th branch (dynamic link use ratio),
- \( v_{ia} \): observed link traffic volume on link \( a \) during time interval \( t \),
- \( g_{iaps} \): prior probability of \( y_{iaps} \),
- \( O_i \): observed traffic volumes generated at origin \( i \) during time interval \( s \),
- \( O_{iaps} \): traffic volumes generated at origin \( i \) during time interval \( s \),
- \( \delta_{iaps} \): binary variable that is 0 if \( y_{iaps} \) is estimated from the time interval \( t \); and 1 otherwise,
- \( I^{\prime} \): set of origins over which a trip generation is observed,
- \( N \): set of nodes,
- \( N^{off} \): set of nodes which have off-ramp,
- \( A_{in} \): set of links leading into node \( n \),
- \( A_{out} \): set of links leading out of node \( n \),
- \( \hat{y}_{iaps} \): estimates of \( y_{iaps} \),
- \( t \): index for the current time interval.

The first term of the objective function attempts to minimise the differences between modelled and observed link traffic volume and the second term tries to minimise the differences between the estimated origin-link specific traffic volume and the calculated ones from the prior information. The third term tries to minimise the differences between estimated and observed traffic generations. The input variables are observed link traffic volumes, observed inflow traffic volumes, link prior probability, and dynamic link use ratio. Prior probability of \( g_{iaps} \) is calculated based on the estimates of the former time intervals, and dynamic link use ratios are calculated by the assumptions (A.1) and (A.2) and the observed link travel time [3]. The model bases on a least squares estimator with linear equality and inequality constraints, which guarantees global optimal solutions. The active set method [4] is applied for obtaining an optimal solution.

*This is a consecutive enumerating process. All the paths are enumerated, and then, if there is on the same path from each origin, it will be the same \( p \)th branch. But when a path diverge and join, it will be numbered with new \( p \).
3. Analysis of Demand Variation

3.1. Research area and data collection

The research area is a part of the Hanshin Expressway from Tsukimiyama to Mukogawa on Kobe route and the distance is about 27.7 km (Figure 1). There is only 1 path for each OD pair in this test network, and so the OD traffic volume can be obtained directly by the estimation. On-ramp traffic volumes, link traffic volumes and link velocities for every 5 minutes are observed by the traffic detector equipped on each link. One-year data from 01/01/2005 to 31/12/2005 are used for estimation except for 25 days (from 08/08 to 31/08 and 22/12) because of the error. The prior probability, $g_{amps}$, is calculated from the previous estimates. Since the demand pattern may be different by the day of a week, the prior probabilities of weekdays, Saturdays and Sundays/holidays were calculated separately.

![Figure 1. Research area.](image)

3.2. Nature of the demand variation

By applying the model, OD traffic volume between each on-ramp and off-ramp is estimated. The estimates are aggregated to four blocks for the simplicity. Block 1 is a western outskirt of Kobe city, and the central business/commercial area locates around Block 2. There is an interchange to intercity highway (Nishinomiya IC) in Block 3, and the expressway further connects to the city of Osaka (Block 4). Figure 2 shows the annual average traffic volume for each OD. According to the figure, traffic volume of OD12 which flows to the centre of Kobe from the western area is the largest. From the figure, there are typically two demand patterns; one peak in the morning and two peaks both in the morning and evening. OD34 has a typical variation where there is only one peak in the morning. Since there are residential areas around Block 3, this variation might be representing a commuting traffic in the morning. On the other hand, the OD23 illustrates the typical variation with two peaks. Since the variations of OD traffic volume within day are different among OD pairs, the OD patterns (or destination choice probability) vary along time.

Figure 3 illustrates the variation of the destination probability for traffic originating from Block 1. The traffic going to both Block 3 (connecting to intercity highway) and Block 4 increase during night. The result matches with the general findings that trips with
longer distance increase at night. The probability of choosing Block 2 starts increasing from around 7 o’clock, and is high during daytime. Since the destination choice probability during daytime is different from the one during night, the result suggests the importance of considering the demand variation within day.

Figure 2. Annual average traffic volume for each OD.

Figure 3. Path choice probability from Block 1 by the flow of time.

3.3. Demand variation owing to the rain

Our experience suggests that a heavier congestion might occur on rainy days. We do not know, however, if this is because of the increase in demand, decrease in network performance, or both. The difference of traffic demand variation on the weather therefore is discussed here. Precipitation data for every hour are obtained for the year of 2005 and we defined the rainy day as the day when the rainfall of more than 1 mm continues more than 2 hours in the morning period. Consequently, 29 days within a year are categorised as rainy days, and the demand variations on the rainy days are compared with the annual variation.

Figure 4 shows the demand variation obtained by the estimation. In order to see the differences from the average condition, average traffic demand calculated using whole data are also shown in the figure.
According to the figures, the traffic volume for OD14 on rainy days seems to increase a little compared with the annual average, but is not true for OD24. Demand variations for other OD pairs are also analysed, and in conclusion, traffic volumes of 5 OD pairs increased whereas the other 4 pairs decreased. To verify the assumption that ‘the average OD traffic volumes on rainy days are the same as the annual average traffic volume,’ the paired t test is applied. As a result, it is statistically proved that the average traffic demands on rainy days are likely to be different from the annual average.

By the above analysis, we can conclude that the demand fluctuations on rainy days are different from the annual average, suggesting that different traffic control or management should be implemented by the weather to increase the efficiency of the network. However, it is difficult to say that the congestion on rainy days is caused by increasing traffic volume. Performance degradation by rain such as a decline of capacity by slippery surface, shorter sight distance and so on, might be major reasons for heavier congestion on rainy days.
3.4. Demand variation owing to commercial custom

In Japan, especially in the Kansai area there is a traditional commercial custom called ‘Gotobi (the date of multiples of 5, i.e. 5, 10, 15, 20, 25 and 30)’ when merchants visit their clients for collecting bills. Traffic volumes are said to be large on Gotobi and the congestion is heavier. Nowadays much of these activities have been replaced to the payment by the bank transfer, but it is said that the traffic congestion on Gotobi is still severe. To confirm this, all OD traffic volumes for Gotobi on weekdays are calculated through a year. Two of the results, OD14 and OD24 are shown in Figure 5.

![Figure 5. Average traffic volumes of OD14 and OD24 on Gotobi.](image)

According to the figures, the traffic demand in the morning peak is larger on Gotobi, implying that more people commute by car on Gotobi. This might be because they are going to use the car for their daytime activities. In the daytime, the tendency for two OD pairs is totally different. For OD14, traffic volume during daytime is smaller compared with the annual average, but for OD 24, it is far larger. As zone 2 covers the major business and commercial district of Kobe, the traffic in the daytime might have increased because of the bill collecting behaviour.
Paired \( t \) test is also applied for this data, and it was found that for most of OD pairs (except OD13), the average traffic volume on Gotobi is statistically larger than annual average. It suggests the importance of handling Gotobi differently from other days.

4. Conclusion

In this study, the demand variation was analysed using estimated OD traffic volume using the observed traffic count and velocity. Demand variation during the day is discussed by the result of estimation using one-year data. Then, the factors influencing the demand variation are explored, and especially the effects of rain and commercial custom are studied here. For the rainy days, it was found that whether the demand will increase or increase depends on the each OD pair, but the average traffic volume is statistically different from the annual average. Also for commercial custom, it was statistically confirmed that traffic volume on Gotobi is larger than the annual average. These results are important for traffic control and management. Increasing of traffic volume is expected on Gotobi, and a traffic control considering this is effective to increase efficiency.

Here, only the weather and commercial custom were considered, and other conditions such as the day of a week, seasonal factor are worth being studied. Also exploring the relationship between variations in traffic demand and performance is very important for better understandings of the effect of variations.

Acknowledgements

We are indebted to the Hanshin Expressway Public Company, Limited for providing the traffic data.

References


BETTER UNDERSTANDING THE POTENTIAL MARKET OF TRANSIT RIDERSHIP AND SERVICES

Kevin J. Krizek 1* and Ahmed M. El-Geneidy2 2

Abstract. Ridership is a key element in the transit industry. Most travel analysis identifies two types of users: captive and choice riders. Captive riders are typically those who lack an alternative an alternative to transit; there therefore use it as their main mode of travel to reach their destination. Choice riders are those who typically choose to use transit or a different mode (car or walking) to reach their destination. Service reliability and availability affects the ridership of both populations. However, substantial increases in ridership are usually assumed to be derived only from choice riders. Non-riders can be further classified as auto captives and potential riders. Auto captives are mainly auto users who don’t have transit as a potential mode of transportation. Potential riders are currently not using transit for certain reasons and/or concerns, but accept the idea of using transit based on certain criteria. This research analyzes results from two surveys conducted in the Twin Cities (Minnesota USA) metropolitan region: one of existing riders and the other of non-riders. The aim is to understand the characteristics of both captive and choice riders, with an eye toward the factors that can increase ridership of the latter population. This research classifies riders and non-riders differently in comparison to previous research. In addition to the captivity to modes, the classification considers regularity of commuting. Accordingly transit riders are classified to four categories: captive riders with regular commuting habits, captive riders with irregular commuting habits, choice

1 Assistant Professor, Urban and Regional Planning. Director, Active Communities / Transportation (ACT) Research Group. University of Minnesota. 301 19th Ave S., Minneapolis, MN 55455- Phone: 612-625-7318 - Fax: 612-625-3513 kjkrizek@umn.edu (Corresponding author)

2 Post Doctoral Research Fellow. Active Communities / Transportation (ACT) Research Group, University of Minnesota. 301 19th Ave S., Minneapolis, MN 55455 Tel: 612-624-8282 - Fax: 612-625-3513 geneidy@umn.edu
riders with regular commuting habits, and choice riders with irregular commuting habits. Similarly, there are four types of non-riders: auto captives with regular commuting habits, auto captives with irregular commuting habits, potential riders with regular commuting habits, and potential riders with irregular commuting habits.

In particular, the research focuses on understanding different classifications of existing and proposed riders (based on preferences from the surveys) and commenting on how using advanced forms of technology could increase the ridership of various populations.

**Introduction and Data**

In 2000, Metro Transit, the local transit provider for the Twin Cities region (Minnesota USA), served around 73,477 unlinked passenger trips. In 2003 Metro Transit served 67,275 unlinked passenger trips. These numbers indicate a decline in demand for public transit service in the Twin Cities during this period, a decline that’s not present in other major transit agencies around the country. In the year 2005, the opening of the new light rail system led to a 30% increase in transit ridership relative to the previous year. This research has two main goals. The first is to provide better knowledge of the composition of the travel market in the Twin Cities metropolitan area, examining in particular those inclined toward transit-related services, while the second is to understand the characteristics of non-riders. Understanding non-riders characteristics aims at defining ways to attract them to the public transit market.

The foundation of this research comes from available data collected by Metro Transit. The data used includes the (a) 2001 survey for current riders and (b) the 1999 survey for non-riders. The former provides information on existing use, the latter gives insights into how to better attract people not currently using the Metro Transit services. These data sources represent extremely rich surveys which, after critical analysis, could yield useful information that would help: (a) increase the efficiency of existing service, (b) suggest ways to use technology enhancements to boost ridership, and (c) learn of different market segments of travelers.

**Research Question**

Travel behavior research classifies travelers mainly in to two groups. The first are regular commuters, which includes mainly workers and students, who tend to travel to certain destinations (work and school) regularly during the week. The second type is irregular commuters, people who tend to have less frequent commuting patterns. This group can include non-worker trips related to businesses and recreational activities. Meanwhile transit research classifies its riders as captives and choice riders, and non-riders as potential riders and auto captives. This research combines the two research approaches to better understand the transit market. It tries to understand the characteristics of each of these groups and the factors affecting
them to use transit responding to the following questions: 1. What do we know of these populations? 2. What are the prospects for increasing the number of choice riders? 3. What role can technology play in attracting more choice riders?

Types of Riders and Non-Riders

Ridership is a key element in the transit industry. Most travel analysis identifies two types of users: captive and choice riders. Captive riders typically lack an alternative to transit, and so use it as their main mode of travel to reach their destination. Choice riders can typically choose to use transit or a different mode (car or walking) to reach their destination. Similarly, Non-Riders can be classified into two groups: first those who are automobile captives or any other mode of transportation, second are automobile dependants or we can name them potential riders, those are people who have choices to use transit or other modes and tend to use their automobile due to several reasons (Jin et al., 2005).

The number of captive riders is usually less than the number of choice riders. Chicago Transit Authority (CTA) reported in the year 2001 around 68% of its riders were choice ones (Chicago Transit Authority, 2001). Similarly, the majority of riders served by the Tri-County Metropolitan District of Oregon (TriMet), the local transit provider in Portland, Oregon, were choice riders (77%). Alternatively, Horowitz (1984) assumed the number of captive riders is much higher compared to choice ones while conducting a demand model for a single transit route. The literature links transit captivity to people with certain types of demographic characteristics, for example: low income, elderly, people with disabilities, children, families whose travel needs cannot be met with only one car, and those who chose not to own or use personal transportation (S. Polzin et al., 2000). It is important to note that new transit users are probably from the potential rider population, while a loss in ridership tends come from choice riders. Loss from captive riders can be linked mainly to changes in demographic characteristics (change in income or car ownership) of the riders. In order to understand the transit market it is important to understand the demand for transit and the factors that affect such demand to switch a potential rider to a choice rider. In the following section discusses the demand for transit service, enabling the understanding of the factors affecting elasticity of demand.

Demand for Transit

The factors affecting passengers’ decision to use transit versus other modes are affected by several costs, including monetary costs (fares), cost of travel time, cost of access and egress time, effort, and finally the cost of passenger discomfort. The Transit Capacity and Quality of Service Manual (TCQSM) provide a comprehensive approach to understanding the transit trip decision making processes, which includes several transit availability factors. These factors address the spatial and temporal availability of service at both ends of the trip (origins and destinations) (Kittelson & Associates, 2003). For passengers, transit trips have three main
components: 1) walking time, 2) waiting time, and 3) trip time. Passengers value their waiting time the most, at a level two to three times that off in-vehicle-time (Mohring et al., 1987). Walking time can be divided to access time to bus stop at the trip origin and egress time at the trip destination. The presence or absence of transit service near origin and destination is found by Murray (2001) to be a major factor in choosing transit as a mode for travel.

Surveys

The current rider survey was conducted by Metro Transit personal in 2001 and included 83 questions. The questions covered a variety of topics including origins and destinations, and rider satisfaction and concerns. Metro transit personal selected random phone numbers and conducted the non-rider survey through phone interviews. The first question was “are you currently using Metro Transit?”. If the respondent answered yes, the interview was terminated, if no the interviewer proceeded with the remaining set of questions. Metro Transit conducted 500 phone interviews in the months of November and December 1999. Each interview included 138 questions directed to non-riders, covering a variety of topics including: reason not using transit, safety and comfort of using transit, concerns related to drivers attitude, concerns related to amenities, concerns related to the commute characteristics, concerns of service reliability if using transit is an option, the level of attractiveness of the current service, and various socio-demographic and economic characteristics.

Analysis Methodology

Using the survey data described earlier in this chapter, the next step is to employ two statistical procedures to uncover the characteristics of the riders and on-riders. The first approach uses factor analysis to learn how each of our measures initially relates to one another. Factor analysis (or principal component analysis) is a statistical technique to extract a small number of fundamental dimensions (factors) from a larger set of intercorrelated variables measuring various aspects of those dimensions. It is used to study the patterns of relationship among many variables with the goal of discovering something about the nature of the measured variables that affect them.

Rider Analysis

First we analyze the rider survey using factor analysis and found 33 questions affecting riders. The results of the factor analysis are shown in Table 1. We first consider the riders survey and examine the data using factor analysis. We found 33 questions affecting riders and categorize them into eight principle components.
Eight factors can be defined through the examinations of the factor loadings. These factors are: (1) driver’s attitude, (2) customer service, (3) type of service, (4) reliability, (5) income and value of time, (6) clean and comfort of buses, (7) safety, and (8) personal. The first factor focuses on driver’s attitude and behavior. A driver’s attitude and behavior includes professional appearance, attitude in terms of operating the bus responsibly and safely, courteousness and helpfulness and calling for stops and intersections. The second factor mainly measures the effect of customer service. This is done through analysis of the accuracy of information being delivered to riders either through phone calls or onboard of the bus. It also includes the courteousness of customer support personal and the amount of time spent on phone to retrieve transit information. The type of service, especially express service, is another factor affecting transit riders. In the Twin Cities region express service operates along various freeways with the right to operate in the freeway shoulders during congestion. Express routes usually serve regular and park and ride stops. Accordingly express service is a major attractor to riders who are trying to avoid congestion. The travel time factor includes the availability of more express service, time of operation during the day, type of bus used as express, location of the park and ride facilities, amenities nearby park and ride facilities, and finally a faster method of payment. Transit service reliability, the fourth factor, is a term used differently between transit riders and operators.

<table>
<thead>
<tr>
<th>Table 1: Cluster analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Driver's Attitude</td>
</tr>
<tr>
<td>Customer Service</td>
</tr>
<tr>
<td>Type of Service</td>
</tr>
<tr>
<td>Reliability</td>
</tr>
<tr>
<td>Income and Value of Time</td>
</tr>
<tr>
<td>Comfort</td>
</tr>
<tr>
<td>Safety</td>
</tr>
<tr>
<td>Personal</td>
</tr>
</tbody>
</table>
Non-Rider Analysis

While for the non-riders a similar approach is conducted to define the major factors affecting them. We selected 36 questions from the non-riders survey that did show an effect on non-riders and how they observe transit as a mode of transportation. Each factor has an Eigen value greater than 1.00. Similar to the riders analysis, Table 3 lists the variables in order of the size of their factor loadings (i.e., coefficients), and shown for each of 11 different blocks of variables (aka factors). Within each of the 11 blocks of variables, the high values (above about 0.5 in absolute value, indicated in bold) are generally all in a single column.

A separate column represents each of the 11 blocks and cumulatively, the 11 factors extracted explaining almost 71% of overall variation in the data. The first factor that affects non-riders is the safety and comfort of the provided service. The safety aspect of the factor originates from questions related to safety at the bus stop and along the trip, while the comfort portion originates from cleanliness of the bus and bus stop, the availability of shelters, and the comfort of seating and temperature inside the bus. Similar to riders, driver’s attitude affects non-riders. Driver’s attitude is derived form questions related to the driver’s courteousness, helpfulness, competency, and accuracy of stop announcement. The third factor is the presence of various amenities the bus stops and special requests related to services that can be provided at work locations. The specialized requests include the availability of service during emergency and late night hours, the availability of service to shopping areas during lunch hours, the presence of a car provided by the employer for work and emergency trips and personalized help in understanding schedules and finding routes. The fourth factor is the characteristics of the
regular commute, which includes questions related to distances and travel time from the non-riders home to various destinations.

The eighth factor affecting non-riders is travel cost and is originated from two questions. The first asks about the actual cost of parking that non-rider pays for parking at work. The second asks non-riders to estimate the current cost of using Metro transit system. Having to transport kids from and to daycare or school is the ninth factor. The tenth factor is travel time, derived from a comparison between travel time using transit and other modes of transportation. The last factor relates to the personal preference of non-riders. This preference originated from a question stating “people like me do not ride transit.” This cluster is important in differentiating between the various types of non-riders (auto captives and potential riders.

Figure 2: Non-riders cluster analysis

Discussion

Based on the previous analysis the transit market can be divided into eight different types of commuters with different preferences. Meanwhile there are two types of commuters, regular and irregular ones. Regular commuters tend to travel on standard basis between origins and destinations everyday, while irregular commuters have commuting needs that vary daily.
Conclusions

Travel market segmentation is a unique way to understand the transit market. We classified eight types of commuters who compose the transit market (refer again to Figure 5). Understanding their attitudes and preferences is an important aspect of retaining current riders and attracting new ones. Captive riders represent users who rely on transit as their primary commute mode. Providing quality service that addresses their needs is important since they ride the system daily. Auto captives rely on their car as a primary transportation mode, likely because transit service is not possible from their origin to destination. Understanding the preferences of the changing area could have the most significant impact on transit ridership. This research used factor analysis to determine a set of factors most affecting rider and non-rider decisions to use transit. Metro Transit should use these factors to understand better the preferences of both choice and potential riders. This research discovered trends between the two groups that, when considered, could attract potential riders and influence choice riders to become more regular transit commuters.
Select References


A GLS ESTIMATOR FOR COMBINED PATH CHOICE MODEL AND O-D MATRIX AGGREGATE CALIBRATION ON CONGESTED NETWORK

Michele OTTOMANELLI¹, Leonardo CAGGIANI² and Domenico SASSANELLI²

Abstract. In this paper a Generalized Least Square (GLS) estimator to solve the simultaneous path choice model (PCM) calibration and O-D matrix estimation using traffic counts is presented. We assume as available information a set of link traffic counts, a starting estimate of the unknown PCM parameters and a starting estimate of the O-D travel demand vector (target O-D- matrix). The problem is formulated as fixed-point model (equilibrium programming) assuming the congested network case, that is we assume the variability of both O-D demand vector and the matrix of link choice proportions (i.e. elements of the assignment matrix) at the same time. In the following, after the presentation of the theoretical aspects of the proposed estimator, two heuristic algorithms, based on the method of successive averages, for the solution of the proposed problem is described. A numerical application shows the statistical performance and capabilities of the approach.

1. Introduction

Estimation of link traffic flows provides transportation planners with important information for traffic network design and control. For this reason, many researchers have been addressing their study to define more and more sophisticated and effective traffic assignment models.

Nevertheless, traffic assignment models reliability and effectiveness depend on other important issues to be faced such as Origin-Destination (O-D) travel demand, network supply model and Path Choice Model (PCM), that is core of traffic assignment model.

¹ Dept. of Environmental Engineering and Sustainable Development, Technical University of Bari. Viale Del Turismo, 4 – 70124 Taranto (Italy), m.ottomanelli@poliba.it
² Dept. of Highways and Transportation - Technical University of Bari.
Via Orabona, 4 - 70125 Bari, (Italy), {l.caggiani,d.sassanelli}@poliba.it
Consequently, others important related problems in traffic modeling are O-D matrix estimation and the calibration PCM coefficients that can, usually, be solved by methods requiring very expensive and time consuming surveys. High costs of disaggregate surveys can be an obstacle to apply traditional and well known methodologies, especially for little or medium sized town whose budget is very modest.

Effective and theoretically consistent methodologies have been proposed to calibrate the parameters of traffic demand models or to estimate O-D matrices by using aggregate data such as traffic counts and/or O-D demand counts. Using aggregate information derived from traffic counts (TC) is very attractive because they are cheap, easy and immediate data to collect.

Few works presented in literature concern calibration of trip distribution and/or modal choice models. In [1], [2], [3] the problem of calibrating systems of travel demand models (four step model) using traffic counts with Generalized Least Square (GLS) Estimators is discussed. In [4], with respect to real sized network, the statistical performance of Generalized Least Square Estimators is analysed in deep. Even less are the works dealing with path choice models calibration, especially for joint calibration with O-D matrix. In this field, Ottomanelli and Di Gangi [5] propose an aggregate calibration of PCM using traffic counts but without O-D calibration with a fixed point and bi-level approach; Russo and Vitetta [6] develop a jointly update of O-D matrix, demand model parameters and link cost functions using TC. Cascetta and Postorino [13] propose a fixed point approach to solve the O-D matrix estimation problem using a GLS estimator.


In this paper a Generalized Least Square (GLS) estimator to solve the simultaneous PCM calibration and O-D matrix estimation based on traffic counts is presented. The problem has been formulated as fixed-point model (equilibrium programming) assuming the assignment matrix as a variable into the optimization steps as well as the O-D matrix. Two heuristic algorithms based on the method of successive averages are also presented to solve the optimization problem.

The algorithms have been applied to a simple network and the statistical performances of the proposed method (estimator) are investigated and discussed.

2. Statement of the problem

In this work the calibration problem has been stated for probabilistic PCM assuming the case of congested networks. Consequent traffic assignment model is the Stochastic User Equilibrium (SUE) [10]. User’s path choice behaviour is modelled by PCM, usually on the basis of random utility theory (RU) [11].

RU theory based PCM estimates the probability $p_k$ that user chooses the path $k$ belonging to his/her path choice set $I$, under the following assumptions:
a) user $i$ knows the paths contained in the \textit{choice set} $I_i$;

b) user associates to each path $k \in I_i$ a \textit{perceived utility} $U_{ik}$ assumed to be a random variable;

c) user chooses the path with maximum perceived utility (or minimum perceived cost).

The perceived utility for the user $i$ is assumed a sum of \textit{systematic utility (s.u.)} term $(V_k = \mathbb{E}[U_k])$ and a \textit{random residual (r.r.)} term $(\varepsilon_k): U_k = V_k + \varepsilon_k \ (\forall k)$.

The r.r. term have zero mean and variance equal to the perceived utility variance:

$$\text{Var}[\varepsilon_k] = \text{Var}[U_k] = \sigma_k^2$$

In this hypothesis the probability $p_k$ that user chooses the path $k$ is:

$$p_k = \Pr[\varepsilon_k - \varepsilon_h \geq V_h - V_k, \ \forall h \in I_i] \quad \forall k \in I_i$$ \hspace{1cm} (1)

Than to define a PCM it is necessary to formulate the expressions of $V_k$ and $\varepsilon_k$. It is possible to specify the s.u. term in the (1) as a linear function of path cost attributes $X_w = -C_w$ and parameters $\beta_w$:

$$V_k(\beta_k, X_k) = -C_k = -\sum \beta_w X_w = -\beta_k^T X_k \quad \forall k \in I_i$$

The simplest and more applied model to specify the (1) is the Multinomial Logit model; this model assumes the distribution of random residual term as a Gumbel variate with parameter $\theta$ that leads to the following PCM formulation:

$$p_k = \frac{\exp \left( \frac{-C_k}{\theta} \right)}{\sum_{h=1} \exp \left( \frac{-C_h}{\theta} \right)}$$

In general, let $\Theta = (\beta, \theta)$ and assumed $X$ fixed, the PCM is a function of $\Theta$:

$$p = p(\beta, \theta) = p(\Theta)$$

The vector $\Theta$ has to be estimated by using a model calibration procedure.

3. \textbf{The proposed calibration model}

The aim of this paper is to calibrate simultaneously the $\theta$ parameter of PCM and the O-D matrix using traffic counts assuming the congested network case.
The most general form to solve the estimation of O-D matrix using traffic counts can be formulated as an optimization problem [1][8][12][13]:

\[
d^* = \arg \min_{d \in S_d} [F_1(d, \overline{d}) + F_2(f, \hat{f})]
\]

where:
- \( S_d \) is the admissible solution set to the problem;
- \( F_1 \) and \( F_2 \) are, respectively, measures of the “distance” between the starting estimate \( \overline{d} \) (target demand vector) and the unknown vector \( d \) and between estimated link flow vector \( f \) and traffic counts vector \( \hat{f} \);
- \( H = A p(\Theta) = \{a_{lk} \ p_k\} = \{m_{od}\} \) is the assignment matrix;
- \( A = \{a_{lk}\} \) is the link-path incidence matrix (\( l = \) link index);
- \( f = H \ d \) with entries \( f_l = \sum_{od} h_{ol} d_{od} \) (\( od = \) OD couple index);
- \( d = \{d_{od}\} \) is the OD travel demand vector.

According with this approach the formulation proposed to calibrate also the distribution parameter is:

\[
\left( d^*, \theta^* \right) = \arg \min_{d \in S_d, \theta \in S_\theta} [F_1(d, \overline{d}) + F_2(f, \hat{f}) + F_3(\theta, \overline{\theta})]
\]

where:
- \( S_\theta \) is the admissible solution set for \( \theta \) parameter;
- \( F_3 \) is measure of the “distance” between the starting estimate \( \overline{\theta} \) (target Logit parameter) and the unknown value \( \theta \).

The problem (2) can be represented as a bi-level programming problem or as a fixed point formulation [5].

In this paper the calibration problem is solved using a fixed point formulation (equilibrium programming):

\[
\left( d^*, \theta^* \right) = \arg \min_{d \in S_d, \theta \in S_\theta} [F_1(d, \overline{d}) + F_2(H(\theta, d) d, \hat{f}) + F_3(\theta, \overline{\theta})]
\]

The distances \( F_1, F_2 \) and \( F_3 \) can be defined following different statistical approaches and assumptions. In this work the Generalized Least Square (GLS) estimator has been assumed because it is robust and the most efficient linear unbiased estimator [12].

For the GLS estimator the measure of distance assume the form of a weighted Euclidean metrics; than the (3) become:

\[
\left( d, \theta \right)_{GLS} = \arg \min_{d \in S_d, \theta \in S_\theta} \left[ (d - \overline{d})^T W^{-1} (d - \overline{d}) + (H(d, \theta) d - \hat{f})^T V^{-1} (H(d, \theta) d - \hat{f}) + (\theta - \overline{\theta})^T Q^{-1} (\theta - \overline{\theta}) \right]
\]
where $W$ and $V$ are the variance-covariance matrices of the error of the target demand vector and the error of the flows, respectively; $Q$ is the variance of the error of $\theta$ parameter.

$W$, $V$ and $Q$ represent the weights of the available information that can be interpreted as the level of confidence (or the reliability) in the available starting data.

In general, these optimization problems are solved with algorithm assuming the assignment matrix $H$ as fixed into the whole procedure \cite{2}\cite{3}\cite{4}\cite{12}. This procedure is suitable when the link cost are known or the network is not congested.

If we assume a congested network, then user's choice is affected by the congestion since link flows and link costs are mutually dependent. For this reason it is necessary to consider in PCM and O-D calibration problem the effects due to congestion on link flow estimation too. For instance, Cascetta and Postorino \cite{13} solve this problem fixing a new one assignment matrix for each iteration but in the optimization step matrix $H$ is a fixed constant.

In this paper the assignment matrix is assumed to vary in each optimization step. In particular, two solving algorithms have been presented.

### 4. Algorithms for problem solution

In the following are proposed two algorithms to solve the simultaneous path choice model calibration and O-D matrix estimation using a GLS specification and SUE assignment model. The difference between the two algorithms is only in the optimization step. In the first resolution algorithm (called A) the demand in the SUE – Logit assignment is fixed to previous iteration value but the Logit parameter is a variable. In the second algorithm (called B) the demand and the Logit parameter are variable in the assignment matrix.

**Algorithm A:**

Step 0  Initialization: $k = 0$

choice of the initial demand vector $d_0$ (i.e. $d_0 = \bar{d}$)

choice of the initial value of $\theta_0$ (i.e. $\theta_0 = \bar{\theta}$)

Step 1  Update $k = k + 1$

Step 2  Solve the problem:

$$(d_k, \theta_k) = \arg\min_{d \geq 0, \theta > 0} \left[ (d_k - \bar{d})^T W^{-1} (d_k - \bar{d}) + (H_k d_k - \bar{f})^T V^{-1} (H_k d_k - \bar{f}) + (\theta_k - \bar{\theta})^2 Q^{-1} \right]$$

with $H_{SUE,k} = f(\theta_k)$ (SUE – Logit assignment problem. The value of the demand vector is fixed to $d_{k-1}$).

The solutions are named $Z_k$ for demand vector and $Y_k$ for Logit parameter.

Step 3  Stop test on the maximum percentage difference between the elements of demand.
\[
\max \left| \frac{z_k^{\text{ed}} - d_k^{\text{ed}}}{z_k^{\text{ed}}} \right| < \varepsilon
\]

If the test is satisfied the algorithm ends otherwise go to Step 4.

**Step 4**

New filtered initial demand vector and new filtered initial value of \( \theta \) parameter are:

\[
d_k = \frac{1}{k} Z_k + \frac{k-1}{k} d_{k-1}; \quad \theta_k = \frac{1}{k} Y_k + \frac{k-1}{k} \theta_{k-1}
\]

go to Step 1.

---

**Algorithm B**:

**Step 0**

Initialization: \( k = 0 \)

choice of the initial demand vector \( d_0 \) (i.e. \( d_0 = \bar{d} \))

choice of the initial value of \( \theta_0 \) (i.e. \( \theta_0 = \bar{\theta} \))

**Step 1**

Update \( k = k + 1 \)

**Step 2**

Solve the problem:

\[
\left( d_k, \theta_k \right) = \arg \min_{d, \theta \geq 0} \left[ \frac{1}{k} Z_k + \frac{k-1}{k} d_{k-1} \right] \quad \left( H_k d_k - \bar{f} \right) \quad \left( H_k d_k - \bar{f} \right) \quad \left( \theta_k - \bar{\theta} \right) \quad Q^{-1}
\]

with \( H_{\text{SUE}, k} = f \left( d_k, \theta_k \right) \) (SUE – Logit assignment problem).

The solutions are named \( Z_k \) for demand and \( Y_k \) for Logit parameter.

**Step 3**

Stop test on the maximum percentage difference between the elements of demand:

\[
\max \left| \frac{z_k^{\text{ed}} - d_k^{\text{ed}}}{z_k^{\text{ed}}} \right| < \varepsilon
\]

If the test is satisfied the algorithm ends otherwise go to Step 4.

**Step 4**

New filtered initial demand vector and new filtered initial value of \( \theta \) parameter:

\[
d_k = \frac{1}{k} Z_k + \frac{k-1}{k} d_{k-1}; \quad \theta_k = \frac{1}{k} Y_k + \frac{k-1}{k} \theta_{k-1}
\]

go to Step 1.
5. Numerical application

The following application aims to test numerically the performances of the algorithms (A and B) proposed in the previous section.

For the application data (supply, demand and assignment model) we referred to Cascetta and Postorino [13] since the problem formulation is similar to the one they propose. The considered network has 5 centroids, 15 nodes and 54 links as depicted in Figure 1. The links are divided into two groups (Table 1), each group has a cost function (Table 2) based on the Doherty separable cost function expression:

\[ c(f) = t_r + t_w(f) \]

Thus, assignment problem is assumed to be on congested network (equilibrium approach).

![Test network](image)

**Figure 1. Test network [13].**

<table>
<thead>
<tr>
<th>First group of links</th>
<th>Second group of links</th>
</tr>
</thead>
<tbody>
<tr>
<td>7-14</td>
<td>8-13</td>
</tr>
<tr>
<td>12-9</td>
<td>12-11</td>
</tr>
<tr>
<td>13-8</td>
<td>13-12</td>
</tr>
<tr>
<td>14-7</td>
<td>14-13</td>
</tr>
<tr>
<td>15-14</td>
<td>17-14</td>
</tr>
<tr>
<td>12-12</td>
<td>17-13</td>
</tr>
<tr>
<td>13-14</td>
<td>18-13</td>
</tr>
<tr>
<td>14-15</td>
<td>19-13</td>
</tr>
<tr>
<td>15-16</td>
<td>19-14</td>
</tr>
</tbody>
</table>

**Table 1. Groups of links.**

<table>
<thead>
<tr>
<th>First group</th>
<th>Second group</th>
</tr>
</thead>
<tbody>
<tr>
<td>6-7</td>
<td>6-15</td>
</tr>
<tr>
<td>6-19</td>
<td>6-15</td>
</tr>
<tr>
<td>10-9</td>
<td>10-11</td>
</tr>
<tr>
<td>10-11</td>
<td>11-10</td>
</tr>
<tr>
<td>11-10</td>
<td>11-20</td>
</tr>
<tr>
<td>12-6</td>
<td>12-15</td>
</tr>
<tr>
<td>13-15</td>
<td>13-14</td>
</tr>
<tr>
<td>14-15</td>
<td>14-17</td>
</tr>
<tr>
<td>15-6</td>
<td>15-16</td>
</tr>
<tr>
<td>16-15</td>
<td>16-15</td>
</tr>
<tr>
<td>16-17</td>
<td>16-17</td>
</tr>
<tr>
<td>17-14</td>
<td>17-16</td>
</tr>
<tr>
<td>14-6</td>
<td>14-17</td>
</tr>
<tr>
<td>13-15</td>
<td>13-18</td>
</tr>
<tr>
<td>14-17</td>
<td>14-15</td>
</tr>
<tr>
<td>15-6</td>
<td>15-16</td>
</tr>
<tr>
<td>16-15</td>
<td>16-15</td>
</tr>
<tr>
<td>16-17</td>
<td>16-17</td>
</tr>
<tr>
<td>17-14</td>
<td>17-16</td>
</tr>
<tr>
<td>18-13</td>
<td>18-17</td>
</tr>
<tr>
<td>19-12</td>
<td>18-17</td>
</tr>
<tr>
<td>19-13</td>
<td>19-18</td>
</tr>
<tr>
<td>19-12</td>
<td>19-19</td>
</tr>
<tr>
<td>20-19</td>
<td>20-11</td>
</tr>
</tbody>
</table>

**Table 2. Cost functions specification.**

<table>
<thead>
<tr>
<th>First group of links</th>
<th>Second group of links</th>
</tr>
</thead>
<tbody>
<tr>
<td>( t_r = 123 \text{ s} )</td>
<td>( t_r = 144 \text{ s} )</td>
</tr>
<tr>
<td>( t_w(f_1) = 7,56 + 0,0006 \frac{f_1}{916-f_1} \text{ se } f_1 \leq 870 )</td>
<td>( t_w(f_1) = 7,56 + 0,0008 \frac{f_1}{655-f_1} \text{ se } f_1 \leq 622 )</td>
</tr>
<tr>
<td>( t_w(f_1) = -772 + 0,94f_1 \text{ se } f_1 &gt; 870 )</td>
<td>( t_w(f_1) = -1084 + 1,84f_1 \text{ se } f_1 &gt; 622 )</td>
</tr>
</tbody>
</table>
The “true” value of the Logit parameter $\theta$ is fixed to 0.71 and the “true” value of O-D matrix is shown in Table 3. The “true” value of link flows has been generated by SUE-Logit traffic network assignment of the assumed true O-D matrix and with the true value of $\theta$. As observed link flow, a set of 10 traffic counts was considered.

As in [13], the numerical analysis was carried out considering 20 target O-D matrices, 20 vectors of measured flows and 20 target values of PCM parameter (actually, in [13] $\theta$ is assumed as fixed, i.e. it is not considered in the calibration problem).

Vectors and values were obtained through random extractions from a normal distribution; in addition, in order to increase the error on target O-D demand matrix the extractions were made by assuming a reduction of 0.5 for the value of each elements of the true O-D matrix. Besides, the target flows vectors were calculated through the assignment of the target matrices to the network, assuming the target values of PCM parameters. The coefficients of variation for demand and flows were respectively: $c_{v_d} = 0.4$, $c_{v_f} = 0.05$. The whole procedure has been repeated two times considering two values of variation coefficients for $\theta$, fixed to $c_{v_\theta} = 0.2$ and $c_{v_\theta} = 0.4$. As many authors suggest [1][2][6], the variance-covariance matrices $W$ (for demand) and $V$ (for flows) can be assumed diagonal.

Let $d$, $f$ and $\theta$ be the true values of the demand vector, flows vector and Logit parameter respectively; the relevant values of variances have been computed through the following expressions:

$$
\sigma^2_{d,od} = (c_{v_d} \cdot d_{od})^2 ; \quad \sigma^2_{f} = (c_{v_f} \cdot f_{i})^2 ; \quad \sigma^2_{\theta} = (c_{v_\theta} \cdot \theta)^2
$$

The traffic assignment model used is SUE – Logit and the relevant Logit PCM is given by the following:

$$
P_{k,od} = \exp\left(\frac{-C_{k,od}}{\theta}\right) \sum_{h=1}^{\text{hed}} \exp\left(\frac{-C_{h,od}}{\theta}\right) \quad \forall o,d
$$

The paths choice set $I_{od}$ , relevant to the O-D couple, is constituted by the path with the minimum cost and by the paths whose cost does not exceed the 25% of minimum cost.

The algorithms A and B have been applied to the test network with $\varepsilon = 0.01$ as stop test and the results have been compared each other.

To evaluate the statistical performances of the proposed methods the following indicators have been used:

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>800</td>
<td>500</td>
<td>1000</td>
<td>300</td>
</tr>
<tr>
<td>2</td>
<td>800</td>
<td>0</td>
<td>1000</td>
<td>500</td>
<td>300</td>
</tr>
<tr>
<td>3</td>
<td>500</td>
<td>1000</td>
<td>0</td>
<td>800</td>
<td>300</td>
</tr>
<tr>
<td>4</td>
<td>1000</td>
<td>500</td>
<td>800</td>
<td>0</td>
<td>200</td>
</tr>
<tr>
<td>5</td>
<td>200</td>
<td>200</td>
<td>200</td>
<td>200</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 3. “True” O/D demand matrix.
1) Mean Square Error (MSE) and Relative Root Mean Square Error (RRMSE) calculated between the following vector:
- the target demand \( \vec{d} \)
- the demand obtained at the first iteration \( \vec{d}^{in} \)
- the final estimate (problem solution) of O-D demand \( \vec{d}^{*} \)
- the target flows on no-counted links \( \vec{t}_{noc} \)
- the first iteration estimated flows on no-counted links \( \vec{t}_{noc}^{in} \)
- the estimated flows on no-counted links \( \vec{t}_{noc}^{*} \)
- the target flows on no-counted links \( \vec{t}_{c} \)
- the first iteration flows on counted links \( \vec{t}_{c}^{in} \)
- the estimated flows on counted links \( \vec{t}_{c}^{*} \)

and their true values.

The formal expressions for the demand vectors are:

\[
\text{MSE} \left( \vec{d}, \vec{d} \right) = \frac{1}{N_{O-D}} \sum_{o,d} \left( \vec{d}_{od} - \vec{d}_{od} \right)^2
\]

\[
\text{RRMSE} \left( \vec{d}, \vec{d} \right) = \sqrt{\frac{\text{MSE} \left( \vec{d}, \vec{d} \right)}{\sum_{i=1,N_{O-D}}^{d_{i}}}}
\]

with \( \vec{d} = \{ \vec{d}, \vec{d}^{in}, \vec{d}^{*} \} \)

The closer to zero the value of RRMSE is, the closer the values are to the true one.

2) Percentage of reduction (% red) relevant to:
- the estimated O-D demand
- the estimated flows on no-counted links
- the estimated flows on counted links

and their target (starting) values for MSE and RRMSE.

For example:

\[
\%\text{red} = \frac{\text{MSE} \left( \vec{d}, \vec{d}^{*} \right) - \text{MSE} \left( \vec{d}, \vec{d} \right)}{\text{MSE} \left( \vec{d}, \vec{d} \right)} \times 100
\]

\[
\%\text{red} = \frac{\text{RRMSE} \left( \vec{d}, \vec{d}^{*} \right) - \text{RRMSE} \left( \vec{d}, \vec{d} \right)}{\text{RRMSE} \left( \vec{d}, \vec{d} \right)} \times 100
\]

6. Results analysis

The analysis of the results was made on the mean values of the indicators obtained. The main values of the results are reported in Table 4 and Table 5.

In both procedures (A and B) all indicators improved a lot since the first iteration.
A part from the results obtained for the traffic count used for the calibration that are trivial, very interesting results were achieved when estimates were compared using the no-counted flows (hold-out sample).

In fact, RRMSE indicators tend to zero and the \( \text{%red MSE} \) reached 96% for the O-D demand and 98.4% for the link flow estimates. In can be note that the greater is the error in the starting estimates of the PCM parameters (i.e. the reliability of the starting estimates) the better is the attained %\text{red}.

It is worthwhile mentioning that algorithm B performs better than algorithm A as it is based on a fixed point formulation with respect both O-D demand vector and PCM parameter.

As reported in Table 5, the achieved results seems to be quite interesting. In fact, it has to be pointed out the good results obtained for the performance indicators evaluated using the hold-out sample (no-counted links) as well as with respect to the O-D demand flows.

Preliminary results from comparisons with methods proposes in literature, such as [13], has shown better statistical performances too.

In addition, to the O-D matrix estimation, the proposed model, allows the consistent calibration of the Logit PCM and then, of the SUE traffic assignment model.

<table>
<thead>
<tr>
<th>( cv_t = 0.2 )</th>
<th>( cv_t = 0.4 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Algorithm A</td>
<td>Algorithm B</td>
</tr>
<tr>
<td>( \theta_{\text{true}} )</td>
<td>0.71</td>
</tr>
<tr>
<td>( \text{mean } \theta_{\text{estimate}} )</td>
<td>0.73</td>
</tr>
</tbody>
</table>

Table 4. Calibration results

7. Conclusion and further researches

In many professional applications, O-D matrix estimation is carried out without given the correct importance to the traffic model parameters. In fact, usually, such parameters are assumed “a priori” on the basis of subjective knowledge or assuming calibration results obtained in different study areas.

In this work, a GLS estimator with fixed-point formulation is proposed with two solving algorithms that allow to estimate simultaneously and consistently the SUE-Logit assignment model parameters and the O-D matrix by using a set of available traffic counts.

The application of the proposed estimator to a test network has shown the effectiveness of the approach that is able to reduce the bias induced on the considered traffic variables. Such a results have been obtained without using expensive and time consuming data and computing effort providing a method that should be helpful for practical application when limited resources for data collecting are available.

The preliminary results of the comparisons with other similar methods are quite promising too. The results obtained on real networks by other authors [3][4] induce to be confident in achieving good result also with the proposed approach.

At this time comparisons with other approaches are in progress such as the method by Lo and Chan [9]. In addition, the effect of the selection of traffic counts to be used in calibration is under investigation, as well as the improvement of the proposed algorithms.
### Table 5. Statistical performances.

<table>
<thead>
<tr>
<th></th>
<th>cv$_{i}$ = 0.2</th>
<th></th>
<th>cv$_{i}$ = 0.4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Algorithm A</td>
<td>Algorithm B</td>
<td>Algorithm A</td>
</tr>
<tr>
<td></td>
<td>demand</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MSE($\tilde{d},\tilde{d}$)</td>
<td>98876.48</td>
<td>98876.48</td>
<td>98876.48</td>
</tr>
<tr>
<td>MSE($d^*,d$)</td>
<td>3999.94</td>
<td>3799.88</td>
<td>3872.40</td>
</tr>
<tr>
<td>MSE($d^*,d$)</td>
<td>4260.30</td>
<td>3786.74</td>
<td>4714.63</td>
</tr>
<tr>
<td>% red MSE</td>
<td>95.7</td>
<td>96.2</td>
<td>95.2</td>
</tr>
<tr>
<td>RRMSE($\tilde{d},\tilde{d}$)</td>
<td>0.567</td>
<td>0.567</td>
<td>0.567</td>
</tr>
<tr>
<td>RRMSE($d^*,d$)</td>
<td>0.114</td>
<td>0.111</td>
<td>0.112</td>
</tr>
<tr>
<td>RRMSE($d^*,d$)</td>
<td>0.118</td>
<td>0.111</td>
<td>0.124</td>
</tr>
<tr>
<td>% red RRMSE</td>
<td>79.2</td>
<td>80.4</td>
<td>78.2</td>
</tr>
<tr>
<td></td>
<td>old-out-sample (no-counted flows)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MSE($\tilde{f}_{\text{no}},f$)</td>
<td>207393.57</td>
<td>207393.57</td>
<td>207547.08</td>
</tr>
<tr>
<td>MSE($f^*_{\text{no}},f$)</td>
<td>4737.94</td>
<td>4006.98</td>
<td>4203.28</td>
</tr>
<tr>
<td>MSE($f^*_{\text{no}},f$)</td>
<td>4775.40</td>
<td>4009.59</td>
<td>4453.91</td>
</tr>
<tr>
<td>% red MSE</td>
<td>95.2</td>
<td>95.9</td>
<td>97.9</td>
</tr>
<tr>
<td>RRMSE($\tilde{f}_{\text{no}},f$)</td>
<td>0.565</td>
<td>0.565</td>
<td>0.566</td>
</tr>
<tr>
<td>RRMSE($f^*_{\text{no}},f$)</td>
<td>0.085</td>
<td>0.079</td>
<td>0.080</td>
</tr>
<tr>
<td>RRMSE($f^*_{\text{no}},f$)</td>
<td>0.086</td>
<td>0.079</td>
<td>0.083</td>
</tr>
<tr>
<td>% red RRMSE</td>
<td>84.8</td>
<td>86.1</td>
<td>85.4</td>
</tr>
<tr>
<td></td>
<td>counted flows</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MSE($\tilde{f},f$)</td>
<td>319967.82</td>
<td>319967.82</td>
<td>319895.57</td>
</tr>
<tr>
<td>MSE($f^*_{\text{no}},f$)</td>
<td>2748.08</td>
<td>282.87</td>
<td>4819.92</td>
</tr>
<tr>
<td>MSE($f^*_{\text{no}},f$)</td>
<td>283.55</td>
<td>287.63</td>
<td>335.42</td>
</tr>
<tr>
<td>% red MSE</td>
<td>99.7</td>
<td>99.7</td>
<td>99.9</td>
</tr>
<tr>
<td>RRMSE($\tilde{f},f$)</td>
<td>0.509</td>
<td>0.509</td>
<td>0.509</td>
</tr>
<tr>
<td>RRMSE($f^*_{\text{no}},f$)</td>
<td>0.047</td>
<td>0.015</td>
<td>0.062</td>
</tr>
<tr>
<td>RRMSE($f^*_{\text{no}},f$)</td>
<td>0.015</td>
<td>0.015</td>
<td>0.016</td>
</tr>
<tr>
<td>% red RRMSE</td>
<td>97.0</td>
<td>97.0</td>
<td>96.8</td>
</tr>
</tbody>
</table>

**References**


THE USAGE OF SEAT BELTS IN TURKEY

Safak BILGIC, Polat YALINIZ, Murat KARACASU

Abstract. Traffic road accidents are one of the most important public health problems all over the world. According to the World Health Organization, almost one million people die because of traffic road accidents each year and 90 percent of the deaths occur in the developing countries. As a developing country, in Turkey, traffic accidents are one of the biggest problems. According to Turkish Parliament Traffic Safety Investigation Commission final report, fatalities are between 10000 and 12000 every year. According to different studies, seat belt usage rate is below the 20% in Turkey. So if this ratio could increase, accident fatalities can be decreased. In this study, causes of this low ratio investigated and some arrangements suggested to increasing the ratio.

1. Introduction

In Turkey, traffic accidents are one of the biggest problems. In 2005, 570419 traffic accidents occurred in Turkey. According to official records, 3215 people died, 123985 people injured and approximately 534 million € crash cost occurred [1]. However, actual fatalities are bigger than official number [2]. Because, statistics prepared by traffic accident reports, which are prepared by traffic policemen and so this number indicates only accident place fatalities. The statistics doesn’t include some of fatalities, which are occurred in hospital or in transport for hospital. According to Turkish Parliament Traffic Safety Investigation Commission final report, fatality number is between 10000 and 12000 [2]. And International Road Traffic and Accident Database, IRTAD (is operated within the framework of the OECD), increases 30% Turkey's official fatality numbers to reach actual number [3].

According to different studies, seat belt usage rate is below the 20% in Turkey. So if this ratio could increase, accident fatalities can be decreased.

1 Eskisehir Osmangazi University, Eskisehir, TURKEY safakb@ogu.edu.tr
2. Transportation system of Turkey

Highways used in Turkey for approximately 95% of freight and passenger transportation. Figure 1 shows changes in numbers of traffic accidents, fatalities and injuries between 1960 and 2005 [1].

![Figure 1. Traffic accidents, fatalities and injuries, between 1960 and 2005.](image)

3. Seat belt usage in Turkey and the rest of the world

In Turkey, from 18.06.1986 onwards, seat belt practice for "drivers and the passenger sitting next to the driver on long-distance highways in cars and land vehicles and minibuses that are put in process in the same way as the cars in terms of registration" was started. From 01.01.1992 on, it was made obligatory to car drivers (excluding taxi drivers), jeep and commercial minibus drivers (excluding passenger minibus drivers) and front seat passengers to fasten seat belts on local roads. At 11.01.1995, the practice of using seat belts in the back seat was passed on to [4]. According to the code, the ones who do not have and use seat belts in vehicles such as automobiles and minibuses are fined to pay money (approximately 30€ in 2006) and with 5 fine points [5]. So the usage rate of seat belt increased to 90% in 1992, with public campaigns, strict controls and education [6]. But then with relaxation in campaigns, this ratio is decreased dramatically. For example, in a research performed by General Directorate of Security (GDS) Traffic Research Centre through observation in Ankara in 1999 at 27 junctions where there is intense traffic, it has been observed that 8557 (21.08%) of personal car drivers out of 40587 wear seat belts. In this study, it has been reported that none of the passengers traveling in the back seat wears...
seat belts. Same research performed in a small city, Cankiri, and it has been observed that 8.19% of personal car drivers wear seat belts. Same research performed intercity roads of Ankara and Cankiri and it has been observed that 70.71% and 61.09% respectively [7].

These rates are very low compared with developed countries. Seat belt wearing rates in the EU15 are between 45% and 95% for front seat occupants and between 9% and 75% for rear seat passengers [8].

According to another study of European Union countries in the mid-1990s found front-seat use of seat-belts of 52–92% and rear-seat use of 9–80% [9]. Use of seat-belts in front seats in the United States rose from 58% in 1994 to 75% in 2002 [10].

Today for increasing seat belt usage, Turkish public agencies using a slogan, which is "To be tied to life" to invite the drivers use the seat belt.

In Eskisehir, at 4 important city center junctions, a seat belt wearing observation has been made. The study was performed by Anadolu University Highway and Traffic Department students on 27th December 2004. 929 driver observed in the study and only 85 (9.15%) of them was wearing seat belt. At the same study 350 front seat passenger has observed and only 23 (6.57%) of them was wearing seat belt.

4. Benefits of using seat belt

Seat belts are very effective in reducing the severity of injuries and the number of fatalities in crashes: About 6000 fatalities and 375 000 injuries could be prevented by universal use (100% of vehicle occupants). The analysis concluded that the maximum achievable seat belt use rate would be 95% with intensive and sustained enforcement and publicity campaigns, resulting in a reduction of 4300 fatalities and 275 000 injuries [11].

When seat belts are used correctly, they reduce the risk of fatal injury by 40-65 percent and the risk of severe injury by 40-50 percent [10].

Several studies on the benefits of seat-belts for drivers and front-seat passengers have found that seat-belts can reduce the risk of all injuries by 40–50%; of serious injuries by 43–65%; and of fatal injuries by 40–60%. Table 7 shows their effectiveness in various types of crash. They are, for example, highly effective in frontal crashes, which are the most common kind of crash and often result in serious head injuries. Their effectiveness for people in front seats is reduced if passengers in rear seats are not also wearing seat-belts or if there are unrestrained objects, such as luggage, in rear seats [9]. Also, as shown in table 1, effectiveness of seat belt, changes according to accident type [12].

<table>
<thead>
<tr>
<th>Crash type</th>
<th>Proportion of all crashes (%)</th>
<th>Driver seat-belt effectiveness in different crash types (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frontal</td>
<td>59</td>
<td>43</td>
</tr>
<tr>
<td>Struck side</td>
<td>14</td>
<td>27</td>
</tr>
<tr>
<td>Non-struck side</td>
<td>9</td>
<td>39</td>
</tr>
<tr>
<td>Rear</td>
<td>5</td>
<td>49</td>
</tr>
<tr>
<td>Roll-over</td>
<td>14</td>
<td>77</td>
</tr>
</tbody>
</table>

Table 1. Injury reduction effects of seat-belts for various types of car crash
5. The public survey

To investigate the causes of this low ratio of seat belt and to produce some arrangement suggestions to increasing the ratio a public survey planned. The survey was performed at 12 important parking lots of Eskisehir by Anadolu University Highway and Traffic Department students between 30th May 2005 and 15th June 2005. 334 drivers questioned in the study. 70 of them were female and 175 of them graduated from university. Cause of selection of parking lots places, 52.40% of the drivers was found as graduated from a university. This ratio is only 7.80% in Turkey (for age ≥25) [13]. So, usage of the seat belt was very high according to other studies. Total usage of seat belt is 36%. The survey results are given with figures below. Questions are below the figures.

![Figure 2. Why don’t you wear seat belt? (Asked to only to non-users, 215 drivers)](image)

![Figure 3. As think of you, what’s the least speed for save the life of seat belt?](image)

![Figure 4. As think of you, seat belt necessary?](image)

![Figure 6. As think of you, Police enforcement is true or false?](image)
Very good 6%
Intermediate 42%
Insufficient 51%

Figure 8. What’s the sufficiency of police controls?

Yes 25%
No 75%

Figure 7. If Police increase the controls, do you use the seat belt often?

Figure 5. Have you ever pay a money penalty for non-use of seat belt?

Figure 9. As think of you, if everyone use the seat belt, how much fatalities decrease?
6. Conclusions

The traffic safety in developing countries is one of the most important problems. Particularly the results of traffic accidents are more serious according to developed countries, since the traffic rules aren’t carried out. The gains of the countries that have comprehended the importance of traffic safety at earlier stages and made arrangements in the direction of this understanding are undisputedly obvious.

The benefits of seat belts are seen in all the accident categories. However, their benefits in crashes/crashing accidents where death and serious injuries frequently occur are much more striking. By means of seat belt, fatalities and injures are decrease to nearly %50 in traffic accidents.

As shown in figures, firstly an education campaign needed to teach of benefits of seat belt. Many Turkish drivers don’t wear their seat belts, because they have wrong information, about seat belt. Then police controls should be increased. Lack of control and education about seat belt causes non-use of seat belts and so fatalities and injures occurs.

References


TRANSPORT NETWORK DESIGN UNDER RISK-AVERSE TRAFFIC ASSIGNMENT: A GAME THEORY APPROACH

Liam O’BRIEN1, Wai Yuen SZETO2, Margaret O’MAHONY3

Abstract. The Continuous Network Design Problem (CNDP) is to determine how best to enhance the capacity of an existing network in a manner which makes efficient use of resources in response to a growing demand for travel. Traditionally the CNDP is posed as a bi-level problem in which link flow is assigned to different routes under a user equilibrium traffic assignment without considering the risk-taking behaviour of users. To capture this behaviour, in this paper we pose the CNDP under a risk-averse traffic assignment through the game theory approach. The CNDP under risk-averse assignment is formulated as a single-level mathematical program by expressing the lower level risk-averse traffic assignment as a set of constraints. This formulation is used to solve for solutions in a simple scenario, and the result shows that when the system is congested, considering uncertainty introduces a higher capacity expansion and a higher toll charge than without. This seems to indicate that ignoring uncertainty will overestimate the network performance and lead to less investment in the network.

1. Introduction

The Network Design Problem (NDP) models the process of determining optimal capacity improvements to a network while accounting for the route choice responses of the users. Comprehensive reviews of the NDP have been carried out in the past, these include:

1 Centre for Transportation Research (TRIP), Department of Civil, Structural, and Environmental Engineering, Trinity College Dublin, Dublin 2, Ireland. obrienli@tcd.ie

2 Centre for Transportation Research (TRIP), Department of Civil, Structural, and Environmental Engineering, Trinity College Dublin, Dublin 2, Ireland. szetow@tcd.ie

3 Centre for Transportation Research (TRIP), Department of Civil, Structural, and Environmental Engineering, Trinity College Dublin, Dublin 2, Ireland. margaret.omahony@tcd.ie
Magnanti and Wong [10]; Boyce [5]; Friesz [7]; and more recently Yang and Bell [13]. However, many of the studies on the NDP to date assume that the network is fixed and certain. In reality, this is often not the case. There are many different events which can cause the network to degrade. These range from events of high impact such as earthquakes, which can cause multiple link failures, or events of lower impact such as road repairs or heavy congestion which may damage certain individual links. Regardless of their magnitude, some or all of these events do happen virtually everywhere from time to time. Since these events threaten the network performance they cause any optimal capacity enhancements introduced under a network improvement scheme to become sub-optimal, if the network uncertainty is ignored.

There are many approaches to consider network design under uncertainty. Some of the previous efforts include Waller et al. [12]; Lo and Tung [8]; Chen and Yang [6]; Lo et al. [9]. However, the existing approaches require knowledge of the link failure probabilities, and some of them do not consider the risk-taking behaviour, which is the behaviour of transport network users in making a trade-off between the travel cost (including travel time, early or late arrival penalty, and so on) and its travel cost uncertainty due to network uncertainty (Yin et al. [14]). In this paper, we adopt the game theoretic approach proposed by Bell and Cassir [4], to model this behaviour in traffic assignment, known as risk-averse user equilibrium (RAUE) traffic assignment. This approach enjoys some advantages over the more conventional approaches. The game theoretic approach assumes that network users are highly pessimistic or naturally very negative about the state of the road network that lies ahead. Therefore, the users tend to behave in a very risk-averse way. This model of route choice behaviour, in turn leads to a very cautious network design with adequate capacity to deal with events which threaten its performance. Another advantage to this approach is that knowledge of link performance frequency distributions (such as delay, travel time or capacity) is not required, unlike in some of the more conventional approaches. This information is very often either absent or not accurate enough to be used (Bell [3]). One more advantage is that it can capture the risk-taking behaviour of users, which is more realistic than the simple user equilibrium approach to model route choice.

In this paper, we incorporate the risk-averse assignment into the NDP. This allows us to explicitly consider the network uncertainty. With this formulation, we can design a network with an expected performance that is more accurate and reflective of real-life scenarios and possibilities such as link degradation. We consider only the continuous version of the NDP (CNDP) for comparison purposes, but this version can be easily extended to the discrete case. We formulate the CNDP under RAUE assignment as a single-level mathematical program by expressing the lower level risk-averse traffic assignment as a set of constraints, and compare it to the CNDP under the normal UE assignment in terms of link tolling and improvement strategies. The result demonstrates that when the system is congested, considering uncertainty introduces a large improvement and charges more than without. Ignoring uncertainty tends to overestimate the network performance and lead to less investment in the network. The rest of this paper is organised as follows: Section 2 describes the formulation of the CNDP under RAUE assignment; section 3 provides some numerical studies on a test network; section 4 reports some conclusions.
2. Formulation

A general transportation network with multiple links, routes, and Origin-Destination pairs is considered. There are $N$ homogenous players between each OD pair. There is also a demon intent on causing maximum delays (and hence increased costs) to the travel players by damaging one and only one link on the whole network. This assumption can be easily relaxed in the future but we consider one demon just for the ease of illustration.

2.1. Risk-Averse Traffic Assignment Constraints

The approach used to model the CNDP under risk-averse user equilibrium assignment (CNDP-RAUE) is the game theory approach, as formulated in Bell and Cassir [4]. The lower level problem comprises two sub-problems: the user problem and the demon problem. The user problem can be viewed as a non-cooperative game in which each homogenous player tries to select the route with minimum expected trip cost.

The user problem can be approximated to deterministic traffic assignment when the number of homogenous players is large (Bell and Cassir [4]). The first-order condition of deterministic traffic assignment can be expressed as the following non-linear complementarity conditions:

$$h_j \left[ \sum_k g_{jk} \delta_{jk} g_k - \sum_{rs} \delta_{rs} \left( \min_m g_m (h) \right) \right] = 0, \quad (1)$$

$$\sum_k g_{jk} \delta_{jk} g_k - \sum_{rs} \delta_{rs} \left( \min_m g_m (h) \right) \geq 0, \quad (2)$$

where $h_j$ is the flow on route $j$; $h$ is the route flow vector; $g_{jk} \delta_{jk} g_k$ represents the cost of route $j$ in scenario $k$ based on the flow vector $h$; $q_k$ is the probability of damage to a link in scenario $k$; $\delta_{rs}$ is the route-OD incidence indicator. $\delta_{rs} = 1$ if $j$ connects OD pair $rs$, and $\delta_{rs} = 0$ otherwise.

Traffic assignment in every scenario includes flow conservation and non-negativity conditions, expressed as:

$$N^{rs} = \sum_j \delta_{rs} h_j, \quad (3)$$

$$h_j \geq 0, \quad (4)$$

where $N^{rs}$ is the travel demand of OD pair $rs$. If route $j$ connecting OD pair $rs$ carries flow ($h_j > 0$), according to (1), the corresponding expected route travel cost $\sum_k g_{jk} (h) \delta_{jk} q_k$ must be equal to the minimum expected travel cost $\sum_{rs} \delta_{rs} \left( \min_m g_m (h) \right) \delta_{rs} \delta_{jk} q_k$ between OD pair $rs$, as the term in the square bracket in (1) must equal zero. If route $j$
carries no flow \((h_j = 0)\), the corresponding expected route travel cost must be greater than or equal to the minimum expected travel cost based on (2).

Route expected travel costs and lowest expected travel costs can be expressed as a function of the route flow pattern \(h\), capacity improvements \(y\), and tolls \(\rho\):

\[
v_a = \sum_j \delta_{ja} h_j, \quad (5)
\]

\[
T_{ak}(v_a) = \frac{v_a}{c_{aik}} \left[1 + \frac{0.15}{v_a} \left(\frac{v_a}{c_{aik}}\right)^4\right], \quad (6)
\]

\[
\tilde{c}_{a,k} = \bar{\lambda}_{a,k} (c_a + y_a), \quad (7)
\]

\[
g_{jk}(h) = \sum_a \left(\psi \delta_{ja} t_{ak} (v_a) + \rho_a\right), \quad (8)
\]

where \(v_a\), \(\rho_a\) and \(y_a\) are the flow, toll and capacity improvement on link \(a\), respectively; \(T_{ak}(v_a)\) and \(\tilde{c}_{a,k}\) are the travel times and capacity of link \(a\), in scenario \(k\) respectively; \(g_{jk}(h)\) is the cost of route \(j\) in scenario \(k\); \(\delta_{ja}\) is a link-route incidence indicator, equal one if link \(a\) is on route \(j\) and equals zero otherwise; \(t_{ak}\) and \(c_a\) are the free-flow travel time and initial capacity of link \(a\); \(\bar{\lambda}_{a,k}\) is a link-damage indicator which equals one if there is no damage to link \(a\) in scenario \(k\) and equals \(\frac{1}{2}\) if there is damage to link \(a\) in scenario \(k\); \(\psi\) is the unit cost of travel time.

Equation (5) states that the link flow is obtained by summing the corresponding route flows on the link. Equation (6) is the Bureau of Public Roads (BPR) link performance function modified here to include different potential scenarios of link damage. Equation (7) expresses the capacity of link \(a\) as the sum of the initial link capacity plus the capacity improvements. Equation (8) computes the cost of route \(j\) in scenario \(k\) based on the total cost on every link on that route.

The demon problem depicts the objective of the demon who seeks a mixed strategy to maximize the total expected trip cost to the users. The problem can be written as the following non-linear complementarity conditions:

\[
q_k \left(\max_w \sum_j g_{wj}(\bar{h}) h_j - \sum_j g_{j}(h) h_j\right) = 0, \quad (9)
\]

\[
\max_w \sum_j g_{wj}(\bar{h}) h_j - \sum_j g_{j}(h) h_j \geq 0, \quad (10)
\]

\[
q_k \geq 0, \quad (11)
\]

where the probability \(q_k\) must satisfy:

\[
\sum_i q_i = 1, \quad (12)
\]

and be non-negative as shown in (11) by definition, and \(g_{j}(h)\) follows the definition in (8). Conditions (9)-(10) actually represent the necessary and sufficient conditions of the Nash equilibrium of a non-cooperative, mixed-strategy game for the demon (see Nash
According to (9), if scenario $k$ is selected ($q_k > 0$), the maximum total expected network cost, $\max \sum \mathcal{g}_{jw}(\mathbf{h})h_j$, must be equal to the total expected network cost in scenario $k$, $\sum \mathcal{g}_{jw}(\mathbf{h})h_j$, through the condition \[ \max \sum \mathcal{g}_{jw}(\mathbf{h})h_j - \sum \mathcal{g}_{jw}(\mathbf{h})h_j = 0. \] If scenario $k$ is not selected ($q_k = 0$), the maximum total expected network cost must be greater than or equal to the total expected network cost in scenario $k$ according to (10).

### 2.2. Cost and Revenue Functions and Cost Recovery Constraints

Exact cost recovery is considered in our formulation to make all improvements and tolling strategies more realistic than they might otherwise be. For simplicity we do not consider subsidies or the case of full cost recovery (where there is a profit/surplus at the end). Exact cost recovery occurs when the revenue exactly covers the construction cost. The exact cost recovery (ECR) constraint may be expressed as:

\[ R_i - K_i = 0, \quad (13) \]

where $R_i$ and $K_i$ are, respectively, the toll revenue and improvement cost.

The toll revenue $R_i$ can be expressed as a function of the equilibrium link flows and tolls, and the improvement cost $K_i$ can be expressed as a function of improvements:

\[ R_i = \sum_a \rho_a \left( n \nu_a \right), \quad (14) \]
\[ K_i = \sum_a u_a \left( y_a \right), \quad (15) \]

where

\[ u_a \left( y_a \right) = b_{a,0} + b_{a,1} y_a, \quad (16) \]

$u_a \left( \cdot \right)$ is the improvement cost function; $b_{a,0}$, $b_{a,1}$ are parameters of this cost function; $n$ converts hourly link flow to annual link flows.

### 2.3. Link Improvement and Toll Constraints

In practice highways rarely exceed more than a few lanes in size. There are very often practical and physical constraints on the extent of any road widening. Furthermore a planning authority may decide for practical or political reasons to only improve and toll certain links in the network. Mathematically these constraints may be expressed as:

\[ c_{a,k} \geq u_a, \quad (17) \]
\[ y_a \geq 0, \quad (18) \]
\[ y_i = 0, \quad (19) \]
\[ \rho_i = 0, \quad (20) \]
\[ \rho_a \leq \rho_{\text{max}}, \quad (21) \]
The total capacity of each link, \( c_{a,k} \) in scenario \( k \) is limited to be less than its maximum allowable capacity \( u_a \), by (17). Link capacity improvements are constrained to be non-negative by (18). Links without improvements and without tolls are described by (19) and (20), respectively. Link tolls cannot be set too high as in (21) where the toll on link \( a \), \( \rho_a \), must be less than or equal to the maximum allowable toll charge, \( \rho_{\text{max}} \).

### 2.4. The Objective Function

For simplicity we adopt an objective function with a straightforward interpretation. We consider the minimization of expected total system travel time (ETSTT). ETSTT describes the expected total travel time spent by users in the network and is very straightforward to compute given the link flows and damage probabilities. Mathematically, it can be expressed as:

\[
ETSTT = \sum_{r} \sum_{s} \sum_{x} v_x(t_x(v_x)) q_x
\]

where \( v_x \) is the link flow as given by equation (5); \( t_x(v_x) \) is the link travel time in scenario \( k \) given in equation (6); \( q_x \) is the probability of damage to a link in scenario \( k \). The product \( v_x(t_x(v_x)) q_x \) is evaluated for every link in each scenario for each origin-destination pair (represented by the subscript \( rs \)) in the network.

The CNDP under risk-averse traffic assignment can then be formulated as:

\[
\min_{h,y,\rho,q} ETSTT = \sum_{r} \sum_{s} \sum_{x} v_x(t_x(v_x)) q_x
\]

subject to:
- the risk-averse traffic assignment constraints (1)-(12)
- the cost-revenue and cost recovery constraints (13)-(16)
- the link improvement and toll constraints (17)-(21)

We have formulated the CNDP under risk-averse assignment and we will demonstrate in the next section its performance in relation to the CNDP under normal user equilibrium assignment (see for example Abdullaal and Le Blance [2]). To obtain solutions to both models this study adopts the Generalized Reduced Gradient (GRG) algorithm [1].

### 3. Numerical Study

For the purposes of illustration, a scenario is set up using a simple network of one Origin-Destination (OD) pair with two parallel links (see figure 3.1). We use this network to compare the different tolling and improvement strategies of the proposed CNDP-RAUE model versus the traditional CNDP-UE model. The following parameters apply:

- Demand Parameters, \( N^{12} = 2000 -12000 \) (increasing in increments of 2000)
- Network Parameters: \( t^0_1 = 10 \) mins., \( t^2_1 = 15 \) mins., \( c_1 = c_2 = 4000\text{ veh/hr} \)
- Parameters of Improvement cost functions: \( h_{1,0} = h_{2,0} = 10000 \); \( h_{1,1} = h_{2,1} = 10000 \)
Value of Time: $\psi = €0.2/\text{min}$
Improvement on link 2: $y_2 = 0$
Toll on link 2: $\rho_2 = 0$

Figure 3.1: Test Network
To clearly illustrate the results without interference from other factors we choose a simple two links network and allow only link 1 to be improved. Under this scenario setting we can more clearly demonstrate the different improvement strategies of the two models. Figure 3.2 shows the different capacity expansions introduced by the proposed and traditional CNDP models for an arbitrary demand range of 2000-12000.

For very low levels of demand (in the range 2000-4000) we can see that the CNDP under UE assignment introduces marginally higher capacity expansions. However these differences appear to be almost negligible. For all other higher demand levels shown here, the CNDP under RAUE assignment consistently introduces greater capacity expansions. In fact for the highest demand level shown here of 12,000 users, the capacity improvements introduced by the CNDP-RAUE model are almost 5 times that of the CNDP-UE model. This simple example illustrates the markedly different improvement strategies of the two approaches and underlines the importance of taking network uncertainty into account, especially when the demand level is high.

Figure 3.2: The improvement strategies of the two models

Figure 3.3 highlights the tolling strategies of the two different approaches. Again we can see that for a very low demand level of 2000 users, the CNDP-UE model charges a marginally higher toll. Of course the reason for this is obvious, since for the same demand level as the CNDP-RAUE model, it must charge a slightly higher toll to cover the cost of the additional improvements, in order to maintain cost recovery. Apart from that, the trend
is very similar to the improvement strategy, with the CNDP-RAUE model consistently charging a higher toll, where the difference in the toll levels becomes even greater with increasing demand levels. This is easily explained by the fact that, in order to maintain cost recovery and fund the improvements, the CNDP-RAUE model needs to generate more revenue through its only source of funding - tolling. This simple study highlights the effect on toll levels of using the CNDP-RAUE model over the traditional model. It suggests that the higher capacity expansions introduced when uncertainty is considered in the network have to be funded at a higher cost to the user. Of course if uncertainty is ignored then we can see that the user will pay considerably less but then there is a trade-off in the expected network performance (which is not shown here due to the space limitation).

4. Conclusions

This paper developed a single-level network design model capturing the risking taking behaviour and network uncertainty in network design problems using the game theory approach. This approach requires no knowledge of the link failure probabilities and hence requires few parameters to calibrate than there would be with some of the other approaches, which is an advantage from a modelling perspective. Compared with the traditional CNDP-UE model, the proposed model provides a more realistic assessment of network improvements by considering the network uncertainty. A numerical study is set up to illustrate that when the network is congested, considering uncertainty introduces a higher capacity expansion and a higher toll charge than without. This seems to indicate that ignoring uncertainty will overestimate the network performance and lead to less investment in the network.

Acknowledgement

This research is funded under the Programme for Research in Third-Level Institutions (PRTLI), administered by the Higher Education Authority.
References


SITLUM – THE STRATHCLYDE INTEGRATED TRANSPORT/LAND-USE MODEL

Andrea ARAMU¹, Andrew ASH², Jim DUNLOP³, David SIMMONDS⁴

Abstract. The Strathclyde Integrated Transport/Land-Use Model (SITLUM) was developed by TRL and David Simmonds Consultancy (DSC) for Strathclyde Passenger Transport (now Strathclyde Partnership for Transport). SITLUM links an earlier application of TRL’s Strategic Transport Model (STM) package for the Strathclyde area (centred on Glasgow) to a new implementation of DSC’s DELTA land-use modelling package. The paper describes the design of the model and presents some example results.

1. Introduction

As part of their strategic modelling capability, Strathclyde Passenger Transport (SPT) have, for some years, used the TRL’s Strathclyde Strategic Transport Model (SSTM) as a “policy filter” in conjunction with a large scale network model, the Strathclyde Integrated Transport Model (SITM). SSTM has also been used to assist policy formulation by the Structure Plan Teams in the Glasgow and Clyde Valley and in Ayrshire. To allow SPT to extend this into modelling the interaction of land use and transport processes, TRL, as lead consultants, and David Simmonds Consultancy (DSC) were commissioned in 2003 by SPT to design and implement the Strathclyde Integrated Transport and Land-Use Model.

¹ David Simmonds Consultancy Limited, Suite 14 Millers Yard, Mill Lane, Cambridge CB2 1RQ, England, aa@davidsimmonds.com
² Transport Research Laboratory, Crowthorne House, Nine Mile Ride, Wokingham RG40 3GA, England, aash@trl.co.uk
³ Strathclyde Partnership for Transport, Consort House, 12 West George Street, Glasgow G2 1HN, Scotland, jim.dunlop@spt.co.uk
⁴ David Simmonds Consultancy Limited, Suite 14 Millers Yard, Mill Lane, Cambridge CB2 1RQ, England, dcs@davidsimmonds.com
(SITLUM). This integrates the state-of-the-art DELTA land-use/economic model with an enhanced version of the SSTM.

Both SSTM and DELTA exist within the SITLUM system as separate programs. Both SSTM and DELTA are run in sequence over the range of years to be modelled. During this process they exchange information. SSTM provides DELTA with accessibility (travel costs) and trip data for movements between study area zones; DELTA generates planning data (influenced by travel costs, taking into account time lag effects) to SSTM. The system also consists of a third program: the SITLUM user interface. This is an extension of an earlier SSTM which allows the user to create test scenarios, create and initiate program control files and view both transport and land-use outputs from a graphic user interface.

2. DELTA

Work on development of DELTA started in 1995 in response to a perceived demand for a new “land-use” modelling package with two key characteristics. The first, practical characteristic was that the model should be suitable for use as an add-on to otherwise free-standing transport models, in particular to “strategic” transport models [3]. The second, theoretical requirement was that the model be constructed in terms of processes of change, drawing on the enormous range of research carried out in urban and regional economics, geography, demography, sociology, etc. Further detail of the original design is published in reference [4] and a recent update in [1].

2.1. Land-use/transport interaction using DELTA in SITLUM

The overall design of a DELTA-based model consists of four components, as illustrated in Figure 1, namely

- the transport model (to which DELTA is linked);
- the economic model;
- the urban land-use model;
- the migration model.

Of these, the transport and urban models work at the level of zones, whilst the migration and economic models work at the broader level of areas. Areas typically correspond to travel-to-work areas, at least within the region of main interest; zones represent finer units within these areas (or within the area we are concentrating on).

The transport model takes inputs which describe activities (different categories of residents and jobs) by zone, for a given year. From this and from input transport system data it forecasts travel by car and by public transport. In doing so, it estimates costs and times of travel between each pair of zones, allowing for congestion caused by the forecast traffic.

The economic model forecasts the growth (or decline) of the sectors of the economy in each of the areas modelled. Its inputs include forecasts of overall growth in output and productivity. The forecasts by sector and area are influenced by

- costs of transport (from the transport model)
- consumer demand for goods and services (from the urban model)
• commercial rents (from the urban model).
Forecast changes in employment by sector and area are passed to the urban model. Freight transport outputs can be passed to the transport model.

The urban model forecasts the zonal location of households and jobs within the areas that are modelled in detail. Locations are strongly influenced by the supply of built floorspace. Locations are also influenced by accessibility, with different measures of accessibility influencing different activities, and by environmental variables. Households are influenced by accessibility to workplaces and services. Businesses are influenced by accessibility to potential workers and customers.

The locations of households and jobs are fed back to the transport model to generate travel demands. Household numbers are also used to calculate consumer demand for goods and services in each area, for use in the economic model. The rents arising from competition for property in each area affect both the economic and migration models. Information on job opportunities is passed to the migration model.

The migration model forecasts migration between areas within the modelled area. (Movements within areas are forecast in the urban model.) The inputs to this model include job opportunities and housing costs, from the urban model. Job opportunities are a strong incentive to migration; housing costs are a generally weak disincentive.

There are complex possibilities for feedback between the four components outlined above. For example, it is possible for an improvement in transport to generate economic growth, which generates additional travel, which may cause increased congestion and some worsening in transport conditions.
2.1.1. Sub-models within DELTA

The original DELTA package was intended to model a single compact area, with a given economic and demographic scenario for the total change in that area. It therefore consisted solely of the urban model (linked to a transport model), which at first consisted of five sub-models:

- the Transition and growth sub-model, dealing with household/population change and employment growth factors;
- the Location and property market sub-model;
- the Employment status and commuting sub-model;
- the Development sub-model;
- the Area quality sub-model.

The name DELTA was discovered as an acronym for these five sub-models. The car-ownership sub-model was added in the application to Greater Manchester [2].

The economic and migration models were added to allow the model system to represent either the wider interactions between a city area and its neighbours (without modelling other cities in any detail) or to model a region containing a number of urban areas. The main linkages between the sub-models within a one-year period are shown in Figure 2. This allows SITLUM to take account of the interactions between the main area of interest and the rest of Scotland.

![Figure 2 Sub-model sequence within a DELTA one-year cycle](image)

3. STM

“STM” refers to the software package; “SSTM” to the Strathclyde application. The STM software was originally developed at TRL to test multi-modal transport policies in the
Greater London area. The key design objective was a high-speed desktop tool which planners and policy makers could use to assess the potential impacts of policy levers applied either singly or in combination. Users would be able to identify, at an early stage, promising strategies on which more detailed investigation could be concentrated.

STM was therefore designed to model travellers’ responses to the many different components of generalised travel cost, but without using a detailed representation of transport networks and route choice which accounts for a substantial part of the run time of large-scale models. An STM study area usually has a relatively smaller number of zones when compared with large assignment models; 173 zones have been used in SITLUM to model the Strathclyde / Ayrshire area. To incorporate long range movements and land-use interactions, the zonal system extends over the whole of Scotland; the transport model, SSTM, also includes a zone for England and Ireland.

The modelled modes include car, public transport and non-motorized modes. There is also provision for the modelling of new Park and Ride schemes, and new Light Rail links and the impact of new railway stations. STM trip purposes cover home-based purposes and the impact of new railway stations. STM trip purposes cover home-based purposes (work, education, social/leisure, etc) and non-home-based trips. Two time periods are modelled: the am peak and interpeak. Levels of travel demand in STM are driven by planning data describing the spatial distribution and composition of population, jobs and car ownership. In SITLUM these are provided by DELTA. SSTM can also be run within SITLUM without interacting with DELTA: either pre-existing DELTA outputs can be used or exogenous planning projections can be created within the STM software using TEMPRO growth factors or they may be user supplied.

To increase robustness, STM is “incremental”, i.e. it uses demand models (distribution/modal split) which derive future patterns of modal share and distribution from cost changes and the corresponding patterns in a base year. Highway travel times are calculated in STM using zonally-based speed flow relationships for different road types. Road and rail trips follow routes through the zonal system based on the actual networks. Provision now also exists for using different highway routings in forecast years arising from changes in the road networks. STM iterates so as to converge towards equilibrium between demand levels and supply costs.

The user can apply a range of transport policy levers over modelled forecast years: cordon charges, parking, public transport supply and fares; as well as explore the implications of scenario changes in terms of demography and car ownership, real earnings and fuel prices. STM can then run in “time-marching” fashion over the specified forecast years to provide outputs for those years, in interaction with DELTA.

A variety of STM outputs are available relating to congestion, modal shift, emissions and road safety. Cost-benefit calculations can also be enabled. A striking feature of STM is its Graphic User Interface (GUI) with an interactive map which gives exportable displays of changes in policy impact measures for comparison scenarios and GIS-style “thematic” mapping of outputs – the latest version of STM (used in SITLUM) also allows land use variables to be displayed.

4. SITLUM application

Figure 3 shows the study area of the Strathclyde Integrated Transport/Land-Use Model.
4.1. Database files

The initial SITLUM population and household database files have been created using the results of the 1991 and 2001 census. The major task in processing the census data is to disaggregate, for each SITLUM zone individually, the households and their members into the household and person categories used in this application of DELTA. The household classification takes account of household composition/age and socio-economic group; the members of these households are identified as children, working, non-working or retired. Employment is categorised by sector and by socio-economic group. The initial database also measures the floorspace available to these activities in each zone, in terms of quantity by type (residential, retail, office etc), the rental value of each type in each zone, the quantity of each type that is vacant, and, for housing, a quality measure.
It is important to note that the base year (2001) situation is directly input to the model and does not have to be reproduced within the model; the model forecasts the changes over time from that situation.

4.2. The economic and demographic scenario

The SITLUM inputs that create the economic and demographic scenarios have been defined so as to reproduce the rates of change in economic activity, employment and population that are used in other aspects of Scottish planning. The concern in using SITLUM is to understand how land-use and transport plans will affect the distribution of employment and population within Scotland.

4.3. Reference case and example results

The base year in the SITLUM application is 2001 and the reference case is run to produce a forecast period of 20 years, until 2021. The following policy test has been devised by DSC purely to show the response of the system to such inputs and does in no way represent policy of any other party.

The policy that is implemented is a new, very high quality LRT route heading east from central Glasgow reaching as far as Airdrie town centre and serving all eight zones that the route passes through. Figure 4 shows this “Airdrie Corridor”. The hypothetical LRT scheme is assumed to start operation in 2005. In addition, from 2005 development of an additional 7000 dwellings is allowed along the corridor. The following inputs/options have been used to define the LRT scheme:
Figures 5 and 6 show summary indicators of the scale and pattern of the main land-use impacts. A wide variety of other indicators are available. The ability to examine the interaction between land-use and transport measures, and to forecasts their impacts over time, is proving valuable in planning practice.

References


ANALYSIS OF THE FACTORS AFFECTING A REGIONAL AIRPORT CATCHMENT AREA

Maria Nadia Postorino, Alberto Andreoni

Abstract: In this paper an analysis of the airport catchment area is made on the basis of two survey campaigns realized at the regional airport of Reggio Calabria. The surveys refer to two different supply conditions mainly in terms of served destinations and fares, thus allowing a kind of before-and-after analysis to be performed. More factors that are supposed to affect the size of the catchment area have been taken into account and their impact has been examined.

1. Introduction and problem statement

The air transport demand at a given airport represents an economic source for both airlines and airport managers, given that the first ones have an economic convenience to offer more flights and more served destinations if there is a minimum amount of air demand to cover their operational costs while the second ones can gain more from passengers/freight and airlines taxes that are computed on the basis of the demand amount.

Given some recent tendencies in terms of airport management rules, it is crucial for each airport manager to attract passengers (but also freight) that in turn means more airlines and offered services.

The purpose of an airport manager is then to move airlines and users (passengers and freight) towards its airport by using various strategies as:

- improvement of land accessibility;
- increase of flight frequencies;
- more served destinations;
- suitable departure time intervals to meet the users requirements;
- availability of competitive air fares;
- good handling systems.

1 DIMET Mediterranea University of Reggio Calabria (Italy), Address Facoltà di Ingegneria, via Graziella, Località Feo di Vito, 89100, Reggio Calabria, Italy, npostorino@unirc.it; alberto.andreoni@unirc.it
The potential geographic area for drawing passengers represents the airport catchment area (Government Accountability Office, [2]). The size of a catchment area varies from airport to airport depending on factors such as how close an airport is to other airports and whether the airport is served by a low-fare airline (and, therefore, it could be attractive to passengers from farther away). The physical individuation of the airport catchment area is not an easy task, and there are different models and methods proposed in the literature to identify it. The airport catchment area is often identified by the number of users that start their trip at the airport, without a specific reference to their geographic origin. However, information on the geographical origins is useful to help airport managers to identify the best strategies for the airport developments (e.g., the budget being fixed, a choice could be made between investments in accessibility rather than in handling systems, and so on).

Another simple, empirical method to fix an airport catchment area is to identify it as the number of users in the geographic area close no more than a given time amount to the airport itself. In other words, this is a measure of accessibility in terms of time to reach the destination (i.e. the airport), where the superior boundary is fixed. As an example, some important European airports consider the accessibility area within two hours by car [4], while in U.S.A. some studies have fixed a boundary of sixty minutes by land vehicles [3]; [1].

In this work different factors are considered to quantify their impact on the catchment area of a regional airport, particularly to analyse if a variation in their values can significantly modify the extension of the catchment area both in terms of physical size and number of users. Particularly, frequencies, kind of aircraft, accessibility and fares have been taken into account for this analysis, performed at the “Tito Minniti” Reggio Calabria regional airport, that could become a central hub for the Mediterranean basin, primarily because of its strategic geographic position serving two known tourist destinations in Italy (Calabria and Sicily) and secondly because of the lack of fast land services linking these regions to the middle Italy and then to Europe.

Till the autumn 2005, the airport had few flights only towards the Italian hub airports, Rome and Milan, operated by two airline companies; particularly, the Rome hub airport was served by only one company under a monopoly system that implied high fares, low flexibility with respect to national special offer campaigns and low frequencies. Furthermore, only hub-and-spoke services were guaranteed, with additional disadvantages for users in terms of both connection waiting time and opportunity to choose different airline companies.

From the end of December 2005, thanks to the efforts of the airport management staff, two low-cost carriers have set up new point-to-point links to Venice, Turin and Bologna and at the beginning of 2006 other links have been started towards Bergamo, Genova, Pisa, Malta and also Roma, thus breaking the initial monopoly system.

Recently, a new weekly international link to Cracovia has been realized thus connecting the town of Reggio Calabria to one important Eastern European city. In October 2005, there were two served destinations with six daily departing flights and 942 offered seats, whereas from March 2006 there are ten served destinations (seventeen daily departing flights) and 1850 offered seats; the comparison of traffic data between February 2006 and February 2005 shows an increase of 38,5 per cent in the number of planned/unplanned passengers (SOGAS, www.sogas.it; ASSAEROPORTI, www.assaeroporti.it).

Furthermore, the airport accessibility has been improved by introducing a bus service, called “Jonica Line”, linking the airport to the various towns of the Ionian strip, and a
combined ship-bus service, called “Volobus”, linking the airport to the city and province of Messina. “Volobus” guarantees connections to flights and during the trip the users can make hand-luggage check-in and get all the information they require about the flight and the airport. Increase of airport accessibility is expected to get more demand at the airport, as some studies in the literature suggest (e.g., Yao and Morikawa [5] point out that the people tend to increase their travel frequency if accessibility improves). This study was partially developed within the research project on “Guidelines for the planning of the development of Italian regional airports”, financed by the Italian Ministry of Higher Education and Research.

2. Data collection

The analysis carried out can be considered a before and after analysis by using data from two RP surveys realized at the Reggio Calabria airport: the first one has been done from 27 to 31 December 2005, when the links to Venice, Turin and Bologna started, while the second one has been realized from 29th March to 4th April 2006, when the links to Bergamo, Pisa and Genova added to those already offered as well as the new flight to Rome, operated by a second company in competition with the initial, monopolistic one. Data of the first survey refer to a before analysis, where three new connections have been established with respect to the initial condition (only hub-and-spoke links) for which, unfortunately, there were no data; data of the second survey refer to an after analysis, giving that new links and possible competition among companies have been introduced. Generally, there could be an initial induced effect of demand growth when new services are supplied (e.g., due to low fare offers to launch the new links), followed by a demand adjustment period till an equilibrium between actual user needs and standard offered services is reached. Then, the aim of the second survey is also to verify the equilibrium condition between demand and the already established supply with respect to the initial induced effects, if any. Table 1 reports the schedule and characteristics of the flights during the first survey period.

<table>
<thead>
<tr>
<th>Destination</th>
<th>Departure</th>
<th>Arrive</th>
<th>Aircraft model</th>
<th>Capacity</th>
<th>Aircraft type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rome Fiumicino</td>
<td>6.35</td>
<td>8.00</td>
<td>M82</td>
<td>163</td>
<td>Jet</td>
</tr>
<tr>
<td>Turin</td>
<td>7.00</td>
<td>9.30</td>
<td>AT42</td>
<td>44</td>
<td>Turboprop</td>
</tr>
<tr>
<td>Milan Linate</td>
<td>8.00</td>
<td>9.40</td>
<td>M80</td>
<td>131</td>
<td>Jet</td>
</tr>
<tr>
<td>Venice</td>
<td>10.35</td>
<td>12.10</td>
<td>F100</td>
<td>94</td>
<td>Jet</td>
</tr>
<tr>
<td>Rome Fiumicino</td>
<td>11.25</td>
<td>12.40</td>
<td>M80</td>
<td>131</td>
<td>Jet</td>
</tr>
<tr>
<td>Milan Linate</td>
<td>14.05</td>
<td>15.45</td>
<td>M80</td>
<td>131</td>
<td>Jet</td>
</tr>
<tr>
<td>Milan Linate</td>
<td>15.35</td>
<td>17.10</td>
<td>B737</td>
<td>167</td>
<td>Jet</td>
</tr>
<tr>
<td>Bologna</td>
<td>16.10</td>
<td>18.10</td>
<td>AT42</td>
<td>44</td>
<td>Turboprop</td>
</tr>
<tr>
<td>Rome Fiumicino</td>
<td>19.10</td>
<td>20.25</td>
<td>M80</td>
<td>131</td>
<td>Jet</td>
</tr>
</tbody>
</table>

Table 1. Departing flights from Reggio Calabria during the first survey period
The sample is made up of 468 interviews; the geographical area where user trip origins are located has been divided into 7 zones: 4 of them correspond to four provinces of the

Figure 1. Distribution of origins in the potential catchment area
Calabria administrative region (Reggio Calabria, Cosenza, Catanzaro and Vibo Valentia); 2 of them include the municipalities belonging to respectively the Ionian and Tyrrhenian strips; the last one includes the users coming from Sicily (mainly Messina and its province). The catchment area of the Reggio Calabria airport is represented in fig.1, for the 7 identified zones.

It is interesting to note that the covered area includes three airports other than that of Reggio Calabria (RC): Lamezia Terme (LT), Crotone (KR), Catania (CT). While the percentage of users coming from areas close to these airports are negligible (less than 1%, see fig.1), however it should be noticed that there are some users willing to cover greater distances to take advantages of point-to-point flights and/or low fares.

The second survey refers to a sample of 1045 interviews and the new supply system reported in table 2.

<table>
<thead>
<tr>
<th>Destination</th>
<th>Departure</th>
<th>Arrive</th>
<th>Aircraft model</th>
<th>Capacity</th>
<th>Aircraft type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fiumicino Rome</td>
<td>6.35</td>
<td>8.00</td>
<td>M82</td>
<td>163</td>
<td>Jet</td>
</tr>
<tr>
<td>Fiumicino Rome</td>
<td>7.20</td>
<td>8.30</td>
<td>B737</td>
<td>125</td>
<td>Jet</td>
</tr>
<tr>
<td>Genova</td>
<td>7.30</td>
<td>9.45</td>
<td>AT42</td>
<td>44</td>
<td>Turboprop</td>
</tr>
<tr>
<td>Fiumicino Rome</td>
<td>8.35</td>
<td>9.40</td>
<td>A320</td>
<td>168</td>
<td>Jet</td>
</tr>
<tr>
<td>Linate Milan</td>
<td>9.15</td>
<td>11.00</td>
<td>M80</td>
<td>153</td>
<td>Jet</td>
</tr>
<tr>
<td>Turin</td>
<td>11.05</td>
<td>12.45</td>
<td>AT42</td>
<td>44</td>
<td>Turboprop</td>
</tr>
<tr>
<td>Fiumicino Rome</td>
<td>11.40</td>
<td>12.55</td>
<td>M80</td>
<td>153</td>
<td>Jet</td>
</tr>
<tr>
<td>Fiumicino Rome</td>
<td>13.00</td>
<td>14.25</td>
<td>AT42</td>
<td>44</td>
<td>Turboprop</td>
</tr>
<tr>
<td>Venice</td>
<td>13.30</td>
<td>15.00</td>
<td>F100</td>
<td>94</td>
<td>Jet</td>
</tr>
<tr>
<td>Linate Milan</td>
<td>14.05</td>
<td>15.45</td>
<td>M80</td>
<td>153</td>
<td>Jet</td>
</tr>
<tr>
<td>Linate Milan</td>
<td>15.35</td>
<td>17.10</td>
<td>B734</td>
<td>167</td>
<td>Jet</td>
</tr>
<tr>
<td>Bologna</td>
<td>15.40</td>
<td>17.05</td>
<td>B737</td>
<td>125</td>
<td>Jet</td>
</tr>
<tr>
<td>Pisa</td>
<td>17.00</td>
<td>18.45</td>
<td>AT42</td>
<td>44</td>
<td>Turboprop</td>
</tr>
<tr>
<td>Bergamo</td>
<td>17.45</td>
<td>19.25</td>
<td>A320</td>
<td>180</td>
<td>Jet</td>
</tr>
<tr>
<td>Fiumicino Rome</td>
<td>19.10</td>
<td>20.25</td>
<td>M80</td>
<td>153</td>
<td>Jet</td>
</tr>
<tr>
<td>Fiumicino Rome</td>
<td>20.00</td>
<td>21.10</td>
<td>B737</td>
<td>125</td>
<td>Jet</td>
</tr>
<tr>
<td>Malta</td>
<td>20.30</td>
<td>21.20</td>
<td>A320</td>
<td>168</td>
<td>Jet</td>
</tr>
</tbody>
</table>

Table 2. Departing flights from Reggio Calabria during the second survey period

As for the first data set, the catchment area has been identified with respect to the 7 prefixed zones. The before-and-after comparison shows very interesting results in terms of the impacts that the factors have both on the extension of the catchment area and the number of potential users. Particularly, while the physical extension of the catchment area is practically the same, the number of users is strongly increased (see table 3). Furthermore, another interesting result refers to a qualitative aspects linked to the “image” of the company: from this point of view, users are willing to pay more to fly with a well
known or hub-based flag company despite the introduction of direct links and low cost fares as well as they prefer to travel by using jets more than propeller aircrafts.

Table 3. Reggio Calabria airport demand trend during the latest months

3. Conclusions

The analysis of the airport catchment area is a crucial feature in the planning of developing policies for regional airports. Important aspects that make an airport to be preferred to another, if there is potential competition in the area, are accessibility, point-to-point rather than hub-based flights, low cost fares, but also the image of the company and the aircraft type.

Starting from this preliminary analysis, future developments are intended to be obtained by using suitable models to represent the expected airport catchment area when variations in the air supply (as frequencies, fares, airport accessibility) have to be simulated.

References


A STUDY ON THE EFFECT OF REMOVAL OF CENTER LINES OF THE ROADS TO EXTEND THE WIDTH OF SIDEWAYS OF THE ROADS

Seiji HASHIMOTO¹, Toshiomi OGURA, Koji IZUHARA

Abstract. The policy to remove centerlines or widen sideways on the roads without footpaths that are supplementary to main lines or are used for daily life is being introduced and progressed to reduce traffic accidents and improve the pedestrians’ environment for walk. This policy has been actively executed since 2000 in Aichi prefecture and such arrangement is being extended to whole country since 2003 but the reality is that there has not been enough verification of the effect of the policy by now. This thesis is to evaluate the effect of the policy on the arranged routes in Toyota city, Aichi prefecture based on objective data such as traffic accidents data, traffic and vehicle speeds and a research on the perceptions of the residents along the arranged routes.

1. The background and purpose of the study

   The policy to remove centerlines or widen sideways on the roads without footpaths that are supplementary to main lines or are used for daily life is being introduced and progressed to reduce traffic accidents and improve the pedestrians’ environment for walk. This policy has been actively executed since 2000 in Aichi prefecture and such arrangement has been extended to whole country since 2003 and is still being progressed. Figure 1 is the image of the arrangement of the policy and the purposes of the arrangement are as follows.

   - Improvement of the environment for pedestrians’ walk and bicycles’ running
   - Decrease of vehicle speed by raising the tension of drivers through the removal of centerlines
   - Decrease of passing traffic
   - Decrease of collision at the cross streets

¹Toyota Transportation Research Institute, 1-1 Wakamiya-cho, Toyota, Aichi, 471-0026, JAPAN, hashimoto@ttri.or.jp
Also as for the effect of the policy, Imoto(2002)\(^1\) has provided the effect of the arrangement on the increase/decrease of short term traffic accidents for some of the routes but substantial evaluations are the subject of future studies. Carol Debell(2003)\(^2\) has introduced the similar project in UK.

This study is to evaluate this policy taking the routes in Toyota city that introduced this policy for the first time in Aichi prefecture and has been executing this policy until 2005 on 69 routes as the major subjects.

2. **Objective evaluation based on the research on traffic accident data and traffic volume**

To objectively grasp the effect of the introduction of this policy, we examined the policy in two routes installed in Toyota City in July 2002. Before and after data of traffic volumes, vehicle speeds and the number of traffic accidents are counted and questionnaire survey for local residents are made to present the features of this policy. In this paper, we describe the result of the research about traffic volumes, vehicle speeds and the number of traffic accidents in one road.

2.1. **Case Study Area - former Toyota-isshiki line (Takemura district, Toyota City)**

This route was a arterial road (managed by prefecture) and there are many buildings along the road including a supermarket. Furthermore, this road is usually used to go to the rail station to go to workplaces or schools having the significance as a main lifeline for the
community. The major function of traffic has already moved to the new bypass thus this road is to be made to a road that serves more role as lifeline of the community. In addition, except the morning hours and evening hours, most of the pedestrians are aged persons thus the acquisition of safe walking space is critical.

As the result of the arrangement, changes in traffic volume, vehicle speed and the number of traffic accidents as shown below were observed.

2.2. Changes in traffic volume, vehicle speed, number of the accidents

Traffic volume, vehicle speed, number of the accidents before and after the arrangement are examined. The results are Figure 3,4 and table 1,2.

Traffic volume and speed are measured by the traffic counters installed on going and coming ways at the point where the speeds of the cars are highest in the road. Figure 3 shows the comparison of car speeds between before and after the arrangement. The average speed difference was very small but the decrease of the vehicle speed that is one of the purposes of this policy was indicated in this route. Also, figure 4 shows the changes in vehicle speed between cases oncoming vehicles existing or not.

As for the traffic accidents, the average number of the accident (before : 5years, after : 1.5years) was available. In that data, the injury accidents happened average 2.8 times/year.
for 5 years before the arrangement greatly decreased to average 0 time/year for 1.5 years after the arrangement.

Figure 3 Before and after comparison of speed distribution

Figure 4 Speed comparison by the existence of oncoming vehicle after the arrangement

Table 1 The car traffic in Toyota-issyiki line before and after the arrangement

<table>
<thead>
<tr>
<th></th>
<th>Before the arrangement</th>
<th>Half year later</th>
<th>One year later</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Week day</td>
<td>Holiday</td>
<td>Week day</td>
</tr>
<tr>
<td>number of vehicles/day</td>
<td>7098</td>
<td>4393</td>
<td>7411</td>
</tr>
<tr>
<td>Day(7:00 - 19:00)</td>
<td>5480</td>
<td>3412</td>
<td>5977</td>
</tr>
<tr>
<td>Night(19:00 - 7:00)</td>
<td>1618</td>
<td>981</td>
<td>1434</td>
</tr>
<tr>
<td>Ratio of night(%)</td>
<td>22.8</td>
<td>22.3</td>
<td>19.3</td>
</tr>
<tr>
<td>Ratio of big cars entered(%)</td>
<td>25.1</td>
<td>22.5</td>
<td>16.4</td>
</tr>
</tbody>
</table>
Table 2 The difference of the number of accidents

<table>
<thead>
<tr>
<th></th>
<th>Crossing point case year</th>
<th>others case year</th>
<th>total case year</th>
</tr>
</thead>
<tbody>
<tr>
<td>before</td>
<td>0.6</td>
<td>2.2</td>
<td>2.8</td>
</tr>
<tr>
<td>after</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

3. Conclusion

This paper examined the effect of the removal of center line based on objective data. Enough decrease in running speed was not recognized and the decreasing effect of traffic volume could not be seen.

As for the number of traffic accidents, the very good effect was observed on the researched road.

Due to space limitation, we cannot describe about the drivers perception and the result of the questionnaire survey for residents in this paper. But many people feels different before and after, and many driver change the way to drive in this road.

References


CARROT AND STICK TACTIC: PARKING TAXATION TO SUBSIDIZE BUS

Vittorio ASTARITA\(^1\) and Marco SALATINO\(^2\)

**Abstract.** Parking taxation is a powerful demand management policy, allowing mobility managers to achieve at least two goals. First, an increase in monetary cost of private modes usually turns into more trips on public transport. Second, collected resources can subsidize free-transit programmes. This is also a suitable strategy for large trip-attractors such as university campuses, as shown in this work. The proposed mode choice model allow mobility managers to identify the optimal pricing strategy for both parking fees and bus fares, in order to encourage transit use and minimize generalized transport cost. Finally, an empirical application of the proposed method on an Italian campus is shown.

1. **Introduction**

Supply management of parking spaces has been long considered one of the most effective tools for transportation demand discipline. By applying strategies based on parking regulation, it is possible to shift users towards more sustainable transportation modes. This is possible because there is a relationship between the number and price level of parking spaces and modal choice. More specifically, a reduction in parking supply (and/or an increase in parking price) has the effect of a decrease in private car use.

Considering data from many world cities it is possible to demonstrate a negative correlation between urban parking space supply, relatively to the number of workplaces, and the use of transit mode [4]. The same relationship can be also found from data collected in different Canadian cities [7]. Not only reducing parking supply, but also the use of appropriate parking pricing, one of more used Transportation demand management (TDM) measures listed in [6], has a relevant impact on modal choice [3]. Parking pricing influences modal choice by acting both directly (increasing driving costs) and indirectly (decreasing car ownership rates).

Another effect can be considered in addition to the direct effects of private car space pricing, that is the indirect effect caused by the reduced motorization rate among

---

\(^1\) University of Calabria, via P.Bucci Cubo 46B Rende, Italy, vastarit@unina.it  
\(^2\) University of Calabria, via P.Bucci Cubo 46B Rende, Italy, m.salatino@unical.it
commuters that would be forced to pay parking space, in fact typically an increased average use of transit mode corresponds to users with a reduced motorization rate [5].

Considered that an excessive in reduction of parking supply could compromise accessibility to urban services, parking price seems to be the more applicable strategy, when some limits are respected: first, parking price level should not be set too high, to reduce the risk that users may shift towards alternative destinations [8]; second users do not appreciate road pricing intervention, unless the collected resources are directly reused in the transportation system to increase level of services.

One possible solution is to introduce parking pricing, finalised to finance programs for public transportation improvement. This parking pricing strategies applied to improve transit service have had a good application in the United States, where university campus have difficulties in finding resources to realise parking structures that are able to receive all the existent car demand. By using this strategy, the university can answer to student discontents that may follow parking supply space restrictions and the introduction of parking pricing, with the promotion of programs that sustain reduction of transit fares [9].

2. The proposed model

The proposed simple model is introduced to evaluate the relationships between policy parameters and equilibrium conditions at a macro level. A more detailed modelling with a network-based supply representation can be found among others in [2].

In the proposed model it is assumed that a certain destination of systematic travel trips can be reached both with bus and car and the travel time on these modes does not depend on traffic flows. In this model public transportation can be financed with resources obtained from parking price. The following notation can be used to express model’s variables:

\[ x_{bus} \] fraction of users travelling on car;
\[ x_{car} \] fraction of users travelling on bus;
\[ f_{bus} \] bus travel fare;
\[ f_{car} \] daily parking fee;
\[ f'_{bus} \] net bus travel fare paid by users obtained from the total travel fare reduced by the incentive obtained from parking pricing;
\[ K \] total number of parking spaces expressed as a fraction of total demand;
\[ c_{bus} \] generalised cost of public transportation;
\[ \Delta f \] reduction of bus travel fare obtained from parking pricing;
\[ c_{car} \] car generalised cost;
\[ t_{bus} \] bus travel time;
\[ t_{p} \] car time to perform a parking manoeuvre;
\[ t_{car} \] car total travel time;
\[ f_{0} \] out of pocket car cost;
\[ \beta \] value of unit of time.

Those variables are connected by the following relationships:

\[ c_{car} = f_{0} + f_{car} + \beta \cdot t_{car} \] (1)
\[ c_{bus} = f_{bus} - \Delta f + \beta \cdot t_{bus} \] (2)
\[
\Delta f = \frac{x_{car} f_{car}}{1 - x_{car}}
\]

(3)

\[
t_{car} = t_0 + t_p
\]

(4)

\[
f'_{bus} = f_{bus} - \Delta f.
\]

(5)

The relationship used in this model to estimate the time needed to perform a parking manoeuvre comprehensive of parking spot search it is simple and analogous to the one proposed in [1] (a more detailed formulation can be found in [2]):

\[
t_p = \frac{\alpha}{1 - \frac{x_{car}}{K}}
\]

(6)

where \( t \) is the average travel time to find a parking spot, \( x_{car} \) is the occupation level of parking supply spaces, \( K \) is the total number of parking spaces expressed as a fraction of total demand and \( \alpha \) is a structural parameter. Cost curves relative to the two considered modes as a function of changing modal split are showed in figure 1. User optimal deterministic equilibrium is obtained at the intersection of the two curves, and can be obtained by imposing equal generalised travel costs:

\[
x_{car} = K \left( 1 + \frac{\alpha \beta}{f_{car} + f_0 - f'_{bus} - \beta(t_{bus} - t_0)} \right)
\]

(7)

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{cost_functions.png}
\caption{Cost functions for ‘car’ and ‘bus’ modes.}
\end{figure}
Owing to the circular relationship $x_{car} \rightarrow \Delta f \rightarrow f'_{bus} \rightarrow x_{car}$ in formulas (3), (5) and (7), we can calculate modal split by simply implementing an iterative algorithm (Figure 2). A set of If...Then instructions assure bus fare values to vary within a desired range, i.e. $f'_{bus}$ neither can be negative nor exceed $f_{bus}$.

3. A case study

The proposed model has been tested in the University of Calabria, for which we can retain all the model’s assumptions. Unlike most Italian universities, which are situated within urban centres, this campus is some kilometres far from the nearest city centre. Since this university is similar to Anglo-Saxon colleges, a fair chunk of students and staff members live in the campus. The rest of university’s population, almost 25,000 people, commutes from the city centre and this causes a huge traffic flow on road network.

Since pedestrian accesses to the university are too long, non-resident students have only two way of approaching the campus, namely private car and public transport. In order to encourage students to use travel by bus, university managers guaranteed reduced fares by drawing up an agreement with the main transport operator.

During thirty years from the birth of the campus, parking lots have grown endlessly. Recently, an unforeseeable boom of matriculations together with constant rise of students’ car ownership rate has overcrowded parking capacity. The university has responded by increasing parking supply further, and this leads demand to grow more towards a higher equilibrium point. Because of this vicious circle, parking demand is not only raised by new matriculations, but also by modal shift from bus to car.

Assuming that parking fees are used to reduce bus fares, we can apply the proposed model to simulate modal split changes ($x_{car}$) depending on parking supply, namely parking capacity ($K$) and taxation ($f_{bus}$). Since no survey has been accomplished to calibrate parameters, we arbitrarily hypothesized the values shown in Table 1 according to common sense. The resulting modal split is shown in Figure 3.
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$f_{bus}$</td>
<td>1 Euro</td>
<td>$t_{bus}$</td>
<td>30 min</td>
</tr>
<tr>
<td>$t_0$</td>
<td>15 min</td>
<td>$\beta$</td>
<td>20 Euro/h</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>1</td>
<td>$f_0$</td>
<td>1 Euro</td>
</tr>
</tbody>
</table>

Table 1. Parameters of the proposed model.

Since bus fare reduction depends on modal split, as in (3), we can only determine the net bus fare for students as a consequence of resources from parking taxation. In this simplified model, mobility managers cannot choose a preferred bus fare, but this is obtained by acting on parking fees, which are totally redistributed on public transport. Resulting bus fares are shown in Figure 4, on the right. Note that these curves are quite the same shape of average generalized costs (Figure 4, on the left). Mobility managers may chose an appropriate combination of both $K$ and $f_{wp}$ so that students can travel free by bus. It is always preferable to chose the right halves of each “U” curve of Fig. 4, because they correspond to less trips by private mode, as shown by Fig. 3.

![Figure 3. Car modal split versus parking capacity and parking fee; darker shades mean higher car use.](image-url)
4. Conclusions

The model proposed in this work makes it possible to estimate modal split between car and bus modes when mobility managers use parking pricing to encourage public transport. A simple case is solved by using an iterative algorithm to bypass circular dependences among involved variables. Results from the model show that in some circumstances such policies can be useful in reducing average generalized transport costs, since they allow mobility managers to shift resources from private to public transport. Such a policy makes it possible to kill two birds with a stone. On the one hand, it is suitable to achieve a more sustainable modal split, by reducing private car use. On the other hand, this policy increase equity in transport, since it reduce transport costs for commuters without their own car.

References


AN INCIDENT DETECTION METHOD BASED ON THE ESTIMATION OF TRAFFIC FLOW PARAMETERS FROM INSTRUMENTED VEHICLES COUNTS

Vittorio ASTARITA¹ and Giuseppe GUIDO²

Abstract. In this paper an incident detection method based on the estimation of traffic flow parameters from instrumented vehicles counts is investigated. The procedure involves the estimation of traffic flow parameters when only a fraction of the total flow is composed of “instrumented” vehicles: "instrumented” vehicles are vehicles with any electronic or non electronic device, that can be counted at specific road sections by a centralized system.

1. Introduction

Many researches have been carried out on the localization of incident on traffic road networks and on real time estimation of non-recurrent congestion phenomena. Already from the 70s some researches have started on the possibility of traffic incident detection based on the use of loop detectors. Traffic road is divided into segment and the parameters of traffic are estimated from traffic counting at given sections. The algorithms used for traffic parameter estimation and for incident detection are based on different methodologies, among them: tree choice models and density values assessment [14] and [16], time series analysis [2] and [7], Kalman filtering [17], catastrophe theory [15]. New technologies and artificial intelligence have made possible to develop new algorithms based on: neural networks [5],[6] and [1], fuzzy logic [9] and [4], genetic algorithm [8] and [10].

In the recent years there has been a widespread use of radio communications tools in the majority of developed countries, and an increasing numbers of customers using mobile phones that allow the exchange of large quantities of data and continuously updated information [13]. There are many possibilities that stem from the use of these technologies; among them, fundamental for the development of improved transport system management and control, is the estimation of traffic flow parameters. The deployment of widespread information and communication networks may now facilitate the extraction of detailed, accurate information on traveler position from the localization of mobile phones and electronic payment tags used in automatic toll collection.

¹ University of Calabria, via P. Bucci Cubo 46/B, Rende, Italy, vastarit@unina.it
² University of Calabria, via P. Bucci Cubo 46/B, Rende, Italy, guido@unical.it
This new technologies can be considered part of the new Intelligent transportation systems (ITS) that include applications of new technologies to traffic management and control. Experience has shown that it is difficult to accurately estimate the true system traffic state and the estimation of traffic flow parameters remains a fundamental objective in the development of new ITS.

For this reasons, alternatively to the above mentioned methodologies, some recent researches have been presented based on enhanced telecommunication system applied between vehicles and road control infrastructure: incident detection system based on radiofrequency [12] and models based on mobile phone calls traffic in emergency conditions [11].

The possibility offered by the knowledge of instrumented vehicle positions on some sections of the road network allows traffic control management to estimate, in real time, traffic flow parameters.

In this paper the algorithm presented in [3] is applied to incident detection.

2. Traffic parameter estimation methodology

The use of new communication technologies, in Italy as in nearly all the industrialized countries, is clearly rising and directives have been supplied, also in Italy, in a recent “Piano Nazionale dei Trasporti” which proposes data transmission to be increasingly present in applications for the study and management of the transport system.

Vehicular traffic flow is a complicated random process, described with parameters that depend on space and time. Information on mobile phones or electronic tags is handled by different organizations, different situations can be found in different countries. In this paper it is supposed that “instrumented” vehicles can be counted at specific road sections by a centralised system. The estimation methodology presented in [3] consent the estimation of traffic parameters on a motorway network. Traffic parameters that can be estimated are: traffic flow, density and speed. A stretch of road between two sections, where “instrumented” vehicles can be counted, is defined as a “cell”. The motorway is subdivided into cells, assuming that “instrumented” vehicles entering and exiting every cell can be counted during the observation period. Moreover, the number of vehicles that enter the first cell of the network and the number that enter and exit on ramps can also be known depending on the situation. The “instrumented” vehicle concentration is obtained and propagated over the network in time and space. This allows one to estimate traffic flow parameters by sampling “instrumented” traffic flow parameters using a concentration propagation mechanism. Some calculation algorithms are introduced for the estimation of traffic flow parameters within the theoretical framework introduced. The presented algorithms are based on numerical methods for the solution of partial differential equations.

The estimation algorithms used in this paper are the result of interaction between the classical theory of traffic flow and the sampling of data that can be allowed by telecommunication technologies applied to the traffic systems. Results are obtained from the combination of traditional traffic counts with traffic counting of instrumented vehicles on specific sections of the network.

Results of the new methodology have shown, on test networks, minimal differences between estimated density values and observed density values. Accuracy of the estimate
improves with an increase in the ratio between instrumented vehicles (mobile phone on board) and total vehicles on the network.

3. Incident detection

The traffic parameter estimation model is applied to obtain densities and speeds for every cell of the motorway network. An incident can be detected by abrupt changes in those variables. Different incident detection algorithms can be proposed and each of them can have different threshold parameters that can be established. An evaluation in the effectiveness of an incident detection algorithm can be obtained by the percentage of incident that are detected on the total number of occurring incident and by the percentage of false alarms. Based on this performance measures some algorithms have been studied on a test network where a motorway and an alternative road are composed into a circular network that is similar to an urban ring toll road. On the test network traffic flow and incidents have been simulated varying conditions as traffic cell length, network extension.

The situations analyzed are among the possible scenarios obtained by any combination of the following:
- Rate of instrumented vehicles stable in time and among o/d pair, rate of instrumented vehicles stable in time but different among o/d pair or rate of instrumented vehicles different in time and among o/d pairs.
- Information also obtained from real time traffic counts at motorway entrances, exits and at some specific location; information obtained from real time traffic counts at motorway entrances and at some specific locations or information obtained from real time traffic counts only at specific locations.

The results obtained in this paper are not only a confrontation between different algorithms, but also an evaluation of time savings obtained by users that are rerouted on the alternative road. Traffic users are rerouted by a single organisation that knowing all “instrumented” vehicles traffic counts is capable to estimate traffic flow parameters and detect incident occurrence. Users are then re-routed by means of variable message signs informing of incident occurrence.

4. Conclusions

Models such as the one proposed in this article can be useful for the analysis and simulation of many different situations on our roads. Customers and managers of the transportation system would find, in the application of these new methodologies, a valid contribution to the resolution of problems relative to road traffic.

Further research efforts should be devoted to an application of the methodology to real traffic data. In this paper only the main aspects of the problem have been presented opening the road for additional on field research.

Many ITS have been applied and designed in recent years, but the great upcoming challenge is to combine on the field different sources of information, already owned by different entities, and to extract what may be relevant to various control systems, with the aim of supplying better, more efficient transportation services to road travellers.
References


TRAFFIC CONFLICT ANALYSIS BASED ON MOVING COODINATES SYSTEM WITH SYNCRONIZED MULTICAMERAS

Hiroshi WAKABAYASHI¹ and Shin-ya MURAMATSU²

Abstract. This paper proposes a moving coordinates system for analyzing traffic conflicts towards ITS-assist traffic safety. First, the significance of the moving coordinates system of analyzing traffic conflict is discussed against fixed coordinates system. This moving coordinates system consists of on-vehicle synchronized five or more cameras shooting ahead and behind of right / left hand side from the observation vehicle, video image capturing subsystem, image processing subsystem, and coordinate-system transforming from pixel coordinate-system to plane coordinate-system on actual roads. After reviewing the previous works on the conventional conflict indicators such as TTC by Hayward (1972) and PET by Allen et al. (1978), the new conflict indicator PTTC is proposed. Finally, TTC, PET and PTTC indicators’ calculation in the expressway merging traffic are demonstrated.

1. Significance of the moving coordinates system

One of the basic countermeasures for traffic safety is to start conflict analysis at frequently weaving occurrence section and merging / diverging sections. The safety measures based on the conflict analysis can be proceeded to ITS technologies that assists drivers’ judgment and operation. In addition, this technology is expected to develop to Advanced Cruise-Assist Highway System (AHS).

To prevent accident, detecting traffic conflict in advance and providing warning information to driver is very effective. To provide conflict warning accurately, it is

¹ Faculty of Urban Science, Meijo Univ., 4-3-3, Nijigaoka, Kani-city, Gifu, 509-0261, Japan. wakabaya@urban.meijo-u.ac.jp
² Ashley Associates, Odakyu-Minami-Aoyama bld. 8F, 7-8-1, Minami-Aoyama, Minato-ku, Tokyo, 170-0062, Japan. bxu03316@nifty.com
important to measure the conflict indices from the distance between the vehicles and the change in behavior of the involved vehicles from detected distance, speed and acceleration. This conflict analysis has been often implemented from the observation point on the roadside such as from pedestrian bridge so far. This observation is called a fixed coordinates system analysis in this paper.

In the case of the fixed coordinates system, distance between vehicles varies depending on the driver and thus there are cases that are dangerous in terms of conflict indices but not be felt dangerous by drivers. This suggests that there is an individual difference among drivers. In addition, the recognitions of the observer and the driver for dangers doesn't always agree. Thus instead of that the dangerous threshold of the conflict index is set by observer, it should be set up depending drivers’ characteristics for their recognitions. Thus it is necessary to observe conflict indices for a certain driver in usual operation to grasp the individual characteristic continuously. The ground coordinates system is, however, difficult for this observation. Therefore, in this study, the multi-cameras observation system is installed on the vehicle. This is called a moving coordinate system observation in this study. Using this system and the conflict indices proposed by Hayward (1972), Allen et al. (1978) and the Wakabayashi and Renge (2003), on-vehicle conflict analysis is carried out in this study.

The conflict analysis with moving coordinates system enables the following analyses: One is the new analysis for traffic accidents. The traffic accident analyses have been carried out using the plane road figure, that is, using the fixed coordinates system. However, it is desirable to analyze the accident based on the moving vehicle. Thus, the conflict analysis using moving coordinates system contributes for traffic accident analysis. The other is the analysis of the individual differences. There is a potential difference among individuals in the recognition for dangerous situation. This can develop to a learning function in the ITS safety assist systems.

2. Traffic conflict analysis based on moving coordinates system

Moving coordinates observation system consists of five cameras, synchronized unit and recording camcorders. Photo-1 shows on-vehicle four cameras those shoot right-forward, left-forward, right-backward and left-backward. The remainder camera shoots a speedometer for calculating conflict indices.

To obtain the actual plane coordinates on the road from the image, the vehicle tracking system with digital VCR

![Photo-1 On-vehicle observation system](image)
(Wakabayashi and Renge, 2003) is used. The vehicle tracking system records the position of the moving object by replaying the images that are captured in a PC.

Coordinates transformation from pixel coordinates system into actual plane coordinates system is then carried out. In this process, projective transformation is used. The transforming equations are as follows:

\[
\begin{align*}
    x &= (a_1u + a_2v + a_3)/(c_1u + c_2v + 1), \\
    y &= (b_1u + b_2v + b_3)/(c_1u + c_2v + 1),
\end{align*}
\]

where \(u\) and \(v\) are pixel coordinates in image, and \(x\) and \(y\) are actual plane coordinates of road, and \(a_1, a_2, a_3, b_1, b_2, b_3, c_1, c_2\) are parameters. These equations are derived from

\[
\begin{bmatrix}
    x \\
    y \\
    1
\end{bmatrix} =
\begin{bmatrix}
    a_1 & a_2 & a_3 \\
    b_1 & b_2 & b_3 \\
    c_1 & c_2 & 1
\end{bmatrix}
\begin{bmatrix}
    u \\
    v \\
    1
\end{bmatrix}.
\]

3. Indicators for traffic conflicts

Use of conflict indicators enables quantitative and objective judgment for what degree of the dangerous traffic condition. The following indicators are calculated.

3.1. TTC indicator

As mentioned in introduction, TTC (Time-To-Collision) indicator is one of the traffic conflict indicator proposed by Hayward (1972). It is defined that the time to collision if the two vehicles continue to drive at the same speed and the same angle without any evasion behavior. Maximum value is infinite and minimum value is 0 second (collision happens).

3.2. PET indicator

PET (Post encroachment time) is identified as the time from the end of encroachment to the time that the through vehicle actually arrives at the potential point of collision (Allen et al., 1978). In other words, the potential place of collision is firstly determined where the vehicle 1 once occupies at the certain time, and then PET is defined as the time that the vehicle 2 reaches the place.
3.3. PTTC indicator

In a fixed coordinate systems analysis, many seriously dangerous conflicts with very short distance in car following situations are found. Even such situations, however, there are many cases that no TTC indicators are calculated, which demonstrate the traffic condition safe. This is because that even in such dangerous car following situation, unless the preceding vehicle decelerates, two vehicles will not theoretically collide as is apparent from the definition of TTC. This suggests that more appropriate indicators are required for such dangerous car following situation.

Thus, PTTC (Potential Time To Collision) indicator has been proposed by Wakabayashi and Renge (2003). This indicator is an "If ...then" type indicator, that is, if the leading vehicle evades dangerous object and decelerates in the situation that two vehicles are traveling with very short following distance on the same lane, the following vehicle needs to take rapid action to avoid collision. PTTC value is a solution of the following equation.

\[ D = \Delta v \cdot \frac{R_{ATC}}{2} + aPTTC^2 / 2, \]  

where, \( D \) is the relative distance of two vehicles, \( \Delta v \) is the speed difference, and \( a \) is the deceleration. This indicator is suitable for high-speed and congested traffic situations.

4. Application to the actual traffic and calculation of conflict indicators

We carried out experimental observation using this moving coordinates conflict analysis system. The experimental route is the Nagoya Expressway Ring(Loop) Line in central Nagoya, Japan. This ring has many merging sections from on-ramp and other radial lines, and has a large traffic every day. At some sections, drivers have to weave from right lane to left lane, and vice versa, depending on their destinations. Examinee is encouraged to use his vehicle because he used to operate it. Photo-2 demonstrates an example of obtained movies those are every two seconds shots from five cameras. In this example, the observation vehicle tried to change lanes from right to left in a congested traffic condition (the speed is approximately 20 km/h).

Using the vehicle tracking system with digital VCR, obtaining the pixel coordinates from these images, then the actual coordinates on road are calculated. Distances between vehicles for longitudinal direction, cross-sectional direction and integrated direction, speed, conflict indicators including TTC, PET and PTTC are plotted in Figure-1.

In this example, no TTC index towards surrounding vehicle 1 is calculated. However, PET index is calculated at the time of approximately 6.5 seconds in Figure 1 among 6 seconds and 8 seconds in Photo-2 that a merging action is accomplished. PTTC index is earlier calculated than PET index among from 5.6 seconds to 5.8 seconds. Because of congestion, the distance between vehicles is approximately two meters at the time of the calculation starting. After that, the distance between vehicles increases.
Photo-2 Continuity images obtained with synchronized multi cameras:
vehicle in a white circle is the surrounding vehicle 1
5. Short concluding remarks and future study

The results of Section 4 suggest that the PCCT indicator is useful as the warning index for the safety. It is necessary for further study to accumulate data and to grasp the characteristic of the various conflict index.

References


TRANSFERABILITY OF DESTINATION CHOICE MODELS WITH
LAND USE VARIABLES: A CASE STUDY ON THE
METROPOLITAN AREA OF NAPLES

Luigi Biggiiero¹, Cinzia Cirillo² and Gaetano Galante¹

1. Introduction

The research presented in this paper is originated by recent work done under the sponsorship of Regional Transportation Research Center based in Campania (TEST - WP6 unit). The general idea is to constitute a knowledge center that can offer services to the administration and or to private bodies active in the domain of land use planning and traffic management. In particular, the sub-group WP6 is in charge of the development of the relation between transport system and land use. The majority of the members are academics, but they are supposed to deliver products and results ready for use and useful to the development of the Region.

As often said in this area of research, the communication between urban planners and transport modelers is very difficult and the attempt to integrate objectives, data and methods remains in practice a dream.

After a number of meetings and discussions a number of common issues were identified:

− All human activities are located in space; residences, working locations, shopping and leisure activities and their spatial distribution constitutes the choice set from which individuals need to travel. In our region the land use is often unplanned, but increasingly local authorities are trying to regulate or at least control the growth of their territory. Some rational plans need to be adopted especially when funding for activities’ development becomes available from the State or from the EU. In that context local and regional authorities can become a useful source of data when an integrated model is developed.
− The implementation of an integrated land use-transport model is very complex and difficult to make in practice; on the one hand the research has not delivered a common

¹ University of Naples - DIT, via Claudio 21, Napoli Italy
² University of Namur - GRT, rempart de la Vierge, 8 Namur Belgium
framework and on the other hand the number of operational systems being applied by planners is not large.

- Integrated models demand an enormous amount of data, their collection and maintenance requires long-term investment and skilled people; those resources are often not available for a single project.
- Land use are often just exogenous variables for transport modelers; however changes in land use can cause change in land use patterns and hence changing activity-travel patterns.
- The choice of destination is often a key factor for the feasibility of particular function at particular locations.

2. The transport model for the metropolitan area of Naples

In that context it was clear that no funding was available for data collection or to start the estimation of a new model system, but that some resources could be allocated to the land use database and that existing surveys and models could be used by the research group. In particular the existing transport model for the metropolitan area of Naples was made available. It was developed to forecast demand on the regional rail system, which is particularly dense in the area around Naples. It is a classical four-stage model recently updated with new surveys and counting; its validation has also given quite satisfactory results. The study area is about 1.171 km² with 3.016.584 inhabitants; the average density is of 2.576 inhabitants per km². The area is divided into 57 zones. The transport network modeled includes both private and public (bus and train) supply.

The model systems consists of the following steps:
- traffic generation;
- destination choice;
- modal choice.

O/D matrices adjustments using traffic counts for each
- time of day periods (two hours);
- mode (private and public) and
- adjustment of choice model parameters from the corrected O/D matrices.

In particular the destination choice model is a multinomial logit model; four models have been estimated for the following purposes: work, primary and secondary school, university, other purposes. The variables included are: the number of workers for each destination zone, the number of students primary and high school, the number of students at university, the number of people employed in commerce and services; all such size variables are in logarithmic form. In addition the distance and a number of dummy variables (i.e. intra-zonal dummy, Naples as final destination) have been found significant.
3. The destination choice model

The destination choice model to be transferred, was originally estimated for the Flemish region (Belgium); the basic geographical unit considered in that project were the statistical sectors. Two novelties characterize this model: the choice set generation and the type of variables used for the estimation. A restricted area where the individuals were supposed to perform all the trips contained in his/her daily activity chain was constructed using the action space concept (Dijst and Vidakovic, 1997). The potential action space being defined as the area containing all activity places which are reachable subject to a set of “temporal and spatial conditions”.

In the original model four sets of explanatory variable were estimated: (1) land use variables, (2) size variables, (3) impedance and, (4) socio-demographics characteristics.

- **Land use variables:**
  - Statistical sector area (total geographical area of the sector [m²])
  - Densely-built housing
  - Built-up housing
  - Housing and other developments
  - Industrial / commercial / port area
  - Infrastructure (highways, district roads, airport and/or railway infrastructure and so on)
  - Green / nature area (broad-leaved, coniferous and/or mixed forests, municipal parks, heath land and moors, dunes and beaches, water)
  - agriculture and meadowland (agriculture and open space, meadowland and orchards)

- **Size variables:**
  - shopping
  - financial (banks)
  - hotel / restaurant / café
  - cinemas
  - sport activities
  - cultural, recreational and leisure activities (museum, library, school of music, zoo, nature reserve, theatres, casino and so on)
  - car retail
  - personal service (beauty center and so on)

- **Level of service variables**
  - time
  - cost

- **Socio-demographic characteristics**
The socio-demographics available for this study were: age, sex, number of children (0, 1 or more), number of cars (0, 1, 2 or more), number of workers (0, 1, 2 or more) and marital status. Interactions of the accessibility with the socio-demographics were explored to measure deterministic taste variation among individuals.
4. The transferability study

This study aims at examining the problem of model transferability. An existing four-stage model for the metropolitan area of Naples will be updated with a destination choice model estimated for the Flemish region in Belgium. The models was estimated on a single day travel diary accounting for all tours made by a single individual, the destination choice set resulting from an action space. The variables estimated include a number of land use variables and a quite rich number of size variables. The estimation of a similar model in the context of the metropolitan area of Naples was impossible for the lack of a detailed travel diary. The study intended to include land use variables by transferring coefficients estimated in another context and test the goodness of the results in terms of O/D matrices estimation.

In particular, a number of issues will be treated in that paper:

− Are the results of this transferability reliable?
− Is the inclusion of land use variables improving the model prediction?
− Which are the policies that can be investigated with the inclusion of land use variables into a four-stage model?

Very few examples of model transferability study exist in the literature, and not many of them deal with real applications. We believe that this is a very good opportunity to improve an existing model with land use variables and compare performances with real counting. This practice will help the dialog between planners and transport modelers and will perhaps enhance the progress towards integrated land use transport models in our Region.

References


A STOCHASTIC APPROACH TO DELAYS AT INTERSECTIONS

Henk J. VAN ZUYLEN¹

Abstract. Travel times and the variation in travel times are important characteristics of the conditions in a road network. In urban networks travel time variability is mainly due to delays at controlled intersection. This paper describes a new model to calculate the probability distribution of delays at a controlled intersection. This is a basis for the estimation of travel times in urban networks.

1. Introduction

Dynamic traffic management becomes more and more a sophisticated fabric of direct control, in-car functions, route guidance and travel information provision. Important information for many management measures is the travel time. Reliability of travel time is becoming an important issue, comparable to the travel time itself. Drivers appear to give a high value to reliability of travel times and appear to prefer longer but reliable routes to shorter routes with irregular travel times (Bogers et al. 2004).

There is still a great deal of unpredictability in travel times due to stochastic processes. Even though we can predict expected travel times rather well (Liu et al. 2006), there is still a lot of uncertainty due to stochasticity. There is not much work reported in literature on the estimation of travel times based on intersection delays. Lin (2004) found that intersection delay in the context of arterial travel time prediction is different from the average delays calculated by existing delay models. Olszewski (1994) provided a systematic computation of vehicle delay for one cycle based on the geometrical relationship. His approach can be used to convert queue lengths to delays.

In this paper one stochastic element of travel time is discussed: the delay at controlled intersections. Travelers have to wait for the green light and for the queue in front to pass the intersection before they can continue their trip. Even if we assume that the traffic control is fully predictable there are still two stochastic elements: the arrival moment at the back of the queue and the length of the queue. This paper describes a model to capture these two stochastic elements.

¹ Delft University of Technology, Department Civil Engineering and Geosciences. h.j.vanzuylen@tudelft.nl. This research has been sponsored by the Transport Research Center Delft.
stochastic aspects and it derives a model to determine the probability distribution of the delay.

Section 2 describes the model, section 3 derives mathematical expressions for the statistical distribution of delays and gives some results of numerical calculations and section 4 discusses the results.

2. The delay model

Urban trips by car are for a great deal determined by delays at controlled and uncontrolled intersections. This is even more true for the variability of travel times. The delay depends on moment of arrival at the back of a queue: the queue that is present at that moment and the time that the green phase starts.

For the calculation of the average delay several models have been developed in the past. The calculation of delays has been recently improved by the modeling of queues as stochastic quantities with a probability distribution (Viti 2006). The probability distribution can be modeled as a Markov chain process.

Less result have been obtained for the calculation of the variability of travel times and delays. However the probability distribution of queues where the Markov chain process is used for the calculation of the distribution, gives a good framework for a model to calculate the variability of the delay.

We assume that the red phase starts at \( t = 0 \) and that the probability of \( n \) arrivals between time 0 and \( t \) is \( A_n(t) \), the probability that at time \( t \) the queue length is \( n \) is \( P^q_n(t) \). For \( P^q_n(t) \) a Markov model has been developed by van Zuylen and Viti (2006): During the red phase the following model is applicable

\[
P^q_n(t) = \sum_{j=0}^{n} P^q_j(0) A_{n-j}(t) \quad \text{(1)}
\]

for the green phase the Markov model is

\[
P^q_n(t) = \sum_{l=0}^{n+s(t-t_r)} P^q_{n-l+s(t-t_r)}(0) A_l(t) \quad \text{for } n > 0, \text{ and} \quad \text{(2)}
\]

\[
P^q_0(t) = \sum_{l=0}^{s} A_l(\Delta t) \sum_{j=0}^{n} P^q_j(t-\Delta t) \quad \text{(3)}
\]

where \( s \) is the saturation flow (for simplicity assumed to be deterministic and constant)

Given a certain queue length \( n \) at arrival, the time until the vehicle can pass the stop line is the time until the green phase plus the time for the vehicles standing in front has passed the stop line. If the queue in front is longer than the number that can pass in the green phase, the vehicle has to wait for a next green phase:

\[
W(t \mid n) = (t_r - t) + \left[(n+1)/s \cdot t_g\right] C + \left(n - \left[(n+1)/s \cdot t_g\right]\right)/s \quad \text{if } t < t_r \quad \text{(4)}
\]

and
\[ W(t \mid n) = \left( \frac{n+1}{s t_g} \right) C + \left( n - \left\lfloor \frac{n+1}{s t_g} \right\rfloor s t_g \right) / s \quad \text{if } t \geq t_r. \]  

(5)

\( C \) is the (fixed) cycle time, \( t_r \) the green time \((t_r+t_g=C)\). The square brackets \([ \ ]\) mean the integer value of the expression inside the brackets. The expression \( \left\lfloor \frac{n+1}{s t_g} \right\rfloor \) is the number of full cycles the arriving vehicle has to wait until the waiting vehicles in front and the vehicle itself can depart. It is assumed that at the moment that no vehicles are waiting any more in front and the signal is green, the delay has ended.

The expectation value \( E(W(t) \mid t) \) of the delay \( W \) for a vehicle arriving at time \( t \) is given by

\[ E\{W(t) \mid t\} = \sum_{n=0}^{\infty} P^q_n(t) W(t \mid n) \]  

(6)

where \( P^q_n(t) \) is the probability that there is a queue of \( n \) vehicles waiting at time \( t \).

3. The probability distribution for the delay

If we want to calculate the probability distribution function for delays \( P^W(W(t)) \), i.e. the probability for a certain delay \( W \) for a vehicle arriving at time \( t \), the mathematics becomes more complicated. The delay functions (5) and (6) have to be converted to a function of \( t \) only. As simplified first step we can look at the following case of an intersection that is not oversaturated and with a neglect of random effects. Let’s assume that the initial queue is zero and the queue builds up proportional to the time in the red phase and decreases proportional to the time in the green phase. In that case the delay is given by the function

\[ W(t \mid n = q, t) = \left( t_r - t \right) + q t / s = t_r - t \left( 1 - q / s \right) \quad \text{if } t < t_r \]  

(7)

\[ W(t \mid n = q, t_r - (t - t_r)(s - q)) = \left( q t \right) \left( t_r - t \right) (s - q) / s = t_r - t \left( 1 - q / s \right) \quad \text{if } t_r < t < t_r/(1-q/s) \]  

(7a)

\[ W(t \mid n = 0) = 0 \text{ for } t < t_r/(1-q/s) \]  

(7b)

In this simplified case the arrival probability during the cycle is assumed to be uniform. The probability that a vehicle has a delay between \( d \) and \( d + \Delta \) \((0 < d < d + \Delta)\) is given by the chance that vehicles arrive in the interval \( t = W^{-1}(d) \) and \( t + \Delta = W^{-1}(d + \Delta) \), where we use the inverse function of the delay function \( W(t) \). It is obvious that \( \Delta t = -\Delta d W(t) / d t \).

Given that the arrivals are uniform over the cycle, the number of vehicles that have a delay between \( d \) and \( d + \Delta \) is equal to the number of vehicles arriving between \( t - \Delta d t / d W \) and \( t \), i.e. \(-q \cdot d(W(t)) / dW \). The normalized probability density for vehicles \( D \) is given by \( D(W) = -\{ dt(W)/dW \} / C \). Both expressions are valid for \( 0 < W < t_r \). The inverse mapping of the delay \( W \) to the arrival time is not a single valued function as can be seen from figure 3. The derivative does not exist at \( W = 0 \). This is a complication that can rather simply be solved by introducing the Dirac delta function \( \delta(x) \) with the following properties:
\[ \delta(x) = 0 \text{ if } x \neq 0, \]
\[ \delta(0) = \infty, \text{ and} \]
\[ \int_{-\infty}^{\infty} \delta(x-x_0) f(x) \, dx = f(x_0) \quad (8) \]

Which makes the probability density function
\[ P(W) = \alpha \delta(W) + \beta \text{ for } 0 \leq W \leq t_r \quad (9) \]

were \( \alpha = 1 - t_r / \{C(1 - q / s)\} \) and \( \beta = \{C(1 - q / s)\}^{-1}. \)

Figure 1 The inversion of the delay function

As a next step let’s assume that an initial queue of \( n_0 \) vehicles exists at the start of the red phase and that the green phase is still long enough to handle all the traffic.

\[ W(t \mid n = n_0 + qt) = t_r + n_0 / s - t(1 - q / s) \text{ if } t < t_r \quad (10) \]
and
\[ W(t \mid n = n_0 + qt = s(t - t_r)) = \{t_r + n_0 / s\} - t(1 - q / s) \text{ if } t \geq t_r. \quad (11) \]

The problem becomes slightly more complicated when the initial queue becomes so large that the green phase becomes oversaturated. The question whether an arriving vehicle has to wait for a next cycle to depart, depends on the cumulative arrivals in the cycle plus the initial queue. As soon as this quantity exceeds the number of vehicles that can depart in the green time, the vehicle has to wait for a following cycle. In a similar way one can deduce when an arriving vehicle has to wait for two or more cycles. The delay becomes now according to eqs. (4) and (5)

\[ W(t \mid n = n_0 + qt) = \{t_r + n_0 / s + [(n_0 + qt) / s t_g] t_r\} - t(1 - q / s) \quad (12) \]

The probability distribution of delays for this case consists of two box shaped functions that may overlap (fig. 4, right).
The height of the box function is given by $D(d) = \frac{dt(w)/dw}{C} = \frac{1}{C(1- q/s)}$ and the begin and endpoints can be easily calculated. Using this function the probability distribution is expressed by

$$P(W) = \frac{1}{C(1- q/s)} \{B(W, C + (n_0 - t_g)/q, l_r + n_0 /s) + B(W, 2 l_r + n_0 /s - C (1 - q/s), C + (n_0 - t_g)/q + l_r)\} \quad (13)$$

Where $B(W, W_1, W_2) = 1$ for $W_1 < W < W_2$ and 0 if $W$ has other values.

3.1. An example with a stochastic initial queue.

As an example we assume an approach to a controlled intersection with a single lane, $s = 1800$ veh/h. The inflow is 600 veh/h, the cycle time is 60 s and the green time is 24 s. Each cycle 10 vehicles arrive and 12 can depart. The degree of saturation is 0. Due to the stochastic character of the arrivals some cycles will be oversaturated and the queue at the start of the red phase will not always be empty. Figure 5 (left) gives the probability distribution of the queue length at the start of red.

The graphical presentation of the calculated probability distribution for the delays is given below (figure 5 right):
4. Discussion

The extension of the conventional delay calculation with the calculation of probability distributions of delays gives the possibility to estimate the reliability of travel times on urban routes with controlled intersections. Using the Markov chain model for the probability of queues gives a rather simple computation scheme that results in numerical results for the probability distribution of delays.

The distribution function (9) and (13) as a sum of box shaped functions and a delta function is easy to calculate. For oversaturated conditions extensions to eq. (13) can be extended taking into account also vehicles that have to wait for more than two cycles.

References


1. Motivation

The paper we propose for presentation is part of a wider research aimed at modelling the impact of a traffic management operation or policy on the congestion in a network. For this we need to know how traffic behaves locally (for example because of a capacity reduction) as well as globally in a network (because of re-routing of informed drivers for example). This is the general framework of the dynamic traffic assignment (DTA) problem.

It is classically solved by searching for an equilibrium resulting from the route choices of all the drivers. This equilibrium search is often based on some trial-an-error like heuristic method. In some cases such methods may happen to be optimized (see for example the solution of Beckmann et al [2] using the Frank and Wolfe optimization algorithm). But in general very few is known about the properties of this equilibrium and how fast it can be reached. From a practical point of view, the problem is then to find the good balance between the precision of the description of traffic (in order to obtain a thin approximation of travel times) and the ease of calculation (in order to maintain admissible computation times).

The present paper is based on two modeling choices: First, the traffic is represented by the LWR model (Lighthill, Whitham [11] and Richards [13]) ; and second, the assignment procedure is progressive, that is to say that precision of the traffic description is all the thinner when the equilibrium is getting closer.

This progressive approach is made possible thanks to the original resolution method we use. Indeed, the classical resolution method of the LWR model is based on the finite difference Godunov scheme (see Daganzo [3] and Lebacque [8]). Such a scheme is very efficient as far as density and flow are concerned, but precision in travel time can only be obtained with a quite thin description of the traffic. On the contrary, the method we propose is based on the explicit tracking of waves in the network. It shows very good properties in travel time approximation, even with rough discretization.

The main contribution of the paper will actually be the presentation of this resolution scheme. The propagation of waves on (isolated) roads has already been presented earlier [6].
and will be briefly reminded here in order to introduce the principles of the Wave tracking method. After this, different issues will be addressed in order to enable the description of traffic in a network for DTA. Those issues include propagation of waves through intersections as well as propagation in the network of other types of information than density (origin-destination, travel time, etc.).

Finally the progressive assignment procedure will be examined and the wave tracking scheme will be proved to be efficient in such a framework.

2. The Wave Tracking method and its application to traffic

The LWR model considers traffic as a homogeneous and continuous stream, characterized by its flow $Q(x,t)$, and density $K(x,t)$. It assumes that flow is always at equilibrium and is given by the fundamental relationship $Q(x,t) = Q_E(K(x,t))$. Thus the classical conservation equation becomes the following scalar hyperbolic equation in density:

$$\frac{\partial K(x,t)}{\partial t} + \frac{\partial Q_E(K(x,t))}{\partial x} = 0$$

(1)

The Wave Tracking method (WT) is a numerical scheme for solving such a hyperbolic conservation law. It is based on the approximation of the fundamental relation $Q_E(K)$ by a piece-wise linear function. As a direct consequence, the solution of the model is unique and does not contain any (continuous) rarefaction fan, but only linear shock waves. Then the method reduces to tracking those waves till they collide with their neighbors, generating new shock waves which will collide later etc.

Even if such a scheme is less usual than the classical finite difference schemes, it seems to receive a renewal of interest in the mathematical society (see for example [7]). It has already been applied to the LWR model for a single road [6] with some boundary conditions [5, 4].

In order to apply this scheme to the resolution of the LWR model in a road network when modelling traffic assignment, we must be able:

1. To propagate traffic through the network (composed of links and intersections),
2. To use the model as a travel time provider for the traffic assignment procedure.

3. Traffic propagation

The propagation of traffic through a road network has been widely studied by different authors in the LWR context. Two main issues have to be solved:

- How to take into account some individual information such as the origin and destination of the trip, the type of traffic information the driver has access to, the level of knowledge of the network...?
- How does traffic behave at an intersection?
The first issue can be solved by splitting the flow into partial flows regarding some ve-
hicular classes (this feature is often referred to as multicommodity). Each class may have
a specific behaviour with respect to route choice. Thus traffic is characterized by its global
density $K$ and flow $Q$, and by its partial densities $K^d$ and flows $Q^d$ for each class $d$.

The equation of conservation (1) then becomes a systems of conservation equations:

$$\frac{\partial K^d}{\partial t} + \frac{\partial Q^d}{\partial x} = 0 \quad \forall d$$

which can be rewritten in terms of global density and composition of traffic:

$$\begin{cases}
\frac{\partial K}{\partial t} + \frac{\partial Q}{\partial x} = 0 \\
\frac{\partial \chi^d}{\partial t} + V \frac{\partial \chi^d}{\partial x} = 0 \quad \forall d
\end{cases}$$

Where $\chi^d$ is the proportion of traffic belonging to class $d$.

Such a system has been studied by [14, 9] in the framework of the theory of hyperbolic
systems of conservation laws. We will explain in this paper how the WT scheme can be easily
extended in order to cope with such a traffic composition. For this a new type of wave will
be introduced which is a discontinuity of composition (but not of density). Its behaviour and
interaction with other waves will be studied extensively.

Furthermore, as for the second issue, we will explain how those different waves behave
when they reach an intersection. This behaviour is based on a continuous model of the inter-
section in terms of flow. It will be shown in particular that a wave arriving at an intersection
(coming from downstream of upstream) may generate (on each adjacent link) one or several
waves moving away from the intersection.

We will further demonstrate that the method is effective (the solution can be calculated
in a finite number of operations) and efficient (the number of waves to be considered is not
tremendous) and finally the results are satisfying (they can be as close as desired to the exact
solution).

4. Travel time computation

As far as travel time computation is concerned, two types of methods could be used, consid-
ering one of the following (trivial) properties:

- The travel time of a vehicle is the difference between its exit time and its entry time of
  a link. Thus travel time can be computed by tracking the trajectory of a vehicle.

- If the FIFO rule is verified, then the variation of travel time on a link depends only on
  the flow at the entry ($Q_{in}$) and the at exit ($Q_{out}$) of the link [1]:

$$\frac{d\tau(t)}{dt} = \frac{Q_{in}(t)}{Q_{out}(t + \tau(t))} - 1$$

553
The integration of this equation in order to compute \( \tau(t) \) directly leads to the classical cumulative flow method [12].

Since density is piecewise constant (by construction of the method), travel time is piecewise linear and then we only need to know \( \tau(t) \) for a discrete set of time. In fact we need to know the travel time for entry times \( t \) corresponding to the discontinuities of \( Q_{in} \) and for exit times \( t' \) corresponding to the discontinuities of \( Q_{out} \).

A hybrid method will be proposed based on the tracking of vehicles entering the link at some variation of the entry flow and on the integration of flows till a variation in the exit flow. It will be proved efficient since it obviates some of the deficiencies of the initial methods (avoiding on the one hand the description of many trajectories and, on the other hand, the sum of numerical errors on a too large period).

5. Application of the wave tracking method into a progressive DTA procedure

Every dynamic traffic assignment can be expressed by a set \( \gamma \) of assignment coefficients \( \gamma_i^d(t) \) defined as the proportion of traffic of a given class \( d \) arriving at the exit of road \( i \) at time \( t \) and willing to use road \( j \) to reach its destination. In general, those sets of coefficients are continuous functions of \( t \), but since in the wave tracking framework, traffic flows are piecewise constant, then the assignment coefficients are piecewise constant as well.

The progressive DTA procedure we propose approximates the set \( \gamma \) by a (converging) succession of sets \( \{\gamma(k)\}_{k=1,2,...} \) obtained at each iteration step \( k \). We use for this two subroutines:

- **WT(\( \varepsilon \))** the wave tracking scheme with a precision \( \varepsilon \). It calculates travel times based on traffic volumes (that is based on traffic assignment coefficients)
- **RC** the route choice process which calculates the assignment coefficients based on travel times. This process may be a logit like in order to reach a stochastic (dynamic) user equilibrium or deterministic for (Wardrop like) dynamic user equilibrium.

The outline of the procedure is the following:

**Initialization:**

The traffic assignment coefficient are calculated based on the free flow travel times:

\[
\gamma^{(0)} \xrightarrow{RC} \gamma^{(0)}
\]

**Loop:**

At each iteration step \( k \) travel times are calculated:

\[
\gamma^{(k-1)} \xrightarrow{WT(\varepsilon(k))} \tau^{(k)}
\]

The descent direction is calculated based on those travel times:

\[
\gamma^{(k)} \xrightarrow{RC} \gamma^{(k)}
\]

The new assignment is a combination between previous assignment and descent direction
\[
\gamma^{(k)} = (1 - \lambda^{(k)}) \cdot \gamma^{(k-1)} + \lambda^{(k)} \cdot \gamma^{(k)}
\]

**While** there is a significant change in \( \gamma \)

Two sets of parameters make this algorithm progressive: the descent steps \( (\lambda^{(1)}, \lambda^{(2)}, ...) \) as well as the precision of the wave tracking scheme \( (\varepsilon^{(1)}, \varepsilon^{(2)}, ...) \).
Several combinations of them will be empirically tested on some theoretical scenarios. The convergence speed will be investigated in order to show the interest of using this progressive precision of the traffic description.

References

A COMPARISON OF MACROSCOPIC AND MICROSCOPIC APPROACHES FOR SIMULATING CONTAINER TERMINAL OPERATIONS

Giulio Erberto CANTARELLA¹, Armando CARTENI¹, Stefano DE LUCA¹

Abstract. This paper presents results of the application of macroscopic and microscopic models for performance analysis of a container terminal, to Salerno Container Terminal (Italy, EU). The results were compared and some guidelines have been proposed.

1. Introduction

Today the efficiency of container terminals plays a relevant role in freight transportation. The transport cost greatly affects commodity prices for some economic sectors; thus the main goal for a terminal operator is the minimization of the terminal through-out time, referring to several goals: unproductive moves number; container dwell-time; average waiting time of shuttle; average waiting time of external trucks; congestion on the roads; quay cranes waiting time; etc. Due to the high number of goals, and their relationships, simulation tools are needed to support decision making, with respect to design issues: strategic/long term decisions (investments; for example types and number of cranes); tactical/medium-short term decisions (for example storage yard lay-out, gate in/out, quay-side); shift-to-shift operations management, or operative planning, (storage allocation, unload/load operations, berth allocation, crane scheduling, yard storage etc…) as well as within-shift one (unload/load operations, yard storage etc…). In this paper we propose some results of the application of macroscopic and microscopic models for performance analysis of a container terminal, to Salerno Container Terminal (Italy, EU).

¹ Dept. of Civil Engineering – University of Salerno, Via Ponte don Melillo, Fisciano (SA)
g.cantarella@unisa.it; acarteni@unisa.it; sdeluca@unisa.it


2. Methodology and results

Design and project appraisal of container terminals can be carried out through macroscopic models, such as continuous flow networks, as well as microscopic models such as discrete event models.

Main advantage of the microscopic approach is that it allows for a detailed analysis in which each single activity is explicitly simulated. Of course this approach may lead to computational problems and is rather computer demanding, especially when resulting models are to be used to support optimization. It seems better suited for operations management. On the other hand, decisions about the container terminal configuration regard types and number of handling means, berth and yard layout, etc. and are better considered within strategic/tactical planning. In this case a macroscopic approach, based on container flows rather than single container movement, could be more effective.

The existent literature faces the problem either managing a container terminal as a system and trying to simulate all elements or managing a sub-set of activities. The main contributions try to maximize the whole terminal efficiency (examples in [1], [2], [3] and [4]), or the efficiency of a specific sub-area (or activity) inside the terminal (container storage and retrieval in the yard operations, space requirement problems, space requirement and crane capacity, re-marshalling strategy, and storage space allocation, stowage of vessels or berth allocation). For a deep literature review the reader may refer to [5].

The most followed approaches are based on deterministic optimization methods. Recently a stochastic optimization model has been proposed for maximize terminal efficiency (see [4]). The proposed model estimates the total workload for a time period and minimizes, for instance, the average time to unload and load a docked vessel. Usually each activity is analyzed using queuing models. An effective and stimulating approach for container terminal system analysis may be represented by simulation. In paper [2] a simulation model of the Pusan east container terminal is developed using an object-oriented approach and estimates container terminal performances, in paper [3] the Witness software is employed to analyze the performance of Hong Kong’s Kwai Chung container.

This paper is part of a more general research project aimed at container terminal analysis, say simulation and optimization. The research project methodological framework includes: model approaches identification (models architecture identification, models specification, decision variables definition and performance indicators specification); applications, object of this paper (test-site choice, survey for the calibration of the model parameters, actual scenario simulations and hypothetic scenario simulations) and conclusions (comparisons and guidelines for applications).

Both macroscopic models, based on continuous flow networks, and microscopic models, based on discrete event models, have been developed and are currently undergoing an enhancements process. A macroscopic model, developed by the same authors, has been presented at the Second workshop on the schedule-based approach in dynamic transit modeling [7]; a microscopic model based on discrete event simulation, developed by the same authors, has been presented at the European Transport Conference [6].

This paper presents results of the application of the already specified models for supporting performance analysis of a container terminal through aggregate indicators. These models have been applied to Salerno Container Terminal (SCT) in Salerno harbour (Italy, EU): estimate the performances of SCT, find out its main inadequacies and critical
points and simulate design scenarios in order to improve the system efficiency. SCT is, at the same time, a small and very efficient one; in fact it operates over less than 10 ha, close to 0.5 millions of TEUs per year, say 50,000 TEUs/ha. In addition the location of Salerno harbour does not allow to further enlarge the area, hence SCT should greatly rely on intensive approach to operation rather than an extensive one in order to keep pace with increasing demand.

An extensive survey has been carried (as reported in [6]), two types of data source have been considered: container monitoring data (available for the terminal info system) and “ad hoc” survey. The first data source refers to all the information that the terminal monitoring office measured every day: from January 2003 to July 2005 more than 1,000 vessels have been monitored. These data has been used for the data analysis of the vessel, gate and yard macro-areas, and particularly for the estimation of the berth-side/land-side demand (per type of container and time period) which involve the container terminal. Jointly with these data, an integrative survey was carried out during the first six months of the 2005. In particular all the berth macro-area activities referred to more than 5,000 TEUs were monitored.

The results of the applications of the macroscopic and microscopic models were compared and some guidelines have been proposed like strengths, weaknesses, fields of application for each approach, most effective indicators etc.

References


OBJECT-ORIENTED MODELING AND SIMULATION OF THE TRANSPORTATION SYSTEM

Hanna SAWICKA, Jacek ZAK

Abstract. The construction of the object-oriented simulation model of the freight, road transportation system is presented in the paper. The system operates at the Polish market of electrical supplies and provides goods delivery between 24 distribution centers spread out uniformly all over the country. An object-oriented computer simulation package Extend OR v. 6.0 has been applied to simulate operations of the transportation system. Its different characteristics have been determined, such as: transportation costs, fleet utilization, vehicles’ queue sizes (vehicles waiting for unloading), crew size in the system, efficiency of the material flow handling, etc. Simulation experiments resulted in drawing certain conclusions regarding the quality and accuracy of the proposed model. The sensitivity analysis has been carried out. The final results were compared with the real system. The authors presented advantages and disadvantages of their approach.

1. Introduction

Transportation, defined in a very synthetic way, is an activity which aims at changing the location of people and products. All those elements which interact between each other to achieve the goal i.e. transportation are described as a transportation system. The authors of this paper consider the transportation system as a set of such components as: fleet, infrastructure (roads, warehouses), human resources, safety and organizational rules which are responsible for planning, implementing and controlling the physical flows of products from points of origin to points of destination [5]. The complexity of a system led authors to the choice of the technique, which could deal with its analysis, modeling and carrying out computational experiments. This is an object-oriented simulation. Its paradigm evolved from the idea of developing a computer representation that is close to the way people think [1]. Thanks to this approach it is possible to model the system in terms of objects that have

1 Poznan University of Technology, Faculty of Machines and Transportation, 3 Piotrowo street, 60-965 Poznan, POLAND,
hanna.sawicka@put.poznan.pl; jacekzak@put.poznan.pl
priorities and behavior, and events that trigger operations that change the state of the objects, which interact between each other [7].

The objective of this research is to construct a simulation model of the existing transportation system operating at the Polish electrotechnical market. The model is constructed on the basis of the system’s analysis. An object-oriented computer simulation package Extend OR v. 6.0 is applied to simulate operations of the transportation system. A thorough comparative analysis of the proposed simulation model and the existing transportation system is carried out.

2. **Analysis of the transportation system**

The analyzed transportation system is a major component of the three-echelon distribution system which operates at the Polish market of electrical supplies and wholesales a full range of electrotechnical products. It is composed of the external (long-distance) subsystem and the internal (short-distance) subsystem. Long-distance transportation process is carried out by road transportation. It is partially managed by the distributor and partially by common carriers. The distributor of electrotechnical products owns 55 transportation means, including 38 vans and trucks. In addition a number of external parties providing transportation services equals 21. The age of the fleet ranges from 6 to 12 years and the technical condition of vehicles varies. The fleet operates between different levels of the distribution network, including: suppliers’ level (SL), central level (CL), regional level (RL) and local level (LL). The movements of products in the transportation network are presented in figure 1.
Most of products are transported from SL (75 suppliers) to one of 12 warehouses on the RL and the remaining materials to the CL (central warehouse). Electrotechnical products are transported from CL to RL, from RL to LL (11 warehouses) and to customers (about 400), and finally from LL (11 warehouses) to customers. The final customers are retailers and wholesalers.

The location and the number of the distribution centers are highly correlated with the density of the population in certain regions of Poland. Their location has also a very important impact on the transportation. The longer is a distance between points of origin and points of destination, the more disruptions in the transportation of goods occur. This is the most frequent on the way from suppliers to the warehouses. The average distance between suppliers and warehouses on CL and RL equals 290 km, while the average distance from warehouses on LL to customers equals 42 km. In the last case the disruptions are very rare.

The internal (short-distance) transportation subsystem focuses on forklifts operating in the warehouses. The total number of those transportation means is 26. They operate between the loading / unloading and storage areas, including picking area. The major load unit is a EURO pallet. Its content is usually a mixture of different products depending on their type, size and quantity. The total number of units equals 38,5 thousands. A very precise analysis led to the union of them into 56 groups. Each group has different number of products. The most numerous groups are sockets for light bulbs, wires, and wiring equipment, while to the least one belong disconnectors, heaters, and fans.

The number of people employed in the transportation system is around 90 persons. This number is based on the estimation of the number of vehicles’ drivers and the total number of the forklifts’ operators.

3. Modeling and Simulation Experiments

3.1. Modeling Assumptions

Based on Banks and Gibson [2] recommendations the authors divided the modeling of the transportation system in the following phases:
- analysis of the transportation system,
- definition of the modeling objectives,
- evaluation and selection of the most appropriate simulation tool,
- input data collection,
- determination of the output data required to fulfill modeling objectives,
- construction of the simulation model.

The modeling of a transportation system started from the recognition of its most important features, such as:
- number of vehicles operating in the system, including long-distance fleet and short-distance vehicles,
- characteristics of those vehicles e.g. age, type, technical condition,
- characteristics of goods delivered e.g. type, number, load unit, specific features,
- number of people working in the system,
- characteristics of the transportation e.g. the average distance of transported materials, service area to cover.
The analysis of a transportation system revealed its the most important problems. The authors identified certain disruptions in deliveries, recognized the sources of high costs of transportation etc.

The next step was concentrated on the selection of the most appropriate simulation tool. The major objective was to find an object-oriented simulation package designed for transportation applications. The authors put their attention on the main parameters of the tools as well as on their characteristics and typical applications [3, 8]. The list of simulation tools, initially very long, was reduced to 3 candidates, including: Extend OR, Flexsim and Powersim. All those packages are based on continuous and discrete-event methodologies. They provide hierarchy structure of a model, which is very helpful when modeling of a complex system is considered. Their libraries provide a variety of objects and their structure can be adjusted to specific problems. The final decision was a tradeoff between the performance i.e. input, output and processing capabilities, the simulation environment, and the price. Finally, Extend OR v.6.0 simulation package was chosen. This tool is a user friendly package which can be used to construct complex models in a graphical form without advanced programming skills [6]. It is typically used in transportation, logistics, business processes redesign, manufacturing, as well as in healthcare, service and communications industries.

The complexity of a transportation system causes the problems of the appropriate set of data collection. If the most important data are not included in the model then the most crucial questions might not be answered. On the other hand, when the set of data is too big, too much time will be consumed during the simulation model construction and experiments. Evaluation criteria of the existing transportation system became very helpful in the set of data creation. The authors took advantage of this set also during the determination of the outputs. Finally, the simulation model of the transportation system was constructed. Its details are presented in the next section.

3.2. Simulation model

The model of the existing transportation system starts from establishing technical parameters, such as: global time units, start and end simulation period, number of runs, time windows in which the system operates. Then the location of a transportation system in the distribution network is established. The scheme of the network is presented in figure 2.

The blocks of transportation are presented on each level. They are linked with the remaining blocks by the arrows which represent material flow, while dashed arrows reflect information flow.

The structure of the model is divided between 4 levels i.e. suppliers’ level (SL), central level (CL), regional level (RL), local level (LL). The blocks called “Deliveries’ generator”, “Warehousing” and “Structure of deliveries” are similar as far as their functions are concerned. The “Deliveries’ generator” is responsible for the generation of deliveries. In particular, this block generates impulses i.e. signals of deliveries.
After that the parameters of those impulses are given, which include:
- name of the supplier (from 1 to 75);
- name of the receiver (number 100 for the warehouse on the CL, from number 201 for the warehouses on the RL, from number 301 for the warehouses on the LL, and finally from 1001 for the customers);
- range of the group of products delivered from points of origin to points of destination;
- max. and min. speed of vehicles on the way from points of origin to points of destination.

The block “Structure of deliveries” on the SL is responsible for information processing. Its elements read all data given in the previous blocks and set new parameters needed for the next stage of the model. Those parameters are as follows:
- name of the group of products (from 1 to 56) and its number,
- number of EURO pallets ordered,
- cost of each EURO pallet ordered.

The transportation blocks read data about the distance and speed from points of origin to points of destination. There are also calculated, during the simulation experiments, costs of each delivery, costs of transportation of EURO pallets and the number of EURO pallets.
transported each day. The statistical information, which is the component of this block, accumulate the following data of transportation:
- arrival time of deliveries,
- name of provider and receiver for each delivery,
- name of transported group of products and its number,
- cost of transportation of each group of products.

The generation of information on the CL is based on the demand and its modeling is similar to the deliveries generation described above. The information processing is included in the warehousing block. There are two ways of information and product flow. The first one comes from the SL. The information about the delivery e.g. the name of delivered products, their quantity and characteristics are read. Then the products are unloaded. This process is characterized by the number of loading platforms, available forklifts, the number of available workers, the average labor intensity etc. The products are finally transported to shelves. The second flow line gets an information from the “Deliveries’ generator” block. There is also included the time of picking and the cost of each ordered EURO pallet. Finally, the number of EURO pallets needed for transportation is prepared and the products are delivered to the RL. At this level the deliveries are generated, the warehousing is modeled similarly to the previous level except for the number of product flow lines, which is doubled. It means that 12 warehouses are supplied by the goods delivered from the SL and the CL. Then the products are transported to the LL, from which they are delivered directly to customers.

Concluding this step of modeling it is worth to mention that all the information comes from the historical data. They are presented as a data base on each level of the distribution network. The selected blocks of the simulation model collect those data and use them during the computational experiments. The elements, which have the most important influence on the accuracy of the simulation model are as follows:
- starting points of deliveries and time of incoming order, which are represented by the real data (they can be reflected by a distribution);
- time of the transportation in the external subsystem, which is constrained by the min. and max. value;
- the number of EURO pallets prepared in the picking area for transportation, which is constrained by the incoming order and the capacity of the available vehicles;
- time of the order fulfillment, which is determined by the labor intensity.

3.3. Results of the simulation experiments

After the modeling phase of the transportation system and providing all the needed parameters the simulation experiments have been carried out for a 6 months period. The results generated by the simulation experiments were compared with the real life data characterizing the transportation system and they are presented in the table 1.

The rows with a real data are not fully filled in. It was very difficult to estimate the utilization of forklifts and their operators. According to the information from employees there are queues of vehicles waiting for unloading, but it happens very seldom.

The authors calculated the average difference between the results presenting real data and data from the simulation experiments, which equals 2,8%.
Levels
Parameters (dataoutput data)
Suppliers’ Central Regional Local
Daily number of departures of vehicles at different levels of the distribution system 

35 5 70 50 

36 4 69 50 

Daily number of EURO pallets dispatched from different levels 

115 23 12 10 

116,8 23,2 12,7 10 

Daily unit transportation costs 

4630 860 180 160 

4635 861 200 158 

Daily total distance of deliveries in km 

10181 878,2 2910 2295 

10150 870 2850 2100 

Utilization of forklifts 

- 0,46 0,3 0,2 

The number of forklifts operators / utilization of operators 

- 9/- 36/- 11/- 

9/0,2 36/0,1 11/0,1 

Average queue length of vehicles waiting for unloading 

- 0,2 - 0,6 0,5 

Table 1. The comparison of the real data and the results of simulation experiments

The analysis of the external transportation system indicates that the number of vehicles’ departures per day from different levels of the distribution system ranges from 4 to 69. This difference is caused by the number of customers the warehouses serve for. The number of EURO pallets dispatched on the SL equals 30, while on the other levels it is around 1 EURO pallet. The highest costs of transportation are on the SL and the lowest are on the LL. Those costs are related to the distance of deliveries, which is the highest between suppliers and warehouses on the CL and RL. The minimum distance is between warehouses on the LL and the final customers.

The results of the simulation experiments of the internal transportation system show that the utilization of forklifts is the lowest on the LL. However, in the best case it isn’t higher than 50%. Very similar situation is observed as far as the utilization of forklifts’ operators is concerned. The lowest number of forklifts operators is on the CL and the utilization is the highest.

The point which links external and internal transportation system is the loading and unloading area. The number of vehicles waiting for unloading is the highest on the RL, but it doesn’t exceed 1 truck. All those results are the basis for the conclusions of the condition of the existing transportation system.

4. Conclusions

The existing transportation system of the electrotechnical products is characterized by a high complexity. To understand the interactions between its elements and to evaluate the system the authors constructed a simulation model and carried out simulation experiments. The most important advantages of the simulation tool are as follows:
- ability to model a huge number of elements of the system,
- definition of the relations between all objects in an appropriate way,
- ability to provide output data which are hardly to estimate without simulation experiments.

On the other hand modeling is time consuming and it requires a precision during data collection to eliminate any mistakes. Moreover, the more elements the model contains the more time a single simulation experiment requires.

The adequacy of the simulation model is evaluated on the basis of the difference between its output data and current state of the transportation system. This difference doesn’t exceed 3.5% and in most cases the results are very similar.

The authors also consider the practical aspects of their research. The financial analysis indicates that the costs of transportation to warehouses and to customers belong to the dominant group of expenses in the company. Based on the results of the simulation experiments the authors propose changes to reduce those costs, such as:
- reduction of distances between suppliers and warehouses on CL and RL,
- elimination of a part of deliveries between warehouses on RL, LL and customers,
- reduction of forklifts in the warehouses.

The authors further research will be based on the following steps:
- construction of different paths of changes of the existing transportation system, from evolutionary scenario to the most radical one, which correspond to the restructuring methodologies [4, 10],
- construction of simulation models of those scenarios and carrying out simulation experiments,
- evaluation of the scenarios, including simulation experiments’ results, based on the MCDA methods [9] and choice of the best candidate.

References

INNOVATIVE TECHNOLOGIES SUPPORTING THE DRIVING CREW IN DEGRADED RUNNING CONDITIONS.

G. Sciutto\(^1\), M. Lucchini\(^2\), D. Mazzini\(^2\), C. Veglia\(^1\)

**Abstract.** The issue of the modal rebalance has concentrated the attention of the European states on the standardization of the railway system, in order to allow the convoys to travel more simply all around Europe. It has been reached, as a result of the creation of international work groups, the realization of ERTMS/ETCS System (European Railway Traffic Management System/European Traffic control System) and of the relative technical specifications of interoperability. In order to support the driver in the on sight operating mode, wherein the man has got the full supervision of the train movement, the possibility to apply radar based technologies to all of the “Automatic Train Control” systems has been estimated. These technologies are able to improve the visibility of objects on track and to bypass the geometric distortion problem in the most complete possible way. Moreover, the radar based technology concurs, if compared to the optical based devices that use cameras, to exceed the problems concerned to the reduction of visibility for atmospheric causes. The objective of the study program is to estimate the possibility to apply in particular to the High Speed system, characterized by a high speed rank and by a line architecture quite simple and linear, resorting to a simulation instrument which will concur, opportunely shaped with different scenes, to estimate if could be applied also to different regimes of circulation.

1. **Introduction**

The existing state-of-the-art does not include any on-the-move optical recognition and identification system for use on-board trains, nor any automated active intervention on the train movement to prevent collision with obstacles lying on the track.

\(^1\) Nitel - Consorzio Nazionale Interuniversitario per i Trasporti e la Logistica, Via Opera Pia 1 la Genova - Italy;

\(^2\) CIELI – Centro Italiano di Eccellenza sulla Logistica Integrata, Via Bensa 1 Genova - Italy
Currently, this aspect of safety relies on human observation by the driver, or prior information. There is no real time observation through optical or other means. Automated braking is triggered only when crossing the wrong signalling signs.

2. Application technology to Railway safety

The content of this document explains how the technology can be utilized for support to the railway safety. The application of new technologies is necessary for a greater exploitation of the railway transport, as indicated from European Community politics. The modal re-balance is the unique instrument for adapt to transport increase. Actually, the application of new technologies instruments to the railway transport, it permits the increase of potentiality as well as of safety for this transport.

In particular, the concept of safety, for the railway transport, it is defined from D.P.R. 753/1980. This legal provision describes the aim of the installation of new technologies, that is to avoid the railway incidents and to eliminate fatality accidents.

In particular this technology is utilized in the “on sight” movement.

2.1. The On-Sight operating mode

In the "automatic train control (ATC)" based systems, through a real time cooperation between the train-borne and the trackside equipments, the trackside subsystem is able to manage and supervise the movement of the controlled trains.

This supervision is applicable when the trackside subsystem has got all of the needed information about the track occupancy. If such information aren't completely available, the train has to be switched to the "On-Sight" operating mode, wherein the driver is fully responsible for evaluating the track occupancy conditions, has to recognize the possible presence of obstacles on the track, and must understand the wayside signalling, apart from the sighting distance conditions.

2.2. Automatic Train Control Systems - State of the Art

The issue of the modal rebalance has concentrated the attention of the European states on the standardization of the railway system, in order to allow the convoys to travel more simply all around Europe. It has been reached, as a result of the creation of international work groups, the realization of ERTMS/ETCS System (European Railway Traffic Management System/European Traffic control System) and of the relative technical specifications of interoperability.
A railway line can be qualified as interoperable when it’s equipped (e.g. with the ERTMS) in such way to allow the transit of convoys coming from either various producers or various countries, not requiring the substitution of the engine and of the machinist.

ERTMS system is based on different operating modes corresponding to different allocations of responsibility between the trackside subsystem (SST), that sends to the machinist all the information for the safe conduct of the train, and the trainborne subsystem (SSB), that elaborates the information automatically and regulates the speed of the convoy in a safe way.

The operating mode correspondent to the total allocation of the responsibility to the SST is called Full Supervision; in such condition the system gives the train all the necessary information for the control of the movement.

To the opposite end it is the operating modality defined On Sight, that it is consequence of a degraded condition, characterized by a lack of the availability of the information concerning the state of the way, in which the train staff is responsible to control that the way is free.

By the technological point of view system ERTMS is able to manage the Full Supervision modality in emergency conditions, and is obvious as the ring weak person of the chain is constituted from the absence of technologies for supporting the man in the On Sight movement phase.

This problem doesn’t concerne to the ERTMS system only, but it also concerns all of the other systems used for the railway “in safety” circulation.

2.3. Prospects of Application of sensor based technologies

Currently technologies based on various kinds of sensors are being applied in the automotive field, characterized by usage conditions radically different from the railway field. Particularly, in this field the survey range is greater of a magnitude order because of the reduced wheel/track friction and of the consequent greater space of arrest in case of emergency stop.

The possibility to apply to system ERTMS radar based technologies has been estimated, technologies able to improve the visibility of objects on track and to bypass the geometric distortion problem in the most complete possible way.

The objective of the study program is to estimate the possibility to apply in particular to the High Speed system, characterized by a high speed rank and by a line architecture quite simple and linear, resorting to a simulation instrument which will concur, opportunely shaped with different scenes, to estimate if could be applied also to different regimes of circulation.

References

Abstract. This paper describes the intersection modelling approaches based on the recent second order macroscopic model developments suggested by Aw and Rascle, and Zhang respectively. Some diverge and converge node models are discussed and simulation results are reported. Using the second order model, congestion and spillback are reproduced in a realistic way.

1. Introduction

An intersection model for a node relates boundary states (conditions) of links upstream and downstream of the node, in such a way as to satisfy conservation constraints and supply and demand constraints. Since satisfaction of the constraints leaves the model undetermined, the model must be completed by a phenomenological component which describes the behaviour of users in the intersection (movements, conflicts).

This program has been carried out in Lebacque and Khoshyaran [1] for the LWR (Lighthill-Whitham-Richards [2], [3]) model. In [1] it is also shown that the behavioural component of intersection models must satisfy an invariance principle in order for them to be consistent. Optimality intersection models have been shown to satisfy this invariance principle.

In the numerical analysis community, the same problem has been addressed [4].

The ARZ second order model [5], [6] has some analogy with the first order LWR model, which makes it possible to extend to this model the supply-demand formalism introduced in

1 smammar@inrets.fr, INRETS/GRETIA, 2 av du Gle Malleret-Joinville F-94114 Arcueil France,
2 lebacque@inrets.fr, INRETS/GRETIA, 2 av du Gle Malleret-Joinville F-94114 Arcueil France,
3 haj-salem@inrets.fr, INRETS/GRETIA, 2 av du Gle Malleret-Joinville F-94114 Arcueil France,
The reader is referred to [8] for a detailed analysis. In contrast with current efforts in the numerical analysis community [9] which aim at modelling nodes using a lagrangian discretization of the ARZ model, the object of this paper is to build node models based on the local traffic supply-demand approach, following the methodology introduced in [1].

2. A brief description of ARZ model

The basic equations of the ARZ model are:

\[
\begin{align*}
\partial_t \rho + \partial_x (\rho v) &= 0 \quad (1) \\
\partial_t v + \partial_x p &= -\frac{v}{\tau} \quad (2)
\end{align*}
\]

(in conservative form.), with

- \(x, t\): the position and time,
- \(\partial_t, \partial_x\): the partial derivative with respect to time and space variables,
- \(\rho(x,t)\): the density at time \(t\) and location \(x\),
- \(v(x,t)\): the speed at time \(t\) and location \(x\),
- \(q(x,t)\): the flow at time \(t\) and location \(x\)
- \(y(x,t)\): the relative flow, i.e. the difference between the actual flow and the equilibrium flow,
- \(Q_x(\rho, x)\): the equilibrium flow (at location \(x\)), i.e. the fundamental diagram,
- \(V_x(\rho, x)\): the equilibrium speed (at location \(x\)),
- \(\tau\): a relaxation time constant.

Let us also define the relative speed \(I(x,t)\) at time \(t\) and location \(x\):

\[
I(x,t) = v(x,t) - V_x(\rho(x,t), x)
\]

If we neglect the relaxation term, we can deduce from equation (2) the following (conservation of relative speed along trajectories):

\[
\dot{I} = \partial_t I + v \partial_x I = 0
\]

The relative speed \(I\) is conserved along vehicle trajectories, it is also conserved through 1-waves and across fixed discontinuities.

The complete resolution of the hyperbolic system (eqns. 1, 2) based on the Riemann problem solutions and a numerical scheme are described in Lebacque et al. [8] in the heterogeneous case (fundamental diagram depending on location). The Riemann problem in the homogeneous case is solved in Mammar et al. [10].
In the heterogeneous Riemann problem, the ARZ model (eqns. 1, 2) can be reduced, on the sector of the \((x,t)\) plane on which vehicles coming from the left are located, to the equation (5).

\[
\frac{\partial}{\partial t} \rho + \frac{\partial}{\partial x} \left( \rho v \right) = 0
\]

\[\text{with } v = I_t + V_s(\rho)\]

The resolution of equation (5) is equivalent to solving a first order model with a modified fundamental diagram \((\rho \rightarrow Q_c(\rho, x) + \rho I_t)\).

The methodology for solving (5) has been developed in [3], and is based on the local demand and supply concept. In the case of a Riemann problem with homogeneous initial conditions left \((l)\) and right \((r)\), the demand \((\Delta)\) and supply \((\Sigma)\) functions have the following expressions (shifted demand and supply):

\[
\Delta_{c,*}(\rho,I) = \begin{cases} Q_c, (\rho,I) + \rho I & \text{if } \rho \leq \rho_{\text{crit}}(I) \\ q_{\text{max}}, (I) & \text{if } \rho \geq \rho_{\text{crit}}(I) \end{cases} \quad \text{for } * = l, r \tag{6}
\]

\[
\Sigma_{c,*}(\rho,I) = \begin{cases} q_{\text{max}}, (I) & \text{if } \rho \leq \rho_{\text{crit}}(I) \\ Q_c, (\rho,I) + \rho I & \text{if } \rho \geq \rho_{\text{crit}}(I) \end{cases} \quad \text{for } * = l, r \tag{7}
\]

where:

\[
\rho_{\text{crit}} = \text{Arg}_{\rho \geq 0} \max \left[ Q_c, (\rho) + \rho I \right]
\]

\[
q_{\text{max}} = \max_{\rho \geq 0} \left[ Q_c, (\rho) + \rho I \right]
\]

The flow and the pressure at the origin (similar formula at the boundary between two adjacent cells in the numerical scheme) are given by:

\[
q_0 = \min \left[ \Delta_{c,l}(\rho_1, I_1), \Sigma_{c,r}(\rho_0, I_r) \right]
\]

\[
p_0 = q_0 I_1 \quad \text{with } p_0 = V^{-1}_c(v_r - I_r) \tag{8}
\]

3. **Node modelling approaches**

Usually, in the macroscopic modelling approach, the network is represented, by an oriented graph including links and nodes. The above ARZ solutions are dedicated to link modelling. By considering the same variables, the question is: how to model a node? Elementary nodes can be of two types:

1. Diverge node including one in-link and two or more out-links
2. Converge node including two or more in-links and one out-link
In the following section models for both kinds of intersections are described.

3.1. An elementary diverge node Model

The diverge node is depicted in figure 1. The used variables are: demand at the in-link ($\Delta$), both supplies ($\Sigma_1 = $ out-link1, $\Sigma_2 = $ out-link2).

![Figure 1: Diverge node case](image)

As the relative speed (\(I_i\)) is preserved along the vehicular trajectories, the supplies $\Sigma_1$ and $\Sigma_2$ can be computed easily by (7), with \(*=1 \text{ resp } 2\) and \(I = I_i\). The flow and the pressure have the following expressions:

\[
q_i = \min\{\alpha_i, \Delta, \Sigma_i\}, \quad p_i = q_i I_i, \quad (10)
\]
\[
q_2 = \min\{\alpha_2, \Delta, \Sigma_2\}, \quad p_2 = q_2 I_i, \quad (11)
\]

Where $\alpha_i, \alpha_2$ are given splitting rates. Physically the meaning of this model is that at the diverge, users choose the downstream link more according to current conditions (as given by variable message signs) than final destination. This model can also be used to represent approximately non FIFO behaviour with link choice based on final destination.

3.2. A less elementary diverge node Model

It can be shown that the above model (10)-(11) does not satisfy the invariance principle. This does not preclude its use as a phenomenological model with a fixed discretization step of the ARZ model. But non compliance with the invariance principle implies non-convergence as the resolution of the discretization tends towards 0. A practical consequence is that a model calibrated for a given discretization which contains a non invariant node model must be recalibrated if the discretization resolution is changed.

In order to avoid this problem the following invariant model (an optimization model), which is close to (10), (11), has been developed, following [1]:

573
Max $\sum_i \Phi_i(q_i) + \Psi(q)$

\[
\begin{align*}
0 & \leq q_i \leq \Sigma_i, \quad \forall i, \\
0 & \leq q \leq \Delta \\
\sum_i q_i - q & = 0
\end{align*}
\]  

(12)

The solution of this model is expressed as a function of the Karush-Kuhn-Tucker coefficient $s$ of the flow conservation constraint:

\[
q_i = P_{[0,\ast]}(\Phi^{-1}_i(s)), \quad q = P_{[0,\ast]}(\Psi^{-1}(s))
\]

\[
\sum_i q_i - q = 0
\]

(13)

In the above formula, $P_{[0,\ast]}$ denotes the projection on the interval $[0,\ast]$. In order for the model to be close to the model (10), (11) we choose:

\[
\Phi_i(q_i) = \frac{q_i^{\ast}}, (t) - \frac{(q_i - q_i^{\ast}), (t))^2}{2 \alpha_i},
\]

\[
\Psi(q) = \frac{q^{\ast}}, (t) - \frac{(q - q^{\ast}), (t))^2}{2}
\]

(14)

The coefficient $s$ is determined by solving the flow conservation equation $q_i + q_2 - q = 0$ (piecewise linear in $s$) as described by the following diagram:
3.3. A converge node Model

The convergent node model (figure 2) is non-trivial. The main problem to solve is the computation of the supply ($\Sigma$) which is function of the local ($I$) of the out-link. On the other hand, the local ($I$) is function of the relative speeds ($I_1$) and ($I_2$).

In order to solve this problem, the following heuristic approach is applied. The computation of the relative speed ($I_{eq}$) for the out-link is split into three steps:

- Resolution of the Riemann problem for the in-link1/out-link : $q_1$, $p_1$
  \[ q_i = \text{Min} \{ \Delta, \beta, \Sigma \{ \rho, I_{eq} \} \}, \quad p_i = q_i I_i \]  
  \begin{align}
    q_1 &= \text{Min} \{ \Delta, \beta, \Sigma \{ \rho, I_{eq} \} \}, \quad p_1 = q_1 I_1 \\
    q_2 &= \text{Min} \{ \Delta, \beta, \Sigma \{ \rho, I_{eq} \} \}, \quad p_2 = q_2 I_2
  \end{align}

- Resolution of the Riemann problem for the in-link2/out-link : $q_2$, $p_2$
  \[ q_2 = \text{Min} \{ \Delta, \beta, \Sigma \{ \rho, I_{eq} \} \}, \quad p_2 = q_2 I_2 \]  

- Computation of the ($I_{eq}$) as:
  \[ I_{eq} = \frac{I_1 q_1 + I_2 q_2}{q_1 + q_2} \]  

The coefficients are partial supply coefficients (see [1], [7]). We use (6), (7), (8) in order to carry out these steps. This first estimate of $I_{eq}$ enables us to calculate new estimates of the in- and out flows by a model symmetric to (12)-(14). Note that $\rho = V_c \left( I_{eq} + V_c (\rho - I) \right)$ with $\rho$ and $I$ the downstream state.

4. Preliminary Results

In this section, nodes are modelled using the models of section 3. The developed node modelling approaches are applied in different traffic conditions: downstream free flow and downstream congestion. The ARZ model for links is discretized according to the Godunov scheme [8].
4.1. Diverge node model results

The considered stretch is depicted in figure 1. The in-link traffic demand is fixed as constant and equal to 2160 vh/h. The splitting rate is constant also and equal to 0.75 toward the out-link1. For all links the considered number of lanes is equal to 1. The time horizon and the time slice of the simulation are equal respectively to 1200 and 2 seconds.

With respect to the relative speed \( (I) \), two main scenarios are considered: fixed constant \( I \) and time varying (figure 4). The obtained results are depicted in figure 3 and figure 5 respectively.

Due to the high demand level, the density at the first cell of the in-link is high (figure 3-a). As a main remark, the congestion propagation are performed properly at the out-link1 (figure 3-b) and out-link2 (figure 3-c)

If the relative speed is time varying, a choc wave appears at the in-link (figure 5-a). Despite the constant value of the in-link first cell, the resolution of the Riemann problem leads to the time variation of the supply and demand of the consecutive in-link cells. With respect to the out-links the same density profiles as in the first case are observed.
Figure 5. Time-space density evolution using time varying $I$

Figure 6. Time-space density evolution ($I$ constant)

Figure 7. Time-space density evolution ($I$ time varying)
Figures 6 and 7 represent the density space-time evolutions in the case of apparition of congestion on the out-link1 with $I$ constant and time varying respectively.

4.2. Convergent node model results

The same hypothesis with respect to the relative speed ($I$) is tested for the converge node model. For the converge node, both profiles must be considered. Figure 8 and 10 depict the density evolutions with $I_1$ and $I_2$ constant and time varying (figure 9) respectively.

![Figure 8. Time-space density evolutions ($I$ constant)](image)

![Figure 9. $I_1$, $I_2$ time varying profiles](image)
In both figures (8,10), the spillback of the congestion is represented in a realistic way. As a matter of fact, due to the high level of the traffic volume at both in-links, the capacity at the first cell of out-link drops leading to the apparition of congestion and spill back to the in-link1 and in-link2.

5. Conclusion

The shifted supply and demand (6), (7) provide an adequate tool for computing boundary conditions and intersection models. Special care must be taken when specifying node models: they must satisfy the invariance principle, in order to be stable with respect to the discretization.

The models presented in this paper are based on simple physical ideas: given directional coefficients (demand in diverges) or constant split coefficients of supply (in merges). They can be improved in several directions: increased efficiency of numerical methods, extension to more complex nodes, improvement of the underlying physical ideas.

Acknowledgements

The authors gratefully acknowledge the support of the ACI (Action concertée incitative) NIM (Nouvelles interfaces des mathématiques) “Modélisation mathématique du trafic automobile” of the French Ministry for Higher Education and Research.
References


RISK INDEX MODEL FOR REAL TIME PREDICTION OF POTENTIAL TRAFFIC CRASH ON URBAN MOTORWAY

Habib HAJ SALEM¹, Christophe RAMANANJAONA ², Jean-Patrick LEBACQUE ³, Salim MAMMAR ⁴

Abstract. The paper focuses on the development of a Risk index model. Based on statistical analysis of a coupled database of accidents and traffic measurements, promising results are obtained. This Risk index could be used as off-line safety evaluation index (evaluation process, off-line simulation) or real-time safety index monitoring for user information.

1. Introduction

In the field of safety analysis, the classical traffic evaluation approaches consist in collecting incident/accidents traffic data during the experimented scenarios (traffic control strategies, modification of the infrastructure etc.), and in proceeding to traffic impact and statistical safety analysis of the number of accidents before and after the implementation of these scenarios. Generally, the collection of the accident numbers must get a statistical significance before undertaking an evaluation process. This remark imposes a long time of field data collection (3-5 years), which is the “price to pay” for having a correct safety evaluation.

This paper aims at developing a risk index based on real-data measurements, which can be used either off-line as an evaluation index during the evaluation process leading to the dramatrical reduction of the field test periods, or in real-time like: a safety monitoring tool (e.g. safety user warning system), a multi-criterion function to be optimized in real time.

¹ INRETS/GRETIA, 2 av du Général Malleret-Joinville F-94114 Arcueil France, haj-salem@inrets.fr
² christophe.ramananjaona@inrets.fr
³ lebacque@inrets.fr
⁴ samammar@inrets.fr
(safety index combined with a traffic index) within several control strategies such as coordinated ramp metering, speed limit control, route guidance etc.

The developed index is based on the collection of measured traffic data synchronized with incident/accidents data on two sites in France: the urban motorway A4/A86 and the ring way of Paris.

2. Data base characteristics

For both sites, traffic dataset and accident characteristics are collected from historical database stored in the SISER/SIRIUS and the Ville de Paris operating system. The considered sites are fully equipped with real traffic measuring sensors located at around every 500 meters apart. The incidents/accidents data characteristics include: time of day, location of the accident, weather conditions, severity. The collected traffic data covers 2 hours (one before and one after the crash) at two upstream and two downstream measurement stations (figure 1.), consisting of traffic volume, occupancy rate and speed (if exists). The time intervals of the traffic measurements are equal to six minutes and one minute, for the SIRIUS and the Ville de Paris systems respectively.

The final constituted database includes the overall accidents occurred and traffic data during 4 years (2001-2004) and (2002-2004), for the SISER and Ville de Paris sites respectively. The total number of accidents collected is around 800 on the SISER site, whereas on the ring way of Paris this number is around 900.

![Figure 1. Topology of the considered stretch measurements for each crash](image)

In order to exclude the effect of several factors, the first investigation step is made on a selected number of accidents with the following criterion: same topology (4 lanes), sunny weather conditions and full luminosities (no night-time accidents considered). Among all collected accidents on both sites, the available accidents where dramatically reduced, leading to 50 and 90 accidents selected on the SISER and Ville de Paris respectively.

3. Methodology

The applied methodologies are mainly based on the statistical analysis of the traffic conditions before the accident. A series of multivariate statistical methods are used, with the aim of finding the relationship between the accident and the traffic conditions. Two
well-known statistical methods are applied: cluster analysis and the most common form of factors analysis [3]. In particular, the principal components analysis is applied to find the non-correlated variables to be used for building the risk model. In our case, the total number of variables characterizing the dataset is equal to 4(stations)*2(volume, occupancy rate)*4 (number of lanes) = 32 variables.

For the clustering analysis, several possibilities are investigated:

- Clustering by upstream occupancy rates/lane
- Clustering by downstream occupancy rates/lane
- Clustering by all occupancy rates/lane

The same clustering method is applied for the measurement stations including the four lanes. Lastly, based on the clustering output results, linear regression and non linear logistic modelling approaches are applied for computing the risk index.

4. SISER Summary results

For the SISER site, a strong correlation between the occupancy rates of the two upstream stations, especially between the occupancy rates of the four lanes of station 1 and the occupancy rates of the four lanes of station 2 is observed from the correlation matrix of the dataset [1]. The same remark is valid for the downstream measurements also. These strong correlations are due to the time sampling of the real data measurements. As a matter of fact, the six minutes time interval is very high for a good observation of the dynamic of the traffic characteristics. For this reason, investigations are focused on the ring way of Paris site where the time sampling of the real data collection is equal to one minute.

5. Paris Ring way clustering results

The hierarchical ascending clustering via SAS is performed, using a Ward’s criterion [2], in order to exhibit the particular class of traffic conditions which prevail at the time just before the accident.

5.1. Clustering by all occupancy rate/lane results

In this case, the application of SAS clustering method leads to find five main representative clusters. The first cluster is characterized by a homogeneous average occupancy (MOcc) on the 16 measurement points. The occupancy rates (Occ) are comprised between 9 and 15%, and characterize a low occupancy value and consequently light traffic conditions. This cluster contains 2191 observations and representing 42.96% of all measurements.

Cluster 2 gathers the observations with higher and inhomogeneous average occupancy. Indeed, the MOcc are lower on the fast lane; their values vary around the critical occupancy (18 to 23%). The two central lanes have higher occupancy rates and correspond to unstable traffic states. All lanes of the last station (St4) are congested. The occupancy rates range from 24 to 27.5%. This cluster represents 27.37% of the samples.
As regards cluster 3 (representing 8.8% of the data), a clear transition is observed between the MOcc of the upstream stations, which are very high (36 to 52 %), and the low MOcc of the downstream stations (6 to 11 %).

The MOcc of cluster 4 are homogeneous on the 16 measurement points, with high values, ranging from 30 to 40%. This cluster represents 19.39% of the population. These states correspond to a high level of congestion.

Lastly, cluster 5 is characterized by average upstream MOcc (16 to 21%), particularly on the first two lanes, and very congested downstream (52 to 68%). Moreover, we observe that station 4 is more fluid than station 3. This cluster is the less representative (1.15% of the data).

Screening the time evolution of the clusters (one hour before the crash) of all records (85 in total), 41 accidents indicate a change of cluster during the last six minutes, i.e. 48% of the cases. If only the last observation before the accident is observed, among the total number of accidents, 39 (46%) are moved to cluster 3. Cluster 3 represents upstream congestion and downstream fluid conditions. The risk modelling is based on the traffic state of this cluster.

5.2. Logistic regression

The database is split into two parts. The first half is dedicated to the calibration of the linear regression using SAS tool. The second half is used for the validation of the found risk model.

During the calibration process, the results given by the clustering are used. The logistic regression is performed by considering that cluster 3 presents the highest level of crash. In this case, the risk model value is set to 1, otherwise to 0. Hence the results given by SAS are the following:

\[
R_{\text{acc}} = \frac{1}{1 \cdot \exp\left(-5.7335 + 0.01107 \cdot st1_{\text{to}(1)} - 0.0827 \cdot st1_{\text{to}(3)} + 0.02601 \cdot st2_{\text{to}(1)} + 0.1102 \cdot st2_{\text{to}(3)} + 0.1886 \cdot st2_{\text{to}(4)} - 0.5798 \cdot st3_{\text{to}(2)} - 0.3851 \cdot st3_{\text{to}(4)} - 0.4483 \cdot st4_{\text{to}(2)} - 0.5809 \cdot st4_{\text{to}(4)} - 0.00407 \cdot st2_{\text{q}(2)} + 0.00663 \cdot st3_{\text{q}(2)} + 0.00449 \cdot st4_{\text{q}(1)} \right)}
\]

As indicated, the obtained Risk model includes 13 parameters and 12 variables. According to these numbers of parameters and variables, the use of this model seems to be very complicated. On the other hand, the results obtained during the validation process are not satisfactory. In fact, applying the Risk model on the second half of the database generates a large oscillation of the Risk value between 0 and 1. In order to reduce the number of parameters and variables, the same approach is applied on the aggregate variables by measurement stations. In this case the number of variable is limited to 8 instead of 32.

6. Model based on the clustering by station averages

In order to simplify the model and its interpretation, we aggregate our variables by averaging the occupancy rates on the lanes of a same station and by summing the flows at each station. Our variables then reduce to one flow and one occupancy rate at each station.
The SAS clustering procedure outputs five clusters. Cluster 1 is the most dense (more than 36%). It is characterized by quite homogeneous Occ and average flow over the 4 stations, (an Occ of 11 to 12% and a flow of 1450 to 1500 vehicles per hour and per lane) characterizing fluid traffic conditions. Cluster 2 presents a very high Occ on the (upstream) stations 1 and 2 and rather average downstream (14 to 18%). As for the flow, it is rather stable and low compared to other clusters. This cluster contains 20% of the time steps. Cluster 3 presents high occupancy rates over all stations. The flow is higher upstream. Cluster 4 has average Occ close to the usual 20% critical value, increasing from upstream to downstream (26.7% at station 4). The flows are higher than the other clusters, up to 1774 veh/h/lane at station 2. Cluster 5 has a high average Occ (around 37%) at all stations and a lower flow (around 1230 veh/h/lane).

When we consider the accidents and attribute to each time step the cluster number to which it belongs, we observe that 43 accidents out of 85 studied (50.58%) present cluster change during the last six minutes. For 60 accidents (i.e., more than 70% of them), the last time step belongs to cluster 2, characterized by a rarefaction shock wave (congested upstream and fluid downstream).

6.1. Logistic regression

The same approach as the one previously described is applied. However, the Risk model is set to 1 for the observations belonging to cluster 2 and 0 elsewhere. The calibration of the Risk model is based on 80% of the full observations. The logistic regression model output by SAS is:

$$R_{acc} = \frac{1}{1 + \exp\left(-7.1677 + 0.2122O_{c, st1} + 0.1061O_{c, st2} + 0.2052O_{c, st3} + 0.0003851_{st1}\right)}$$

The dataset rest (20%) of the observations are used for the model validation. The same obtained model is applied on 1000 observations which are not used for the calibration. The output results of this model are depicted in figures 2 and 3.
The time evolution of the Risk index is more realistic. In particular, the oscillations are removed. These results are promising. The Risk index value is high (around 1: maximum value) 3 minutes before the occurrence of the accident and this value is maintained until the final time of the accident.

7. Conclusion and next steps

The obtained results are very promising. The number of parameters was limited to 5 which can minimise the effort for the calibration process. However, more investigations are needed in order to take into account other parameter conditions such as the weather, luminosities (night) and the modification of the geometric topology (different lane number of the upstream and downstream measurement stations).

References

EXISTENCE AND STABILITY OF EQUILIBRIUMS IN REGIONAL TAX COMPETITION

Megan M. KHOSHYARAN1, Jean-Patrick LEBACQUE2

Abstract. In this paper a simple model for tax competition between countries is described and analyzed. Due to the effects of variable population densities, complex transportation costs, and consumer preferences, the behavior of the system is very complex. A general framework is developed to analyze this behavior, as well as simplified analytical and semi-analytical models.

1. Introduction

Within recent years tax competition has increasingly become an instrument of economical concurrence between regions or countries. The growing diversification of authorities and the resulting jurisdictional competition leads to problems which are similar in structure. The basic question is: does such a competition between conflicting authorities lead to stable and effective operations. The tax competition problem can be considered as an archetype for complex systems with two operators or players.

Economists have been studying the problem of tax and toll competition for a long time, see [1], [2], [3] for example. Nevertheless simplifying hypotheses, notably in the pioneering paper [1] by Kanbur and Keen, have a strong impact on the results. What are the possible effects of tax and toll competition between countries, regions or operators on the system? Can equilibrium be reached? The aim of the paper is to provide some elements of answer to this question.

1 etclinic@wanadoo.fr, ETC Economics Traffic Clinic, 34 avenue des Champs-Elysées, F-75008 Paris France,
2 lebacque@inrets.fr, INRETS/GRETIA, 2 avenue du Général Malleret-Joinville F-94114 Arcueil France.
2. The model

2.1. Introduction

The model is based on the early work of Kanbur and Keen [1], expanded in [4]. Elements for treatment of intra-regional competition with more players can be found in [5].

The model considers two countries (regions) $i = 1$ and 2 and addresses the problem of trans-border shopping (or any other taxable economic activity). The attractiveness of activities is affected by the tax level $t_i$ imposed by each country ($i$), and each user maximises his net value (the value of the activity including tax minus the cost of trip to access the activity). Trip cost is neglected when the customer stays in his country.

The model can be made one dimensional as shown in [1] and [4], with the cost of trip to the border as the fundamental factor. The population density being $\rho_i(x_i)$ and the density of users at distance $x_i$ to the border with travel cost $\tau$ being $\varphi_i(x_i, \tau) d\tau$, the density of the population with respect to the travel cost to the border is given by:

$$P_i(\tau_i) d\tau_i = d\tau_i \int_{x_i}^{\infty} \rho_i(x_i) \varphi_i(x_i, \tau_i) d\tau_i$$

Further, the variability of travel behaviour with respect to activity value can be taken into account, by introducing the probability of making a trip to the other country given a trip cost $\tau_i$ as a function of tax difference $t = t_i - t_j$:

$$\text{Prob}[i \to j / \tau_i] = G_i(t - \tau_i)$$

As suggested in [4], a logit model could be used in this context. Defining $g_i = G_i'$, a modified population distribution is obtained:

$$\Sigma_i(z_i) dz_i = d\tau_i \int_{(i)} g_i(\lambda) P_i(z_i - \lambda) d\lambda$$

which yields the following expression for the number of customers of region ($i$) who buy the commodity in region ($j$) as:

$$\int_{-\infty}^{z_i} \Sigma_i(s) ds$$

Let us denote (1) the region on the left-hand side of the border, and (2) the region on the right-hand side of the border, and the equivalent modified population density is introduced:

$$\Sigma(s) = \Sigma_i(-s) + \Sigma_j(s)$$

The figure below illustrates this concept. The border is located at the origin.
Note that the population is grouped in clusters (urban communities).

Let us introduce further the total population \( N \) of the two regions:

\[
N = \int_\infty^{\infty} \Sigma(\zeta)d\zeta
\]

and the population distribution function:

\[
n(t) = \int_\infty^{\infty} \Sigma(\zeta)d\zeta
\]

Then the number of customers who pay taxes in country \((i)\), \( N_i(t_1,t_2) \), is given by:

\[
\begin{cases}
N_1(t_1,t_2) = N - n(t_1 - t_2) \\
N_2(t_1,t_2) = n(t_i - t_2)
\end{cases}
\]

The tax revenue in region \((i)\) is given by

\[
R(t_1,t_2) = t_i N_i(t_1,t_2)
\]

It follows that the problem of tax competition can be parameterized by the tax difference \( t = t_i - t_j \). This fact is fundamental.

2.2. Reaction functions

Revenue functions are not concave, as the example below shows (figure 2), displaying a cross-section of the revenue of country 1 (tax 2 given). Therefore global optima are impossible to characterize, and the Nash equilibrium can only be local. The reader is referred to [3], in which this concept is systematically applied under the terminology second order locally consistent equilibrium, and also to [4].
In other terms, the local first order optimality conditions for the revenues are expressed with respect to the parameter $t$ (the tax difference) and yield so-called reaction curves,

$$
\rho_i(t) = \begin{cases} 
\rho_{ii} = \frac{N - n(t)}{\Sigma(t)} \\
\rho_{ij} = \frac{N - n(t) - \Sigma(t)}{\Sigma(t)} 
\end{cases}, \\
\rho_j(t) = \begin{cases} 
\rho_{jj} = \frac{\Sigma(t) + n(t)}{\Sigma(t)} \\
\rho_{ji} = \frac{n(t)}{\Sigma(t)} 
\end{cases}
$$

The meaning of the reaction curves is the following.

- **Reaction curve $\rho_i$:** If the country 2 applies toll $t_2 = \rho_{2i}$, then country 1, in order to optimize (locally) its revenue, should apply toll $t_1 = \rho_{1i}$. Implicitly and formally, $t_1$ is a function of $t_2$: $t_1 = \rho_{1i}(\rho_{2i}(t_2))$. This relationship is mediated by parameter $t$: $t_1 = \rho_{1i}(t)$ and $t_2 = \rho_{2i}(t)$.

- **Reaction curve $\rho_j$:** Its interpretation is symmetric of the interpretation of the reaction curve $\rho_i$. If the country 1 applies toll $t_1 = \rho_{1i}$, then country 2, in order to respond optimally, should apply toll $t_2 = \rho_{2i}$. Implicitly $t_2$ is a function of $t_1$: $t_2 = \rho_{2i}(\rho_{1i}(t_1))$. This relationship is mediated by parameter $t$: $t_1 = \rho_{1i}(t)$ and $t_2 = \rho_{2i}(t)$.

Note that $\rho_{1i}(t) - \rho_{2i}(t) = \rho_{2i}(t) - \rho_{2i}(t) = t$. 

Figure 2
Actually, (10) expresses 1\textsuperscript{st} order optimality conditions, which must be completed by 2\textsuperscript{nd} order conditions (local concavity of $R_1$ and $R_2$):

$$\frac{\partial^2 R_1}{\partial t^2} \leq 0, \quad \frac{\partial^2 R_2}{\partial t^2} \leq 0.$$ 

On the reaction curves, the following expressions result:

$$\frac{\partial^2 R_1}{\partial t^2} = -2\Sigma(t) \frac{(N - n(t))\Sigma'(t)}{\Sigma(t)} = \Sigma(t) \frac{\partial \rho_1}{\partial t},$$

$$\frac{\partial^2 R_2}{\partial t^2} = -2\Sigma(t) + n(t)\Sigma'(t) \frac{\Sigma(t)}{\partial t} = -\Sigma(t) \frac{\partial \rho_2}{\partial t}.$$ 

We introduce the following quantities, called discriminants, which when positive characterize the fact that revenues reach a local maximum:

$$\Delta_1(t) = 2\Sigma'(t) + (N - n(t))\Sigma'(t)$$

$$\Delta_2(t) = 2\Sigma'(t) - n(t)\Sigma'(t)$$

Potential equilibriums are obtained by intersecting the reaction curves. Indeed such intersection point will have a value of $t$ common for both $\rho_1$ and $\rho_2$. Thus a local Nash equilibrium is characterized by $\rho_{1i} = \rho_2$ or $\rho_{2i} = \rho_1$ and positive discriminants:

$$\Xi(t) = 2n(t) + \Xi(t) - N = 0$$

$$\Delta_1(t) \geq 0, \quad \Delta_2(t) \geq 0$$

The equation in (13) admits at least one solution (possibly many) as shown by the plot of the $\Xi(t)$ function (figure 3) corresponding to the population density of figure 1:

![Figure 3](image)

The behavior of the discriminants is also irregular implying that possibly all solutions to the equation in (13) actually do not satisfy the inequalities in (13) and there may be no Nash equilibrium at all (even local).
3. Existence and quality of equilibriums

3.1. Complex behavior

The situation analyzed by Kanbur and Keen [1] was very simple (only the respective population size was taken into account) thus leading to simple behavior with respect to equilibriums.

As shown in [4], the more complex setting described in this paper, in which population density, travel costs and consumer preferences are taken into account, leads to complex behavior. This behavior is illustrated by the end of the previous section. It is also illustrated in figure 5 by the reaction curves obtained for the population of figure 1 (reaction curve 1 in blue, reaction curve 2 in red).

For any given toll value imposed by country 1, there are many possible responses of country 2. Let us list the main reasons for complex behavior:

- non concave revenue functions
- complex reaction curves, implying multiple-valued reaction functions
- reaction curves intersect at least once, they may intersect many times
- intersection points of reaction curves do not necessarily satisfy the inequalities in (13) (non positive discriminants)
- intersection points of reaction curves satisfying (13) may correspond to unstable local Nash equilibriums.
3.2. Simplified analytic approaches

A first possible approach to simplification is to consider piecewise constant population cluster densities. Discriminants are positive; the reaction curves are lines which may intersect. The possible intersection points are then stable local Nash equilibriums, as shown by the following figure:

In figure 6 iterates of optimal toll values chosen by each country in turn are shown to induce a sequence on the reaction curves that converges towards the local equilibrium.
A better way to approximate any population density is to consider piecewise linear clusters. Such an approximation can be made uniformly as precise as required up to first derivatives of $\Sigma(t)$. The complex behaviour described above becomes tractable in this setting. In particular, it can be shown that local Nash equilibriums satisfying to (13) belong to 4 classes and depending on the class and some numerical values the equilibriums may be stable or not.

4. Conclusion

As pointed out in [4], the complex and possibly unstable behaviour of the system has important economic and practical consequences. It pleads in favour of cooperation and negotiation between the agents, as the only practical way to control the system. Further, the system presented in this paper is an archetype for many economic and transportation systems with competition, which depend critically on the difference between two control variables, say tolls or taxes. For instance, transportation systems with two competing operators would fit into the framework developed into this paper and could be analyzed accordingly.

Acknowledgements

The authors gratefully acknowledge the support of the ACI (Action concertée incitative) NIM (Nouvelles interfaces des mathématiques) “Modélisation mathématique du trafic automobile” of the French Ministry for Higher Education and Research.

References

TRAFFIC FLOW MODELLING AND INTER-VEHICLE SPACING CONTROL IN THE MONT BLANC TUNNEL

Vincent HENN*   Jean-Baptiste LESORT*

1. Introduction

In the long road tunnels crossing the Alps Mountains, traffic control and enforcement have taken an increased importance after the Mont-Blanc accident in 1999, which resulted in 39 fatalities. Speed control has been implemented for long, together with a flow control at the tunnel entrances, but it is only now that technology makes it possible to control and enforce the vehicle spacing regulations which have been decided after the 1999 accident.

The idea of controlling vehicle spacing raises several questions:

- How can flow control at the tunnel entrance, combined with speed control inside, provide the best conditions for users to comply with spacing regulation?
- What can be the effect on traffic flow of a local spacing enforcement system if spacing rules are not respected upstream (congestion inside the tunnel)?
- What can be the effect of slow vehicles on spacing?

In order to give a first answer to these questions without involving huge studies or experiments, it has been decided to develop a very simple traffic model devoted only to evaluate these questions.

2. Description of the model

2.1. Why a macroscopic model?

The first idea was to use a microscopic car-following model, able to take account of the vehicles diversity and to represent accurately the motorists behaviour vs spacing rules. Unfortunately numerous problems were raised by this microscopic approach:

* LICIT (INRETS-ENTPE), rue Maurice Audin, 69518 Vaulx-en-Velin cedex, France, vincent.henn@entpe.fr, jean-baptiste.lesort@inrets.fr
Calibrating motorists behaviours is very difficult, and there is little possibility to make measurements, and particularly spacing measurements, inside the tunnel; The available data are mainly based on users inquiries and give only qualitative rules.

It seemed difficult to draw conclusions from a limited number of scenarios based on microscopic simulations, for the effect of changing a parameter in such a model is not straightforward.

It was decided to use a completely different approach, based on a simple macroscopic model.

The model used is a mere Lighthill-Witham-Richards [1,2] model. The first interest of such a model is that the phenomenology of traffic is totally described by a flow/density (or speed/spacing) relationships and that it is thus easy to represent various user compliance to spacing rules. The second interest is that extensions have been made to this model to represent the interaction of the flow with moving obstacles [3], which happens to be the main modelling question in tunnels like the Mont-Blanc: most traffic perturbations are caused by vehicles having lower acceleration possibilities and causing traffic accumulation behind them.

These two points will be explained in the next part and some results given.

### 2.2. The flow density diagram

It is agreed by the tunnel operators that speed limits are mostly respected. The maximum speed is 70 km/h in the tunnel, and there is a minimum cruising speed of 50 km/h.

The spacing rules assume that vehicles must keep between them a distance of 150 m and 100 m when stopped (note that this safety distances are not related to collision prevention, but to fire security). The flow density diagram corresponding to a perfect compliance with these rules is thus the following:

From a user survey conducted at several long road tunnels [4], the following observations have been made:
About one third of the users comply with these rules,

One third respect spacing of about 40 to 50 meters, decreasing when the flow speed is low,

One last third use very short distances, even at maximum speed.

The corresponding diagrams are triangular diagrams with different values for critical and maximal densities. The combination of the three types of users yields to the following fundamental diagram:

Other different combinations are possible and provide various types of compliance with the spacing rules. This makes it possible for instance to represent a stricter compliance in the vicinity of a spacing control system.

2.3. Influence of slow vehicles

Some vehicles in the flow have a special behavior and their speed is not restricted by traffic itself, but rather by their mechanical characteristics. Those "slow vehicles" are not necessarily constrained in speed (and may be fully capable of reaching the speed limit of 70km/h) but may have some poor acceleration capabilities so that at the entry of the tunnel their speed is particularly low. Therefore they behave like a moving bottleneck.

The idea of the proposed model is to represent exogeneously the trajectory of those "slow" vehicles while the remaining traffic is described by the LWR model.

3. Scenario analysis

Several theoretical scenarii can be build in order to answer the different issues raised by the vehicle-spacing control problem.

For example let us consider the influence of a zone where the inter-vehicle spacing is controled. Inside that zone the traffic is supposed to follow strictly the inter-spacing rules and use the first fundamental diagram, while outside the zone, the second diagram is used.
Let us further consider a truck driving off after the toll gate. Because of its limited acceleration, it will not reach the maximal speed before a certain distance. During the time of its acceleration, some "faster" vehicles will accumulate after it causing some congestion. At the entry of the control zone, this platoon cannot pass through and generates stronger congestion (see figure below).

Space-time diagram around a zone where inter-vehicle spacing is controlled (the darker the color, the denser the traffic)

Hence it can be seen on such a simple example that spacing control enforcement can actually lead to stronger violation of rules, with an opposite effect to the objective of the security spacing rules.

4. Conclusions

Such simple traffic model like the proposed one can lead to several simple yet important conclusions. In particular, while the natural behaviour of traffic is far from the strict respect of the spacing rules, there is congestion (and rules violation) upstream any spacing control zone. Two types of solution can be envisaged in order to avoid such a problem:

- Help the "natural" behaviour of traffic to get closer to the safety rules (by mean of a better information given to drivers: Why those rules? How to measure spacing? How
to maintain safe spacing...), that is change the fundamental diagram. But this issue is not really a traffic management one, but rather a psychologist one. From a traffic point of view, its is nevertheless interesting to note that a platoon caused by a limited acceleration vehicle is dissipated by a control zone and will not reappear after the control zone as far as the speed limit is verified (which is more or less observed and can be enforced by speed control). Hence a spacing control zone helps the traffic to verify spacing rules, even downstream that zone.

- A second important solution lies in the management of the vehicles entrance inside the tunnel (at the toll gate for example). Especially, this management has to be further investigated around limited acceleration vehicles.

References


A SHOCKWAVE-BASED METHODOLOGY FOR RELIABLE ESTIMATION OF ORIGIN-DESTINATION TRAVEL-TIME IN URBAN AREAS

by

Manoj K. JHA¹, Franklin OKONKWO², Wolfgang KUHN³

Abstract.

Reduced travel-time delay is a key measure of effectiveness between specified origin-destination (O-D) pairs in the morning rush hour in urban areas. Reliable travel-time estimates are critical to analyze the traffic network efficiency and effectiveness. In this paper, we develop a reliable estimation procedure for travel-time delay using a shockwave-based methodology for a typical driver. From the study of the commuting pattern of typical drivers we divide the commuting path into three segments, to separately calculate travel-time delay into those segments for improved reliability. Segment I represents the residential streets, Segment II represents freeways while segment III represents the central business districts. Formulations for travel-time delay calculation in the three segments are presented. A sensitivity analysis is performed to examine the relationship between commuting start-time and travel-time delay. It is found that: (1) the proposed segmentation approach may improve the reliability of travel-time estimation; (2) in the morning rush hour travel-time delay varies non-linearly with the start-time of the commute; and (3) shockwave methodology generally overestimates the travel-time delay in segment II (freeways).

Introduction

In recent times, there has been increased desire for an improved estimate of travel-time of an average driver between given origin and destination (O-D) pairs. Travel-time delay and queue lengths are generally estimated by cumulative arrival-departure models and shockwave methodology [4-6, 10, 12]. In this paper we focus on improving the reliability of travel-time delay calculation in the morning rush hour for an average driver. In order to accomplish this we examine the nature of the traveling path of an average driver. Typically, an average driver leaves his/her residence and navigates through arterial streets before getting on a freeway and finally takes the arterial streets again in the vicinity of his workplace to eventually arrive at the workplace (which is typically in a Central Business District (CBD)). Thus, there are three distinct segments of the commuting path: arterial residential streets, freeway, and arterial streets in the vicinity of the CBD. We label these as segments I, II, III, respectively and show them in Figure 1.

1 Morgan State University, Baltimore, MD, USA, mkjha@eng.morgan.edu
2 Morgan State University, Baltimore, MD, USA, frankolisa@yahoo.com
3 University of Leipzig, Germany, kuehn@wifa.uni-leipzig.de
A typical driver may face different congestion effects in the three traveling segments. For example, in segment I he/she may be stuck behind a slow forming queue due to intermittent stoppage of school buses. Please note that U.S. traffic regulations require traffic to stop on both sides of the street when a school bus is stopped for boarding and alighting young students. It is common to have recurring and/or non-recurring congestion in segments II and III. Due to the speed, flow, and congestion variability in the three segments it is desirable to study the travel-time delay effects separately.

Figure 1. Segments of an O-D path of an average driver

The objective of this paper is to develop shock waves in segment I due to the intermittent school bus stoppage and also perform a sensitivity analysis to examine the effect of commuting start time on travel-time delay. A precise calculation of the travel-time delay can help accurately estimate origin destination trip pattern as well as predict to a reasonable extent, an expected time for an average driver to reach his destination.

Shockwave Application in Modeling Traffic Delay

Many shockwave applications for modeling traffic delay have been reported in the literature [4-6, 10, 12]. The queuing phenomena along freeways have been studied using the shock wave theory [6]. However, its application to calculate arterial delay has been limited. Wirasinghe [12] used shockwave theory to develop formulas for calculating individual and total delays upstream of incidents. The formulas were based on areas and densities of regions representing different traffic conditions (mainly congested and capacity regions) that are formed by shockwaves in the time-space plot. Chow [2] compared two methods for calculating total incidence delay on a highway section: shockwave analysis and queuing analysis. He assumed a unique flow-density relationship and derived the equation of total delay, which were found to be identical for both methods. He concluded that if he had used a time-dependent flow-density relationship, which is more realistic, then the two methods would yield different results. McShane and Roess [8] suggested that queuing analysis and shockwave delay methods yield different results when applied to the analysis of the bottleneck. They showed that the use of deterministic queuing analysis method might underestimate the overall magnitude of delay when compared to the shockwave method. Chin [1] indicated that both queuing and shockwave analysis yield similar result, provided
the problem under consideration is modeled appropriately in each method. The validity of his conclusion is weakened by the fact that he used the actual number of vehicles that arrived at the study bottleneck to make its comparison rather than the expected number of vehicles causing its proposed shockwave model to underestimate the delay. Michalopoulos et al [9] presented a mathematical model for the formation and dissipation of traffic queue at signalized intersection. The analysis presented possible combinations of shockwave development that are most likely to occur at the intersection. Hurdle and Son [5] showed that although shock wave and cumulative arrival and departure curve method define queue length differently, both models when used together instead of individually, gives a greater level of understanding of what is happening on the roadway. Computation of queue length at traffic signal is briefly discussed by Lighthill and Whitham [7] and analyzed in more details by Rorbech [11]. No particular application was proposed by Lighthill and Witham while Rorbech’s analysis is limited to a linear case and is based on geometric arguments that are not entirely rigor.

**Shockwaves Due to Intermittent Stoppage**

Most of the shockwave applications [5, 6, 10, 12] for modeling travel-time delay are due to lane blockages at a fixed time. It is assumed that after temporary shutdown the blockage is removed resulting in a dissipation of the queue. Such phenomena are common along freeways where a sudden lane blockage may occur due to an accident. In the morning rush hour along arterial streets queue may be formed due to intermittent stoppage of school buses. Several shock waves may be generated in those situations and due to the intermittent stoppages traffic may not attain free-flow speed resulting in considerable increase in the travel-time delay. In the literature we found only one study [4] in which arrival-departure pattern due to intermittent stoppages and time shifts as well as generated shockwaves are described. For brevity here we describe an arrival-departure pattern due to two school bus stoppage (Figure 2). It is assumed that the queue clears at every stoppage before entering into subsequent stoppages. More complex scenarios may arise, such as queue spill back at subsequent stoppages under oversaturated conditions. Such scenarios will be studied in future works.

From the geometric relationship of the triangles (Figure 2) we can determine the maximum queue length, the time of queue dissipation as well as the total and average delay experienced. Considering the first stop, the time required for the queue formed to dissipate is given by

$$t_D = \frac{q_{d1} \cdot (t_{1-db} - t_{1-a})}{1 - \frac{q_{d1}}{q_{d1}}}$$  \( (1) \)

Maximum delay experienced by individual vehicle is given by

$$d_m = t_{1-a} - t_{1-db}$$  \( (2) \)
Figure 2. Arrival departure pattern due to two school bus stoppages

Total time of vehicle delay (D) is represented as

\[ D = q_{a1} t_{1-a} - q_{d1} t_{1-de} \]

(3)

The maximum number of vehicles in queue

\[ q_a = q_{a1} t_{1-db} - q_{d1} (t_{1-de} - t_{1-a}) \]

(4)

The maximum number of delayed vehicles is

\[ q_d = q_{a1} (t_{2-db} - t_{1-a}) + t_{a} \]

(5)

The average individual vehicular delay is given by

\[ d = \frac{D}{q_a} \]

(6)

For N number of stops, average delay experience is given by

\[ \Delta t' = \frac{\sum_{i=1}^{N} d_i}{N} \]

(7)

where: \( N \) = Number of bus stops; \( d_i \) = Delay experienced by vehicle i
Solution Approach

For the commuting path represented in Figure 1, we represent the free flow travel-time (uncongested) for the O-D path as

\[ T = t_1 + t_2 + t_3 \]  
(8)

Considering the delay experienced in each segment, the actual travel-time is represented as

\[ T_a = (t_1 + \Delta t_1) + (t_2 + \Delta t_2) + (t_3 + \Delta t_3) \]  
(9)

where \( \Delta t_1 \), \( \Delta t_2 \) and \( \Delta t_3 \) represent the travel-time delay in the three segments, respectively. \( \Delta t_1 \) can be obtained from Eq. (7). \( \Delta t_2 \) can be calculated as:

\[ \Delta t_2 = \frac{-W_r (t_1 - t_2)}{2q_2} \left[ t_2 (k_s - k_{sc}) - t_1 (k_{sc} - k_{sc}) \right] \]  
(10)

where \( W_r \), \( q_2 \), and \( k_s \) represent speed, flow, and density, respectively. An explanation of Eq. (10) is provided in Rakha and Zhang [10].

\( \Delta t_3 \) can be calculated by Eq. (6).

Example

In this example we calculate the travel-time of an average driver who commutes to work daily from Rosedale area to Essex in Baltimore Maryland. The driver leaves his home at 7:00 a.m. and is faced with school bus delays along the arterial streets in segment I, recurrent congestion delays on the freeways in segment II and delay due to heavy traffic along the arterials in segment III. Average speed, flow, density, and other relevant data were collected along the commuting path. Travel-time delays in the segments were calculated using the above equations. Following results were obtained:

\[ t_1 = 6 \text{ minutes}, \ t_2 = 15 \text{ minutes}, \ t_3 = 10 \text{ minutes} \]
\[ \Delta t_1 = 4 \text{ minutes} \ \Delta t_2 = 20 \text{ minutes} \ \Delta t_3 = 2 \text{ minutes} \]

The actual travel-time of the driver for the commuting path under consideration is given by

\[ T_a = (t_1 + \Delta t_1) + (t_2 + \Delta t_2) + (t_3 + \Delta t_3) \]
\[ = (6 + 4) + (15 + 20) + (10 + 2) \]
\[ = 57 \text{ minutes}. \]

If for the travel path under consideration, no delay was experienced, then the travel-time for the same path is given by

\[ = 6 + 15 + 10 \]
\[ = 31 \text{ minutes}. \]
Sensitivity to Starting Time of Commute

In order to examine the sensitivity of travel-time delay on starting time of commute relevant data (including speed, flow, density, and travel-time delay) were collected at 5 minute intervals between 7:00-9:00 a.m. (see, Table 1) The estimated travel-time delays obtained from the above equations were plotted against the observed travel-time. It was found that: (1) travel-time delay varies non-linearly with the commute start-time; (2) highest travel-time variation is noted in segment II; (3) shockwave methodology generally overestimates the travel-time delay in segment II.

Conclusion
A shock wave based methodology has been developed to estimate travel-time delays for a typical driver in the morning rush hour. The proposed segmentation approach allows a reliable calculation of travel-time delays in urban areas. The results suggest that while school bus stoppage significantly affects the travel-time delay, the largest share of the delays come from recurring freeway congestion. It is also found that travel-time increases non-linearly with commuting start-time between 7:00-9:00 a.m. until about 8:10 a.m. after which it starts decreasing. A reliable travel-time delay estimation can help accurately estimate origin-destination (O-D) trip pattern as well as provide some level of assurance of the expected time to reach a given destination. This finding will be valuable for transportation planners and traffic engineers in improving traffic operation. In future works, extensive sensitivity analysis using real-world data will be carried out. The results will be compared against that obtained from standard traffic simulation software such as Corsim.

Table 1. Travel-Time data for the sensitivity analysis

<table>
<thead>
<tr>
<th>Time Interval (a.m.)</th>
<th>Observed Travel-Time Delay (minutes)</th>
<th>Estimated Travel-Time Delay (minutes)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Segment I</td>
<td>Segment II</td>
</tr>
<tr>
<td>7:00</td>
<td>2</td>
<td>12</td>
</tr>
<tr>
<td>7:05</td>
<td>2</td>
<td>14</td>
</tr>
<tr>
<td>7:10</td>
<td>4</td>
<td>14</td>
</tr>
<tr>
<td>7:15</td>
<td>4</td>
<td>18</td>
</tr>
<tr>
<td>7:20</td>
<td>5</td>
<td>18</td>
</tr>
<tr>
<td>7:25</td>
<td>8</td>
<td>20</td>
</tr>
<tr>
<td>7:30</td>
<td>9</td>
<td>25</td>
</tr>
<tr>
<td>7:35</td>
<td>11</td>
<td>25</td>
</tr>
<tr>
<td>7:40</td>
<td>11</td>
<td>27</td>
</tr>
<tr>
<td>7:45</td>
<td>12</td>
<td>27</td>
</tr>
<tr>
<td>7:50</td>
<td>15</td>
<td>30</td>
</tr>
<tr>
<td>7:55</td>
<td>11</td>
<td>31</td>
</tr>
<tr>
<td>8:00</td>
<td>10</td>
<td>31</td>
</tr>
<tr>
<td>8:05</td>
<td>10</td>
<td>34</td>
</tr>
<tr>
<td>8:10</td>
<td>10</td>
<td>35</td>
</tr>
<tr>
<td>8:15</td>
<td>10</td>
<td>35</td>
</tr>
<tr>
<td>8:20</td>
<td>10</td>
<td>38</td>
</tr>
<tr>
<td>8:25</td>
<td>6</td>
<td>30</td>
</tr>
<tr>
<td>8:30</td>
<td>7</td>
<td>25</td>
</tr>
<tr>
<td>8:35</td>
<td>7</td>
<td>20</td>
</tr>
<tr>
<td>8:40</td>
<td>4</td>
<td>15</td>
</tr>
<tr>
<td>8:45</td>
<td>4</td>
<td>10</td>
</tr>
<tr>
<td>8:50</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>8:55</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>9:00</td>
<td>1</td>
<td>4</td>
</tr>
</tbody>
</table>
Figure 3. Comparison of observed and estimated travel-times

References

PIACON: ROBUST VEHICLE TRAJECTORY BASED ARTERIAL MULTI-CRITERIA TRAFFIC SIGNAL CONTROL

Andrzej ADAMSKI

Abstract. New robust representation of real uncertain traffic phenomena by family of vehicle trajectories models estimated in real-time by dedicated robust estimators is presented. It offers unique capabilities of modeling and predicting complex traffic behavior to be result of hybrid vehicle-trajectories interactions. These models are used in PIACON method for multi-criteria vehicle trajectory based robust traffic control on an artery. To illustrate the robust nature of controlled traffic processes the control reactions on discrete events generated by controller actions are presented.

1. Introduction

The nowadays high spatial resolution satellite and air based imagery integrated with GIS platform may be a vital source of information about vehicle trajectories random field (VTRF). The available 1m spatial resolution data make it possible precisely assign the detected and tracking vehicle trajectories to appropriate lanes. The aerial video recorded from helicopters equipped with digital video camera and GPS receiver creates another remote traffic sensing platform. The wide spectrum of available conventional means of traffic information collection (e.g. GPS-related locations, video-detectors, lasers, vehicle probes, radar) create supplementary data sources for estimation of VTRF. The real-time estimation of city operational related environment traffic paths characteristics (e.g. paths flows, O-D volumes) together with estimation of city-region traffic flows characteristics (i.e. VTRF oriented intersections, routes, sub-areas level-of-service measures) create the most natural traffic representation for surveillance, control and management purposes. The first original implementation of VTRF traffic representation for multi-criteria real-time adaptive traffic control purposes was developed in PIACON (Polyoptimal Integrated Adaptive CONtrol) method [1-5][7][9-11]. In paper [11] it was proposed to integrate multiple existing surveillance platforms (airborne, super-network, conventional) to real-time accurate estimation and prediction of VTRF realizations from the traffic control, management and congestion

1 Institute of Automatics, AGH University of Science and Technology, 30-059 Cracow Al. Mickiewicza 30, Poland., aad@ia.agh.edu.pl
mitigation implementation point of view. It is evident that many traffic problems can be solved sensibly only on the basis of VTRF traffic representation [9-10]. The dynamic traffic assignment problems are a good examples. VTRF representation of traffic in control enables to use completely new travel smooth measures as performance criteria. It create the possibility of direct representation in control problems until now not available LoS (Level of Service) standards represented in natural way by family of trajectories and their characteristics (e.g. various trajectories related measures representing stops, delays, queues, unsafe interactions, driver discomfort) in terms of route segments and even O-D [2]. In addition by robust VTRF trajectories estimators the system uncertainty of various parts of transportation system (e.g. links, intersections, arteries, sub-areas) may be decreased. The operational knowledge included in VTRF gives unique opportunity for real-time intelligent adaptive traffic robust control actions dedicated to real-time on-line recognized operational functionality of various parts of the network e.g. completely novelty PIACON artery multi-criteria traffic progression synchronizing control with VTRF related on-line real-time estimated control performance measures. The pro-ecological and safety control modes offered by PIACON uses essentially the VTRF knowledge about types of vehicles [6][9]. The congestion mitigation efficiency may be very high due to exploration of VTRF knowledge concerning trajectory related congestion generators. In this context congestion avoidance (CAV) PIACON mode reduce in anticipative way congestion sources. Similarly, the traffic monitoring and surveillance actions with VTRF representations are more efficient e.g. due to more early and robust vehicle trajectory based traffic incident detection [8][10-12] and longer available reaction time for anticipative surveillance actions. Finally, the traffic network management and ITS co-ordination actions realized over functional space of vehicle trajectories enables in general more easily to explore and generate beneficial network synergic effects than by conventional events-based traffic models [7-8][10].

2. Traffic systems representations

Traffic phenomena posses both continuous and discrete spatio-temporal nature and have different scales therefore demand for macro and micro representations of traffic flows in natural way arises. The classical traffic control methods that are based on local information are not fitted in general to complex traffic phenomena like congestion, over saturation and traffic incidents [12]. There are proposals of multi-loops network detection systems with the aim of mitigation of crucial discrepancy between spatio-temporal context of controlled traffic processes and representativity of local information feedback offered by inductive loops. In this context we observe fixed-spatial like fixed-time control proposals. The idea of new traffic representation is presented in Fig. 1. The real traffic system is a 2-D dynamical system which interacts with its environment (e.g. traffic network users). The specifications of these interactions i.e. attributes are represented by the system variables (e.g. speeds, accelerations, passivity, energy) with relationships resulting from the system laws. The laws specify the restrictions on the possible manifestations of the phenomenon underlying the system and yield the system behaviour (i.e. the set of all admissible system trajectories). In consequence behaviour of the real traffic system is sensible represented by a family of vehicle trajectories. The first original such type traffic model was presented in [1] and refined in [8-11]. In this model (2-D) Traffic Dynamic System is formally represented as follows:
**TDS** = *(P, A, B)*: where **P** denotes set of parameters (e.g. \( P = T \times S \) with **T**– time, **S**– space); **A**– signal set; **B**– behavior of the system. This system describes a traffic phenomena evolving over two-dimensional domain **P**. These phenomena are specified by attributes which take their values in the set **A**. Thus each manifestation of phenomenon corresponds to an element \( a \in A^P \) (2-D system trajectory). Phenomenon is governed by certain laws, **B** consists of all trajectories compatible with the laws governing phenomenon. These trajectories are estimated in real-time to base on available traffic data sources and grid computing technologies. The image processing technologies will offer for example the following specifications of 2-D discrete systems: \( P = \mathbb{Z}^2; A = \mathbb{R}^q \) with positive integer \( q > 0 \) equal to the number of attributes. The behavior \( B = \{ a : a \in A \text{admissible} \} \) can satisfying certain qualitative properties e.g. to be LSIC i.e. linear (L: B is linear subspace of \( A^P \)), shift-invariant (SI): \( a \in B \Rightarrow \nabla_i \cdot \vec{a} = \vec{a} + k \delta (i - 1), p_z + k \delta (i - 2) \) where \( \delta() \) is Kronecker delta, and complete (C: \( \{ a \in B \} \iff \{ a_{ij} \in B_{ij} : \forall \text{FI} \in \mathbb{Z}^2 \} \) where FI is finite interval (“window”). The main advantages of such proposed traffic models are as follows:

- **2-D traffic system is characterized by compatible family of all 2-D vehicle trajectories.**
- **Lack of a’priori decomposition of system variables and system functions mappings so many possible functional mappings governed by real traffic behavior are covered**
- **Behavior of Traffic Dynamic System is unique (e.g. in the contrast to functional models)**
- **Representation flexibility: orderings (causality, recursivity), local/global properties (internal: controllability, observability; external stability), canonical forms/ initial conditions.**

![Figure 1. New traffic systems representations](image-url)

Formally the curve to be graphical representation of position of vehicle over time and space (constrained to existing roads) plane is called vehicle trajectory. In general several Vehicle Trajectories Random Field (VTRF) traffic representations are possible [11]. In Fig. 2 the vehicle trajectories direct traffic information is presented. The individual/family of trajectories provide fundamental information about individual/stream of vehicular motion characteristics starting from simple travel direction, speeds, accelerations, densities, normal/abnormal traffic situations recognition and on operating vehicle characteristics and behavioral patterns ending. In general the admissible trajectory originally have to be monotonic Lipschitz functions with constrained by admissible vehicle speeds cone slopes.
Let us define the random field $RF$ as a mapping in the form: $X: P \times \Omega \rightarrow S$ where $P$ – is a set of parameters (e.g. $P = T \times S$ with $T$ – time, $S$ – space); $\Omega$ - random events / variables; $(\Omega, F, P)$- Probabilistic Space; $S, P \subseteq \mathbb{R}^n / Z^n (n \geq 1)$. The RF may be also defined as a mapping in the form $X: \Omega \rightarrow F(P, S)$ e.g. $F(P, S)=L_2 (P, S)$ Hilbert space. In this general definition wide spectrum different traffic representations are possible: dynamic $(T \subset P)$, static $(P \cap T = \{t_0\})$, stochastic $(\Omega \neq \mathbb{R})$, deterministic $\Omega = \{\omega_0\}$. We can in general distinguish the following traffic fields. **Deterministic Fields (DF):** Given function $X: P \rightarrow S$ mapping Borel sets is called $(\dim S)$ DF with $(\dim P)$ space of field parameters e.g. $X \in L_2 (P, S)$ Hilbert space if it is square-integrable over $P$. **Stochastic Fields (SF):** Assuming that $(A_1)$: $L_2 (P, S)$ Hilbert space of DF; $(A_2)$: $(\Omega, F, P)$- be a Probabilistic Space: the map $X: \Omega \rightarrow L_2 (P, S)$ is $(\dim P)$ SF if preimage of $X^{-1}(A) \in F$, for $\forall A \subseteq L_2 (P, S)$. As special cases of SF the purely spatial/temporal process (traffic models) are available. Similarly the fields with special properties in $P (P \subset P)$ like homogeneous (partially homogeneous), white (partially white), Poisson, Markov can be used as special traffic representations [11]. The parameterization of $P$ space can be selected in different ways e.g. in accordance with data sources. For “probe vehicles” used to collect real-time travel information (creating the reference trajectories) the VTRF field can be represented locally by a family of lines $\{l_i (p, \Theta_i): i=0, \pm 1, \pm 2,...\}$ where $l(p, \Theta): p \cos(\Theta)-t \sin(\Theta)$; and the angle with $t$-axis $\Theta \in [0, \pi/2]$; $p \in \mathbb{R}$ is the signed length of the perpendicular to the line from origin. (Fig. 2). In this case the robust traffic congestion/incidents detection and robust travel time estimation was presented in [10][12]. In this paper dedicated AR autoregressive system representation for vehicle trajectories estimation and prediction is proposed. The behaviour $B(R) = \ker(\nabla_i^a)$ i.e. is governed by polynomial shift-operator equations of the form: $R(\nabla_i^a) a = 0$ with $R(.)$ polynomial matrix $\in \mathbb{R}^q$ defined by $R(\nabla_i^a) = \sum_m \sum_n \nabla_i^m R_{mn} \nabla_j^n$ i.e. 2-D polynomial matrices with $q$-columns parameterizes the LSIC system.

**Figure 2. Vehicle trajectories-based traffic attributes** (flows, densities, speeds, spacing/platoon dispersion, headways/travel times, stops/queues length/delays)
3. PIACON vehicle trajectories based control modes

For the spatio-temporal data $z(s_j,t)\; j=1,...,n;\; t=1,...,T$ we build model for estimation/prediction of vehicle trajectories at $s_j$ with no constraining assumptions. We use VAR (vector autoregressive) models with spatial structure in the form: $z_t = \alpha + \sum_{l=1,p} R_l z_{t-l} + \epsilon_t$ where $\alpha = \{ \alpha(s_j) \}_{j=1,n}$ is a spatial trend; $R_l (l=1,...,p)$ parameter matrices; $\epsilon_t \sim N(0, \Sigma_e)$ white noise process; $K_z(t) = \text{cov}(z_t, z_{t-\tau})$ covariance matrix that can be computed from $R_l$ and $\Sigma_e$ e.g. for $p=1$ we have $K_z(t) = R_1^T K_z(0)$ where $\text{vec}(K_z(0)) = (I_p - R_1 \otimes R_1)^{-1} \text{vec}(\Sigma_e)$. Spatio-temporal trends can be modeled by VAR: $\nabla z(s,t) = \alpha(s) + \nabla y(s,t)$ where $\alpha(s) = \nabla \frac{1}{2} x(s,t); \nabla y(s,t)$ is time stationary process. We assume that the vehicle trajectories are approximated by polynomial function in $t$ with coefficients to be function of location $s$ i.e. $x(s,t) = \sum_k a_k(s) t^k, k=0,1,...,N$. Renaming $x_{j+(m-1)n} = \{ z(s_j, t-m) \}$ where $j=1,...,n; m=1,2,...$ we use partial correlation function (PCF) for places $s$ defined as: $r_s(h) = \text{corr}(z(s,t), x_h | x_{h-1},...,x_1)$ to selection of appropriate model i.e. identify representative time lags $h_m: r_s(h_m) \neq 0$ and $r_s(h) = 0$ for $h_m < h < m$. To illustrate applicability of these vehicle trajectories models in traffic control area the results of PIACON artery synchronization control modes concerning main Krakow artery (Al. Trzech Wieszczy) are presented in Tables 1-3. In the first “platoon trajectories” robust control mode the green band generated by PIACON enables movement of real-time detected and estimated platoon of vehicles along artery in the wide range of operational conditions (uncertainty, hybrid operation of controllers, directional anisotropy). In addition special traffic events option can offer so called “express service” e.g. in the case of a priori known ambulance trajectory (Table 1). In the second “reference trajectory” robust control mode the green band is generated around this trajectory (Table 2). In the third “multi-criteria” mode the artery synchronization is optimized according to many criteria estimated in real-time from estimated trajectories and at successive intersections (Table 3). As can be seen all proposed PIACON original control modes are very beneficial and robust.

4. Summary and conclusions

The efficient solutions of traffic problems are stimulated by potential of new technologies which presently enables on essential increase of possibilities and abilities (traffic responsiveness and adaptability, robustness, fault-tolerant, congestion mitigation) of traffic surveillance, control and management actions. The transportation systems are a very complex plants i.e. large-scale anisotropic, geographically distributed complicated systems with complex traffic phenomena (randomness, uncertainty, very complex interactions, diversified non-linear dynamics, human behavioral anisotropy, structural instabilities). This is why the adequate representation of traffic phenomena should be a prerequisite of all planed beneficial actions in transportation systems. This aspect is a main critical point in most existing approaches. In this context the proposal of the VTRF traffic representation with real-time robust estimation of vehicle trajectories offers essential progress on the way to development of adequate traffic phenomena models.
References


Table 1: PIACON: artery robust control with “platoon trajectories” mode
Table 2. PIACON artery synchronization “reference trajectory “ robust control mode
Table 3: PIACON: Multi-criteria artery synchronization
MODELING ACTIVITY CHOICE DISTRIBUTION AND ROAD CHOICE BEHAVIOUR ON A NETWORK WITH SIDE CONSTRAINTS

Tai-Yu MA¹, Jean-Patrick LEBACQUE²

Abstract. Modeling travel demand based on activity distribution is an important issue in transport demand analysis. In this work, an activity-based destination and road choice model is proposed as an optimization problem. A multi-agent like ant colony algorithm is proposed to reflect travelers’ adaptive behavior to activity distribution and traffic condition dynamics.

1. Introduction

Modeling traffic assignment based on activity distribution is an important issue in many activity-based models [1], [3]. The dynamics of activity/destination choice are usually modeled within the utility maximization framework. But the underlying assumption of perfect human cognition has been subject to criticism [5]. Recently, interest into multi-agent modeling has been increasing in the transportation research domain [11] as an alternative or a complement to the utility approach. Related work in this direction has been proposed in [8], [12].

In the present paper, we propose a link capacitated traffic assignment model based on the concept of concurrence of activity and route choice. We formulate the traffic assignment, and activity location choice problem as an equilibrium problem. Various optimization solution techniques have been proposed in [6], [7], [9]. In this work, a multi-agent type ant colony algorithm is proposed, because it provides an optimization rationale and at the same time has the potential for including complex activity behavior.

The ant colony based optimization algorithm (ACO) [4], [10] is a general meta-heuristics for combinatorial optimization problems. It consists of many ant-type agents who

¹ INRETS/GRETIA - 2, Avenue du Général-Malleret Joinville F-94114 Arcueil France. tai-yu.ma@inrets.fr
² INRETS/GRETIA - 2, Avenue du Général-Malleret Joinville F-94114 Arcueil France. lebacque@inrets.fr
cooperate by depositing pheromone trails to solve optimization problems. By means of an indirect communication mechanism i.e. pheromone trails, ants are capable to find shortest paths from their nest to a food resource. The focus of this work is on how to build heuristic guides for ant-type agents to find a good quality solution in a capacitated constrained network. The pheromone scheme itself should be extended towards more general agent-to-agent or system-to-agent communication schemes, emulating information diffusion or communication schemes. The ant-like agent model can also be developed into a simulation of the transportation system.

2. The model

2.1. Description

In this work, we propose a model based on activity distribution to optimize travelers’ road and destination choice in a network with link capacity constraints. This idea extends the earlier work [12], which was based on static assignment with a probabilistic opportunity choice model of activities. Contrarily to common static assignment models, the traffic distribution of any origin with respect to destinations is assumed to be unknown. The network is represented by a directed graph in which links are constrained by their capacity (traffic flow via each link can’t exceed its capacity). In this first version of the model, we assume that travelers engage in a single activity and we neglect the activity chain problem. Travelers’ behavior is modeled by analogy with the search behavior of artificial ants. They depart from an origin node to a destination node to engage some activity, the reward being the activity’s economic value, minus the travel cost (the difference being the net economic value). At each potential destination the activities are distributed according to a given distribution of economic value. Although this model is not completely realistic, at this point it serves as a benchmark for testing network modeling methods and it will be used as an archetype for more realistic models based on activity programs.

For the probability distribution of gross activity value \( v \) at each destination \( d \), we assume that it is an exponential distribution function \( P_v(v) = \lambda_d \exp(-\lambda_d(v - m_j)) \) where \( v \geq m_j \) and \( \lambda_d \geq 0 \). The related cumulative distribution function \( H_d(v) \) can be deduced as \( H_d(v) = 1 - \exp(-\lambda_d(v - m_j)) \) if \( v \geq m_j \), otherwise 0. Therefore, the function of average gross value of serviced activities at destination \( d \) is defined as follows:

\[
S_j(G_j) = \int_0^v H_d^{-1}(1 - \alpha) d\alpha = G_d \left[ \frac{1 - \ln G_d}{\lambda_j} + m_j \right]
\]

where \( G_d \) is the proportion of serviced activities in destination \( d \). The total gross serviced activity in destination \( d \) can be obtained by \( N_d S_j(G_j) \). In this respect, we formulate our activity-based link capacitated traffic assignment problem as follows (Beckmann type transformation [2]):
\[ \text{Min} \sum_{e} f_e(s, C_e) \, ds - \sum_{d} N_d S_d \left( \frac{\sum_{o} d_{od}}{N_d} \right) \]

St.

\[ d_k - \sum_{e \in M(o(k))} x_{ek} + \sum_{e \in N(o(k))} x_{oe} = 0 \quad \forall k \in K \]

\[ d_k - \sum_{e \in M(d(k))} x_{ek} + \sum_{e \in N(d(k))} x_{ed} = 0 \quad \forall k \in K \]

\[ \sum_{e \in M(i)} x_{ek} - \sum_{e \in N(i)} x_{ek} = 0 \quad \forall i \notin O, \forall i \notin D, \forall k \in K \]

\[ \sum_{d} d_{od} = d_o \quad \forall o \in O \]

\[ \sum_{o} d_{od} \leq N_d \quad \forall d \in D \]

\[ \sum_{k} x_{ek} \leq C_e \quad \forall e \in E \]

\[ x_{ek} \geq 0 \quad \forall e \in E, \forall k \in K \]

where

- \( E \) set of links
- \( O \) set of origins
- \( D \) set of destinations
- \( K \) set of origin and destination pairs \((o,d)\)
- \( o(k) \) origin of \( o-d \) pair \( k \)
- \( d(k) \) destination of \( o-d \) pair \( k \)
- \( M(i) \) set of links whose head node is \( i \)
- \( N(i) \) set of links whose tail node is \( i \)
- \( C_e \) capacity of link \( e \) (data)
- \( x_{ek} \) total flow on link \( e \) whose \( o-d \) pair is \( k \) (unknown)
- \( d_o \) total demand for origin \( o \) (data)
- \( d_k \) total demand for \( o-d \) pair \( k \) (unknown)
- \( N_d \) number of activities at destination \( d \) (data)
- \( S_d(\cdot) \) function of average gross value of serviced activities in destination \( d \) (data)
- \( f_e(x_e, C_e) \) travel time function of link \( e \) determined by traffic flow via link \( e \) and its capacity \( C_e \) (data)
The link cost function is a simple linear function of the link flow. The link flow is bounded by a constant link capacity (Equation (8)). Equation (3), (4), and (5) describe the conservation of flow. Equation (6) is total travel demand constraint at origins. Equation (7) is the bound on activity supply at destinations.

Concerning the search of the solution, each agent (ant) chooses the activity with best net value (economic value minus path cost): the activity choice competition at destinations. The resulting equilibrium is a user equilibrium, structurally close to an assignment problem. This destination choice is modeled with the introduction of virtual nodes and links. The virtual links represent the distribution of activities at destinations. The flow on a virtual link measures the quantity of occupied activities at the corresponding destination. The number of available activities at the destination serves as the capacity of the corresponding virtual link. It is taken into account by introducing the concept of penalty value on the virtual links. A formula similar to (12) applies.

Further, in contrast with the utility maximization framework, in which the travelers’ cognition capacity is perfect, we assume that travelers have no global information of the system. Travelers use local information concerning path quality experienced by other travelers and local link travel cost to choose their search direction in the network. After an agent has engaged an activity, the pheromone trail of his routing path is updated corresponding to the quality of the routing path. The quality of this path is measured by the net value of the activity chosen at the end of the path.

The travelers’ destination distribution and routing path choice are influenced by activity value distributions and by the dynamic traffic conditions determined by the artificial ants’ behavior. In the dynamic version of the model the constraints of supply and demand and the constraints of flow conservation are imposed at nodes.

2.2. The ACO approach

The general idea of the ACO algorithm is that every ant constructs a partial solution by searching a vacant activity in the network. Ants mark their paths by depositing some quantity of origin-based pheromone on the edges of paths. Subsequent ants move to better paths by the attraction of previously found best solutions. Communication layers between ants from different origins are independent. Besides, two important effects that have to be considered are: how capacity (resources) constraints are satisfied, and how ants choose activities located at different destinations.

Capacity constraints are treated by adding penalties to the travel cost function. The distribution of values of activities at destinations is given, as mentioned above.

By this algorithm, an approximate solution is obtained and we have compared it with the exact result obtained by traditional mathematical programming methods (Frank-Wolfe-like [6] and also [7]).

The general framework of this ACO algorithm is composed of three steps:
1. initialize all links with small amounts of pheromone;
2. the probability for ant \( m \) at node \( r \) to choose next link \((r, s)\) is given by the stochastic decision rule (after [4]):

\[
P^m_r(t) = \begin{cases} \left[ \frac{\tau^m_{rs}(t)}{\sum_{s \in J^m_r(t)} \tau^m_{rs}(t)} \right]^\beta & \text{if } s \in J^m_r(t) \\ 0 & \text{otherwise} \end{cases}
\]  \( (10) \)

where parameters \( \alpha, \beta \) control the relative importance of normalized pheromone value \( \tau^m_{rs}(t) \) and normalized heuristics information \( \tilde{\eta}_{rs}(t) \), defined by normalized inverse link cost at link \((r, s)\) at iteration \( t \). \( J^m_r(t) \) is set of non visited outgoing links from node \( r \) for ant \( m \).

To reflect congestion effect, link travel cost is calculated with a penalty term that activates when traffic flow approaches its link capacity.

\[
f_m(x^m_e(t), C_e) = a + \frac{b}{C_e} x^m_e(t) + \Gamma_e(t)
\]  \( (11) \)

where \( a \) and \( b \) are constant, \( C_e \) is link capacity. \( x^m_e(t) \) is actual traffic flow on link \( e \) just after ant \( m \) enters link \( e \) for iteration \( t \). \( \Gamma_e(t) \) is penalty function, defined by

\[
\Gamma_e(t) = \begin{cases} \psi \left( \frac{1}{C_e - x^m_e(t)} - \frac{1}{\mu C_e} \right) & \text{if } C_e(1 - \mu) \leq x^m_e(t) \leq C_e \\ 0 & \text{otherwise} \end{cases}
\]  \( (12) \)

where \( \mu \) is a parameter controlling activation of penalty function. \( \psi \) is a parameter defined by \( \psi = C_e b \zeta \), in order to eliminate influence of link capacity. \( \zeta \) is a parameter.

3. After an ant accomplishes its path to some destination, the quantity of pheromone is immediately updated by the local update rule:

\[
\Delta \tau^m_{rs}(t) = \begin{cases} Q / L_m(t) & \text{if } (r, s) \in P^m_m(t) \\ 0 & \text{otherwise} \end{cases}
\]  \( (13) \)

where \( P^m_m(t) \) is the path used by ant \( m \) at iteration \( t \). \( Q \) is a constant and \( L_m(t) \) is solution quality function, defined by ant’s path travel cost minus the obtained activity value and adding a constant defined by globally maximum activity value such that \( L_m(t) > 0 \). When all ants have constructed their paths, the pheromone trail is updated by the normalized global update rule:

\[
\tau^m_{rs}(t) = (1 - \rho) \tau^m_{rs}(t - 1) + \rho \omega \sum_{s \in J^m_r(t)} \sum_{\theta \in \Omega} \frac{\Delta \tau^m_{rs}(t)}{\sum_{s \in J^m_r(t)} \sum_{\theta \in \Omega} \Delta \tau^m_{rs}(t)}
\]  \( (14) \)

where \( \rho \in (0, 1) \) is a constant evaporation rate, \( \omega \) is a parameter to adjust the amount of added pheromone, \( o \) is the origin and

\[
\Delta \tau^m_{rs}(t) = \sum_{\theta \in M(o)} \Delta \tau^m_{rs}(t)
\]  \( (15) \)

represents the total pheromone value at link \((r, s)\) for type \( o \) pheromone at iteration \( t \), and \( M(o) \) is the set of ants departing from origin \( o \).
This global update rule is based on the minimum cross-entropy pheromone update concept [10] to avoid the problem of scale in updating the pheromone amount when the number of ants and the path quality function change.

A simplified capacitated test network is represented by nodes and directed links (see figure 1). Origins with ants are nodes 1 and 2, and destinations with vacant activities are nodes 3 and 4. We test different settings of parameters and number of ants. In figure 2, pheromone amount evolution on each link for ants from node 1 is illustrated. It shows that after a certain number of iterations the profile of quantity of pheromone in links stabilizes and we obtain a near optimal solution. The performance of the algorithm depends on adequate setting of the parameters in computing the transition probability (see figure 3).

3. Conclusion

This model performs traffic assignment, destination and activity choice based on activity distributions. The travelers’ behavior is modeled as artificial ants’ food searching process. By an indirect communication mechanism between ants of same origin, a set of paths emerges yielding close to optimal routing and destination choice. The algorithm can be also extended to the dynamic traffic assignment problem with activity destination and departure time choice. Future extensions of this work concern taking into account travelers’ activity programming, dynamic adaptative behavior to traffic congestion and different supply side operations and strategies.

Figure 1. A simplified road network
Figure 2. Evolution of quantity of pheromone on links

Figure 3. Comparison of objective value with different settings of parameters
Acknowledgements

The authors gratefully acknowledge the support of the ACI (Action concertée incitative) NIM (Nouvelles interfaces des mathématiques) “Modélisation mathématique du trafic automobile” of the French Ministry for Higher Education and Research.

References


Abstract. This paper examines the possibility of applying fractal measures in the modelling of traffic networks, and presents an investigation of the effect of the representation scale of the road hierarchy on the network equilibrium. Some complexity aspects of the network topology, and their effect on the user equilibrium, are illustrated using a numeric example. Conceptual and methodological issues that could be addressed by further research in transport planning are suggested.

1. Introduction

Transportation is a wide human-oriented field with diverse and challenging problems waiting to be solved. Transport systems are comprised of many elements – services, prices, infrastructures, vehicles, control systems and users, which are been taken as a whole. Characteristics and performances of transport systems and their components are usually defined on the basis of quantitative evaluation of their main effects, considering the objectives and the constraints of the transport system. Such systems are complex by nature. The planning and operation of transport systems may vary much by its focus and methodologies. A detailed level of representation may be used in the functional design and control management of transport systems. In the appraisal a new transport scheme, transport planners and policy makers may use a larger scale when representing the transport system.

Segments of the transport system (such as road links) can be classified into hierarchies according to the size and type of traffic flow they are supposed to carry. Public roads may be classified by types: motorways (freeways), trunk roads, principle roads and local B and C type roads. The design of such roads is based on their purposes and their roles in accessibility and mobility. This leads to specific recommendations on the level of service provided by different types of road, and guidelines on the geometric design of such roads (such as the number of lanes and the lanes width, and the design of other elements such as intersections, parking, traffic calming). Table 1 represents the length of public roads in

---

1 Centre for Transport and Society, University of the West of England, Coldharbour Lane, Bristol, United Kingdom BS16 1QY, erel.avineri@uwe.ac.uk
Great Britain (based on 2001 national statistics). Although motorways account for only 1% of the total length of the transport network, they carry about a quarter of the traffic volume.

<table>
<thead>
<tr>
<th>Road Type</th>
<th>Length (km)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Motorways</td>
<td>3,465</td>
</tr>
<tr>
<td>Trunk - non motorway</td>
<td>11,742</td>
</tr>
<tr>
<td>Principal - non motorway</td>
<td>34,822</td>
</tr>
<tr>
<td>B roads</td>
<td>30,057</td>
</tr>
<tr>
<td>C roads</td>
<td>83,272</td>
</tr>
<tr>
<td>Unclassified roads</td>
<td>228,349</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td><strong>391,707</strong></td>
</tr>
</tbody>
</table>

Table 1: Public Roads in Great Britain - Length (kms) by Class
(based on 2001 national statistics).

One of the important aspects of the transport planning process is assigning the predicted traffic volumes on the road segments. This is followed by an analysis of the transport network and its performance, mainly by observing its ability to satisfy the travel demand of persons and goods in a given area.

Many applications of traffic networks models do not carry a full representation of the transport network. Representing the whole network may not be effective or in some cases not feasible, for the following reasons:

(i) transport modellers do not have access to detailed information about the minor roads, and are not able to calibrate the travel time functions on these segments;
(ii) transport modellers and planers may not have the required resources to represent the whole transport network; budget, time and human resources are largely constrained;
(iii) when modelling a planned transport system, some of its elements may not be fully planned or designed. For example, the planning of some minor roads in residential areas may be based on a conceptual level only;
(iv) the modelling process may be mainly focused on the effects on the major roads.

Due to the above and other reasons an obvious conflict appears between the simplicity of the abstract and partial representation of the traffic system, and the accuracy and the completeness of the whole system. Faced with this conflict, many transport modellers choose a compromised solution, representing the network by its major roads, and ignoring (or representing in a more abstract way) the minor roads and links of the system. Many transport modellers may believe that this under-representation of the transport network may not have a major effect on the traffic assignment. This follows an implicit assumption that omitting minor roads will not make much difference on the assignment of traffic volume on the traffic roads, and the performance of the transport system. The main problem with incomplete networks arises in heavily congested areas where some of the medium and long-distance trips use minor roads as ‘rat runs’. This is considered to be one of the major limitations of classic models of the road network (see p. 365 in [5]).
The transport network can be described a complex system. It has been observed that transport networks exhibit properties of regularity and self-similarity [8; 1]. Moreover, some researchers take it a step further and argue that the complex characteristics of urban traffic networks can be explained by a fractal growth of cities (or vice versa) [7; 3; 4; 10]. These concepts are reviewed and discussed in the full paper.

2. Numeric Example

Assume that the link volume-travel time functions are linear, and given by:

\[ t_{ij} = \alpha_{ij} + \beta_{ij}f_{ij} \]  

(1)

where

- \( t_{ij} \) is the travel time on link \( ij \)
- \( \alpha_{ij} \) is the free flow time on link \( ij \)
- \( \beta_{ij} \) is the delay parameter for link \( ij \) (the increase in travel time per unit increase in the flow on link \( ij \)), and
- \( f_{ij} \) is the flow on link \( ij \).

(above notation is taken from [6]).

Let us consider a simple network problem, presented in Figure 1. There are three possible paths to get from the origin to the destination, and the total traffic, \( Q \), is equal to the sum of traffic volumes on these paths:

\[ Q = F_1 + F_2 + F_3 \]

(2)

where:

\[ f_{ab} = F_1 + F_3; f_{bd} = F_1; f_{ac} = F_2; f_{cd} = F_2 + F_3; f_{bc} = F_3 \]

(3)

and:

\[ T_1 = T_{ab} + T_{bd}; T_2 = T_{ac} + T_{cd}; T_3 = T_{av} + T_{bc} + T_{cd} \]

(4)

The parameters were used in this numeric example are presented in Table 2.

Following Wardrop’s principles [9], in equilibrium, no user can decrease his/her route travel time by unilaterally switching routes. This condition may be represented by

\[ T_1 = T_2 = T_3 \]

(5)
Figure 1. The Traffic Network and Three Possible Paths.

Table 2: Parameter Values of the Road Segments.

<table>
<thead>
<tr>
<th>link (ij)</th>
<th>$a_{ij}$</th>
<th>$b_{ij}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>ab</td>
<td>0</td>
<td>0.015</td>
</tr>
<tr>
<td>bd</td>
<td>25</td>
<td>0</td>
</tr>
<tr>
<td>ac</td>
<td>25</td>
<td>0</td>
</tr>
<tr>
<td>cd</td>
<td>0</td>
<td>0.015</td>
</tr>
<tr>
<td>bc</td>
<td>8</td>
<td>0.015</td>
</tr>
</tbody>
</table>

The user equilibrium solution for this 3-path network is $T_1 = T_2 = T_3 = 38$ minutes, while the user equilibrium solution for a 2-path network (by omitting the link $bc$) is $T_1 = T_2 = 32.5$ minutes. This illustrates the Braess’ Paradox [2]: adding new capacity (such as an extra link) in a congested network does not necessarily reduce congestion and can even increase it. This situation happens because the users of the network users do not face the true social cost of an action, reflected by a ‘User Equilibrium’ rather than a ‘System Optimum’.

The above example illustrates one of the important aspects of choosing a scale in network modeling: ignoring the minor road segment, represented here as link $bc$ will cause a completely different traffic assignment, followed by a different (in this case – worse) travel times for all users, and a different distribution of negative impacts of transport; for example, residents of the areas located next to the link $bc$ may suffer from high levels traffic congestion, air pollution and noise; this distribution of transport impacts may not be captured if the network is simplified by omitting the third link $bc$.

In the full paper the following concept is illustrated: by reducing the scale the representing the traffic network in a more detailed, we may found out that the network performance is described by different traffic equilibria, and the traffic assignment on the minor links might be much changed. This concept can be demonstrated by a hypothetical network, generated by a recursive iteration of the (nested) Braess’ Paradox, as presented in...
Figure 2. This network can be described as a fractal system; it is characterised by self-similarity, symmetry across scales and an emphasis on the position of the observer and scale of measurements. A detailed representation of the effect of the fractal geometry of the road network on the assigned traffic volumes is given in the full paper.

This paper examines the possibility of applying fractal measures in the modelling of network equilibrium, and presents an investigation of the effect of the fractal dimension of the road hierarchy on such equilibrium. Conceptual and methodological issues that could be addressed by further research in transport planning are suggested.

3. Conclusion

This paper describes conceptual and methodological issues for presenting and exploring complex traffic systems and illustrates the application of complexity techniques in the study of traffic equilibrium.

The paper was motivated by the hypothesis that transport systems are complex by nature. Although it is not necessarily considered to be a major concern by practitioners, strong empirical evidence, reviewed in this paper, highlights its importance in transport modelling.

Facing rather large and complex traffic networks, with this conflict, many transport modellers may compromise on representing the network by its major roads, and omitting the minor roads and links of the system. They should be aware of the main limitations of this approach:

1. small changes in the network (such as omitting minor links) may lead to major changes in the traffic equilibrium, and therefore to the users of the traffic system;
the distribution of the negative effects of the transport is sensitive to the level in which the network is represented. While transport planners cannot completely avoid the above situation, they should be aware of the following:

(1) due to the simplification of the transport network, the predictive value of the traffic flow on the road may be wrong (even on roads which hierarchical level is fully represented, such as roads that are classified as motorway);

(2) further empirical and theoretical investigation of the fractal patterns urban road networks may help in the conceptual representation of the minor parts of the road network.

References


STOCHASTIC DELAY AT TRAFFIC SIGNALS: INFLUENCE OF DEMAND PATTERNS

Estelle CHEVALLIER1, Ludovic LECLERCQ1

1. Motivations

It is well-known that most traffic delay is due to signalized intersections in urban networks. An accurate description of delay at traffic signals is therefore needed to refine assignment models and travel time estimates. Delays at traffic signals are composed of two parts: (i) a deterministic service time representing the time a vehicle has to wait before crossing the junction (which depends on both signal timing and the ratio of arriving flow over saturation rate) and (ii) a random queuing component due to non zero initial queue length when vehicles cannot discharge from previous cycles. Well-known steady state formulae (e.g. the Webster’s one [11]) or time dependent equations [3] have been proposed to deal with these two phenomena for specific arrival distributions.

In macroscopic travel time estimation models linked with dynamic assignment models, one of the stakes is to accurately reproduce both the uniform delay for the average arriving flow over the cycle and the average effects of stochastic fluctuations. The deterministic part of the delay is correctly reproduced by explicitly modelling the capacity restriction due to the traffic signal for an aggregated average demand. However, the average effects of stochastic variations should be added on an ad-hoc basis by artificially delaying vehicles. Indeed, at a macroscopic scale with long simulation steps, inter-arrival times are assumed constant over a cycle and overflow queues cannot occur if demand is constant and below capacity. An explicit formulation of the average penalty due to random arrivals is, therefore, very appealing in the area of traffic simulation models.

In this article we show that delays at signalized intersections are mainly influenced by traffic demand patterns defined as the time evolution of vehicle arrivals. We will demonstrate that a major part of the stochastic delay can be directly incorporated into the simulation model by a finer representation of demand, allowing for variations within one or a few signal cycle(s). To this end, the first section of this paper focuses on the method used to study the impact of the demand pattern into the delay at a fixed signal timing intersection. Then, we highlight the main results before discussing the outcomes of our findings.

1 LICIT (ENTPE/INRETS), Rue Maurice Audin 69518 Vaulx-en-Velin Cedex, France, estelle.chevallier@entpe.fr; leclercq@entpe.fr
2. Methodology

Existing steady state delay formulae share the same basic assumptions for representing the demand function: (i) a constant demand rate (reference demand) for the deterministic component of the delay and (ii) random oscillations around this demand rate according to the arrival distribution for the stochastic term.

However, in urban networks, there is little evidence of a constant mean arrival rate at a signalized intersection over a signal cycle. Firstly, vehicles are bunched as a result of upstream control, and, secondly, the average number of arriving vehicles during green or red can vary. To reproduce these phenomena, we have modelled the demand pattern as a step function with two uniform rates. This seems to be an accurate description of the platooning and the dynamic effect of adjacent control.

Such a demand pattern can be described by three parameters:

a) The demand cycle which is the time period for exactly one replication of the step demand function. It is close to the signal cycle if there is a signalized junction upstream on the network. On the contrary it increases when the upstream control is an unsignalized intersection.

b) The intensity which is the magnitude between the two levels of the demand pattern. It depends on the assignment coefficients for major and minor streams at the previous junction and weakens with the distance travelled. Hence, high intensity demand patterns are characteristic of a major arterial with dominant straight movements whatever the entry of the junction, while low intensity demand patterns stand for a junction where turning rates for minor streams are important.

c) The offset which is the initial lag between the beginning of the red period and the highest level of demand. It depends on the signal coordination between consecutive signals. It can also follow a uniform distribution if the step demand function is the result of a local increase of traffic flow upstream the junction.

The impacts of these three parameters (represented in figure 1) on the average delay per vehicle are studied in the remaining of the paper.

![Figure 1: Parameters of a step demand pattern](image-url)
The cumulative count curves method was used to compute the average delay per vehicle over each cycle. The left part of figure 2 shows the cumulative number of vehicles arriving at the intersection and served by the traffic light when the offset is equal to zero. The area between the two curves, divided by the total number of arriving vehicles over the simulation period, gives the average delay per vehicle. It represents the first point in the right part of the figure. Each other dot corresponds to a replication of the method with a specific offset. The offset reflects the impact of adjacent intersections in terms of signal coordination quality (if the demand cycle is approximately equal to the signal timing) or the consequence of a local peak period (for longer demand cycles). By varying the offset between zero and the demand cycle, both favourable and unfavourable situations are encompassed. The bold solid line shows the average delay obtained over all offsets (simply referred as the average delay in the sequel). It should be compared with the deterministic delay calculated for a constant demand (noted as the reference delay) represented by the dashed line. The average delay is the delay we considered in the sensitivity analysis we carried out in section 3.

![Cumulative curves of the number of arriving and served vehicles (left), delays obtained by varying the offset (right)](image)

### Figure 2: Cumulative curves of the number of arriving and served vehicles (left), delays obtained by varying the offset (right)

### 3. Results

In this section, the signal cycle is set equal to 100s, starting with a red period of 50s.

#### 3.1. Offset effect

The offset is probably the main influencing factor of delay. When the whole red period is subject to the upper level of the step demand pattern, the proportion of blocking vehicles increases compared to cases with lower arrival rates. Vehicles cannot discharge during a signal cycle and queue length grows up. Such a scenario may represent, for instance, poorly coordinated signals. On the contrary, the system performance can be optimized by limiting
the number of arrivals during red and maximising flow during green (well coordinated signals).

In figure 2, we can notice that the average delay is about 1.6 times higher than the reference delay (41s instead of 25s). In other words, unfavourable cases seem to be more disadvantageous in term of delay than the benefits of signal coordination. This is a convincing claim for optimizing signal cycles by accounting for adjacent control.

Moreover, even for the well-coordinated case, the average delay with the step demand pattern is still higher than the reference delay. Incorporating the dynamic effect of traffic by allowing for flow variations within the signal cycle triggers additional delays compared to the deterministic term of the steady state formula (reference delay). Thus, if we want to model a traffic scenario with consecutive well-coordinated intersections, we can give higher weights to favourable offsets (instead of using a uniform distribution) and still have higher delays than for the reference demand.

Multiple simulations for other intensity and demand cycle values have been carried out and it appears that the delay computed over all offsets is always higher than the reference delay. Hence, accounting for the offset effect with a step demand pattern is a powerful means to incorporate the main part of the stochastic delay due to random arrivals.

3.2. Frequency effect

For this study we chose a step demand pattern with maximum intensity (the first level of demand is set equal to 0.5 veh/s while the second level is zero). However, similar conclusions could be drawn for other traffic conditions. A relevant parameter for this section is the frequency of the demand pattern defined as the ratio of the demand cycle over the signal cycle. Note that, the reference delay does not vary with frequency since the reference demand is equal to 0.25 veh/s in each case. This demand corresponds to the capacity of the signalized intersection.

By increasing the frequency, the average delay rises up and the discrepancy with the reference delay is magnified (see figure 3). Indeed, for longer demand cycles, the number of signal cycles for which a higher-than-capacity rate applies is greater. Therefore, the length of the queue increases as well as the queue clearance time.

![Figure 3: Frequency effect on the average delay](image-url)
3.3. Intensity effect

We compared delays for a high intensity demand pattern (defined in the previous section) and a low intensity one (with the upper level equal to 0.3 veh/s and the lower level equal to 0.2 veh/s) for different frequencies. As shown in figure 4, the high intensity demand pattern always gives higher average delays than the smoother one (which gets close to a constant arrival rate). The discrepancy is all the more sticking as the frequency increases. One should draw the attention on the specific case when the demand cycle is equal to the signal cycle. For all values of the intensity, the unfavourable cases for certain offset values are exactly balanced by the benefits of good coordination. In this case, the delay for a dynamic demand pattern is equal to the reference delay (uniform arrival rate).

![Figure 4: Intensity effect on the average delay for different frequencies](image)

3.4. Evolution of the simulated delay with the saturation degree

The demand pattern is calibrated so as to represent the effect of an upstream signalized intersection. The demand cycle $D_c$ is set equal to 125s. We note $Q_0$ the reference demand and $Q_1$ and $Q_2$ the levels of the demand pattern.

$$\frac{D_c}{2} Q_1 + \frac{D_c}{2} Q_2 = D_c Q_0$$

(1)

The two levels of the step demand function ($Q_1$ and $Q_2$) correspond to 80% of straight movements for the main flow and 20% of turning movements for the minor flow for the same arrival flow ($q$). We assume that the upstream intersection has the same signal timing as the studied junction (the signal cycle is $C$). By noting $Q_m$ the saturation flow and $G_a$ the queue clearance time (which is the same for stream 1 and stream 2):

$$Q_1 = 0.8 \left[ G_m Q_m + \left( \frac{C}{2} - G_a \right) q \right]$$

(2)

$$Q_2 = 0.2 \left[ G_m Q_m + \left( \frac{C}{2} - G_a \right) q \right]$$

(3)

Moreover:
\[ G_a Q_a + \left( \frac{C}{2} - G_a \right) q = Cq \]  

(4)

From (4) we can derived an expression for \( G_a \) in terms of \( C \), \( Q_a \), and \( q \). Substituting \( G_a \) in expressions (2) and (3) and solving equation (1) gives a value for \( q \) (and so for \( Q_1 \) and \( Q_2 \)).

As usual, the saturation degree is defined as the ratio of the reference demand over the capacity of the signalized intersection.

One can see in figure 6 that the average delay computed for a demand cycle of 125s (frequency equal to 5/4) is close to the Webster’s formula for low saturation degrees (up to 0.6), when the likelihood of overflow queue from previous cycles is low. It should be noted that in this range of saturation degrees, consistency of all delay models have been shown [5], and hence, this result is not surprising.

The gap rises up when the saturation degree increases due to the predominance of the random delay component in the Webster’s formula. This term represents the time spent in an overflow queue for a M/D/1 system [9]. However, assuming a Poisson arrival for its derivation does not reflect real traffic patterns since headways (respectively flows) between vehicles can be unrealistically small (respectively high) as we can see in figure 5 for classic value of reference demand. A main consequence of this assumption is that the steady state delay tends to infinity as traffic demand approaches saturation. This hypothesis was considered as a weakness by many researchers [2] since it does not fit real traffic observations.

Time dependent models, although without any theoretical foundation [3], better describe how delay estimation models should behave under varying demand levels. They are obtained thanks to the coordinate transformation technique, which transforms the equation defining a steady-state stochastic delay model so that it becomes asymptotic to the deterministic over-saturation model for saturation degrees near 1 and higher. The HCM 2000 equation is an example of such models. It is derived from an approximation of the Miller’s steady-state equation [8] which, like in Webster’s formula, assumes Poisson arrivals for the overflow term [1]. Even if the delay curve does not tend towards infinity when we approach capacity, the exponential hypothesis for describing inter-arrival times, again, leads to an overestimation of delay. As a consequence, the gap observed in figure 6 between our simulated delay and the HCM equation cannot be viewed as a limitation of our method.

To overrun the drawback of the Poisson arrivals assumption, a shifted exponential distribution can be used to represent the vehicle arrivals (noted shifted demand pattern in the remaining). The M3 model proposed by Cowan [4] has been used frequently [7] [10] as it takes into account both a minimum headway and the platooning effect. However, the shifted exponential distribution seems to be a good compromise to avoid the over-estimation of the number of very short headways and the calibration of the proportion of bunched vehicles used in the Cowan’s distribution. With this distribution, one can see in figure 6 that the delay does not tend towards infinity as traffic flow approaches the capacity rate. Moreover, it fairly well approximates the average delay obtained with the step demand pattern for a demand cycle equal to 125s. Note that if we have used a longer demand cycle,
the average delay with the step demand pattern would have been closer to the HCM formula.

Figure 5: Irrelevance of the Poisson arrival distribution

Figure 6: Comparison of different average delay functions depending on the saturation degree
4. Conclusive remarks and discussion

In steady state models, the system is assumed to have been running long enough to allow it to have reached an equilibrium state. Therefore, these kinds of models predict delays over long periods on which demand is aggregated around a mean constant arriving rate and where the average effect of random arrivals are incorporated thanks to the stochastic component.

In order to improve the accuracy of assignment models, several authors currently focus on dynamic travel time estimation models able to account for the average effect of randomness in the number of arrivals [6]. At first sight, correcting delays on an ad-hoc basis by using one of the delay estimation functions can be appealing. We can imagine, for instance, incorporating the stochastic term of Webster’s like formula in the simulation by artificially stocking vehicles on the network. However, using the results of a steady state formula in a dynamic assignment model where arrival rate can vary is a tedious work. Besides the arguable assumptions concerning the arrivals distribution used to derive steady state equations, the exogenous integration of an additional term to account for random processes poses the risk of a double counting of delays. Indeed, in steady state models, the effects of demand variation are entirely encompassed into the second term through random oscillations around the mean arrival rate, whereas in dynamic traffic simulations, the explicit traffic light modelling already gives delays which encompassed a dynamic demand variation. To summarize, the scaling effect of the demand pattern should be carefully taken into account. If the demand flow is aggregated over a long time period, one should incorporate the stochastic delay component exogenously. Otherwise, if demand variations during the signal timing are already described by the simulation model, there is no need to correct the simulated delay.

To clear up this ambiguity, our study tried to demonstrate that it is possible to obtain endogenous accurate delay estimates (only by modelling the signal cycle) by refining the demand function. In order to account for traffic dynamic, we used a step demand pattern varying over one to six signal cycles, which is relevant in an urban context where adjacent controls are likely to alter the vehicle arrivals. We showed that the higher the intensity and the cycle demand are, the higher the simulated delay is. This can be explained by both the dynamic pattern of demand (which can vary within a cycle) and the offset effect which can trigger unfavourable cases with large queue lengths unbalanced by more favourable cases.

Interestingly, stochastic effects of arrivals, encompassed into Webster’s like formula, can be reproduced by using a deterministic uniform distribution of the offset. More precisely, the average delay obtained with a step demand function by varying the offset looks like the delay for vehicles arriving with a shifted exponential distribution. As a result, the stochastic arrival processes leading to local overflow queues when traffic intensity is below capacity can be put away if we use a finer demand description in the order of the signal cycle. Therefore, in a traffic simulation model, it is possible to replace the random draw of inter-arrival times by a loop over a deterministic uniform distribution of the offset to forecast the average vehicle delay. Hence, there is no need of several simulation runs with different random number seeds to impede for stochastic processes.
References


CAR POOLING CLUBS: VIABILITY OF THE CONCEPT APPLIED TO LISBON’S METROPOLITAN AREA, USING GIS AND OPTIMIZATION TECHNIQUES

Gonçalo CORREIA¹, José Manuel VIEGAS²

Abstract. Growing traffic congestion and the associated pressure in car parking, require the study of innovative measures to reduce the number of cars traveling every day to the city centers, specifically single occupant vehicles. Car pooling is a system by which a person shares his private vehicle with one or more people that have similar or aligned destinations. Studies were conducted to analyze the viability of car pooling clubs in Lisbon’s Metropolitan Area. A simulation program was built based on statistical data from the Commuter trips of single riders and the population characteristics in each borough surrounding Lisbon, analyzing the probability of matching between those people, having in consideration time and capacity constraints. Using an optimization program integrated with a GIS tool one is able to identify some Geographic areas that are better candidates for providing a strong initial case study in carpooling clubs.

1. Introduction

The rising of auto usage deriving from suburban occupation and car ownership growth is making traffic congestion more frequent in urban areas. This results in air pollution, energy waste and unproductive and unpleasant consumption of people’s time. In most Metropolitan Areas of the United States the majority of trips in individual transport are single occupant vehicle trips (SOV). In 1990 approximately 90% of work trips and 58% of

¹ Instituto Superior Técnico, Universidade Técnica de Lisboa, Departamento de Engenharia Civil, Secção de Transportes e Vias de Comunicação, Av. Rovisco Pais 1049-001 Lisboa Portugal, gcorreia@civil.ist.utl.pt
² Instituto Superior Técnico, Universidade Técnica de Lisboa, Departamento de Engenharia Civil, Secção de Transportes e Vias de Comunicação, Av. Rovisco Pais 1049-001 Lisboa Portugal, viegas@civil.ist.utl.pt
the other trips were done in SOV [1]. Some measures have been tested in the last years like passenger transfer to Public Transport and car pooling systems.

Some studies have been conducted in the past to evaluate the viability of carpooling prior to the implementation of a real case study. In 1994 Krishnamoorthy [2] started with the assumption that an 80% match rate would be the minimum to make carpooling viable, and then he divided Melbourne in zones and formulated lists of acceptable trips between zones. The main conclusion of this work was that a 2.5% of population participation rate would result in an 80% success rate. But this is a result made with various simplifications not considering the several time and capacity constraints.

In 1999 Tsao J. [3] used a matrix of 2 mile sided squares to simulate the trip generation and attraction in the city of Los Angeles. The objective was to find an upper-limit for the ride matching probability. He considered a minimum distance for carpooling to be interesting and used a gravitational method for the trip distribution; it was also considered that the density of workers and jobs in each square was a constant. His conclusion was that carpooling wasn’t viable for this city because it is to much sprawled.

To evaluate the minimum scale for these systems viability in Lisbon’s Metropolitan Area a Simulation on carpooling matching, using geographic, time and capacity constraints was built. The method used was to generate random trips for a specific geographic area (typically the borough) and evaluate the possibility of associating those people in a carpool to reach their destination. This method needs to use three main tools: optimization software to analyze the possibility of carpooling between members of the club, a GIS program to generate feasible origin and destination coordinates for the random trips, and a Visual Basic Program to make the interface between these two.

2. The Method

The Long Term Carpooling Problem is a well known NP-Complete problem [4] that’s why some heuristics have been used to find the optimal matches between participants, like the ANTS Heuristic [5], this search for the best optimal solution for sets in a range of 50 to 225 clients, but at the cost of a significant CPU time.

The objective of this simulation was not to find the best optimal solution, but to find acceptable feasible matches between participants in order to evaluate the matching probability for large number populations, typically all the single riders of an entire borough. Hence this had to be done with some rigour but also with considerable speed.

The method that we found to do this is to consider a typical four-indices formulation using binary variables and the time and capacity constraints of the LCPP problem [5], but instead of using a classical heuristic to try finding an optimal solution for the whole set of clients, this set was dived into clusters, of up to Nmax clients which were analysed independently afterwards.

In summary, these groups are analysed in a first iteration for compatible clients, and afterwards the clients which are part of a group which complies with the capacity constraint, are saved as grouped clients and the others are gathered for a next iteration where clusters are again formed for analysis.

In the next sub-section the LCPP Problem is revised, and then the process of clustering the users and finding the matches for a large population is explained.
2.1. The LCPP Problem

The LCPP (Long Term Car Pooling problem) can be defined as follows. A number n of users must reach their work destination, and latter on the day get back home. The problem objective is to partition the set of users into subsets, or pools, such that each pool member in turns will pick up the remaining members in order to drive together to the work place and back, not necessary the same work place (this is different from the past formulation). The present approach considers that the server picks up his colleagues, takes them to his workplace and then they have to walk from this point to their destinations. In the afternoon they return to this point were they are picked up by the server and taken to their homes.

Each user i enlisted in the car pooling program specifies:

- The maximum extra time $T_i$ he is willing to accept, when picking up colleagues, in addition to the time needed to drive directly from home to the workplace or back;
- The minimum time $e_i$ acceptable for leaving home;
- The maximum time $u_i$ acceptable for arriving at work;
- The minimum time $e_i'$ acceptable for leaving work;
- The maximum time $u_i'$ acceptable for getting back home;
- The capacity $Q_i$ of his car;
- The maximum distance $DistMAX_i$ user i is willing to walk from the server destination to his workplace;
- Another parameter associated with each user is the penalty $p_i$ incurred when the user cannot be pooled with anyone else.

The objective is to define user pools such that as few cars as possible are used and that the routes to be driven by the drivers are as short as possible, subject to time and capacity constraints. Note that pools are supposed to be stable over a period of time and will not change every day. This means that the number of people in a pool will be at most equal to the capacity of the smallest car among those owned by pool members, since each member will eventually pick up all other ones, in a day-to-day or week to week basis.

The LCPP as defined above is actually a multiobjective problem, requiring to [5]:

- Maximizing car usage, thereby minimizing the number of cars travelling to/from work;
- Minimizing the length of the path to be driven by each employee, when acting as a driver;

The problem structure suggests that it is possible to combine these two objectives in a single objective function, as follows.

Let $k$ be a pool of clients. Each of them, on different days, will use his car to pick up the other pool members and go to work (and latter come back), thus he has to define an Hamiltonian path on the partial subgraph of $G$ identified by $k$, starting from the node associated to his house, passing through all the other nodes and ending at the workplace.

Let $Hpath(i, k)$ be such an Hamiltonian path, starting from $i \in k$, connecting all $j \in k\{i\}$ and ending at the workplace of each server.

---

Hpath \((i,k)\) is a feasible path if \(|k| \leq Q_j, \forall j \in k\), all user constraints are met.

The minimum path, \(\text{min}_{\text{path}}(i,k)\), for \(i \in k\) is the shortest feasible Hamiltonian path for \(i\).

\[\text{min}_{\text{path}}(i,k) = \left(40+20+40+63\right) + \left(60+45+40\right) = 163\]

\[\text{min}_{\text{path}}(2,k) = \left(40+45+40+57\right) + \left(60+45+40\right) = 182\]

\[\text{min}_{\text{path}}(3,k) = \left(40+45+40+57\right) + \left(60+45+40\right) = 182\]

\[\text{min}_{\text{path}}(4,k) = \left(40+20+40+63\right) + \left(60+45+40\right) = 163\]

\[\text{cost}(k) = \frac{163+182+182+163}{|k|} = 172.5\]

In this formulation, it is assumed that the shortest paths are chosen. The cost of a pool \(k\) in then defined to be:

\[\text{cost}(k) = \begin{cases} \sum \text{min}_{\text{path}}(i,k) / |k| & \text{if } |k| > 1 \\ c_{io} + p_i & \text{otherwise} \end{cases}\]

\[\text{(1)}\]

If a person is alone, this has an increased penalty, whose amount is associated with him. The cost of a complete solution is the sum of the costs of all the pools in it, that is,

\[\text{cost}(\sigma) = \sum_{k \in \sigma} \text{cost}(k)\]

This perspective optimizes both objective functions. Provided that the penalty \(p_i\) of a client is sufficiently bigger than 0, it is more convenient to pool clients together than to leave them alone.

2.1.1. A four indices mathematical formulation

The four indices formulation that translates the LCPP defined above can be summed up by the following variables and constraints:
\( x_{ij}^{hk} \): binary variable equal to 1 iff arc \((ij)\) is in the shortest Hamiltonian path of a server \(h\) of a pool \(k\);

\( y_{ik}^{h} \): binary variable equal to 1 iff client \(i\) is in pool \(k\);

\( z_{i}^{h} \): binary variable equal to 1 iff client \(i\) is not pooled with any other client;

\( S_{i}^{h} \): non negative variable denoting the pick-up time of client \(i\) by server \(h\);

\( F_{i}^{h} \): non negative variable denoting the arrival time of each client \(i\) at his workplace when travelling with server \(h\);

\( H_{i}^{h} \): non negative variable denoting the departure time of each client from his workplace travelling with server \(h\);

\( L_{i}^{h} \): non negative variable denoting the arrival time of client \(i\) at home, driven by server \(h\);

\( l_{i} \): the maximum time acceptable for arriving at work;

\( e_{i} \): the minimum time acceptable for leaving home;

\( l'_{i} \): the maximum time acceptable for arriving home;

\( e'_{i} \): the minimum time acceptable for leaving work;

\( Q \): The capacity available in each car;

\( T_{i} \): The maximum driving extra time user \(i\) is willing to accept, when picking up colleagues, in addition to the time needed to drive directly from home to the workplace or back;

\( t_{ij} \): time travel between trip origins of clients \(i,j\);

\( t_{OD_{ij}} \): time travel between trip origin of client \(i\) and destination of client \(j\);

\( t_{0ij} \): time travel between trip destination of client \(i\) and destination of client \(j\) – this is a walking distance;

\( DistMAX_{i} \): Maximum walking distance that each client is willing to walk from the server destination to each work place;

\( d0 \): destination distance between clients;

\( K \): index set of all pools;

\( C \): index set of all clients
Objective function (LCPP):

\[
Z_{LCPP} = \min \left( \sum_{k \in K} \left( \sum_{h \in H} \sum_{j \in A} c_{ih} y^k_{ij} / y_j + \sum_{k \in K} \sum_{h \in H} \sum_{j \in A} c_{ih} y^k_{ij} / y_j + \sum_{i \in C} p_i \xi_i \right) \right)
\] (2)

\[
\sum_{j \in C \cap \{0\}} x^h_{ij} = y_{ik}
\] (3)

Force a client i to be declared to be in pool k, if there is a path originated in h going from i to j;

\[
\sum_{i \in C \cap \{0\}} x^h_{ji} = y_{jk}
\] (4)

Force a client j to be declared to be in pool k, if there is a path originated in h going from j to i;

\[
\sum_{j \in C} \sum_{i \in C} x^h_{ij} = \sum_{i \in C} x^h_{ij}
\] (5)

Continuity of the paths;

\[
\sum_{k \in K} x^h_{ik} + \xi_i = 1
\] (6)

Force each client to be assigned to a pool or to be penalized;

\[
\sum_{i \in A} x^h_{ij} \leq Q_h
\] (7)

Car capacity limitation in each group;

\[
T_h + \tau_{h0} \geq F^h - e_h
\] (8)

Maximum extra travel time;

\[
\sum_{k \in K} \sum_{i \in C} x^k_i \leq \sum_{j \in C} y^k_j
\] (9)

Disables the possibility of forming groups of only one element.

Other constraints are needed to establish feasible pick-up and drop out times for the groups, those are combinations of the variables: \( S^h_i, F^h_i, H^h_i \) and \( L^j_i \).

### 2.2. The Process for determining the possible groups

As it was mentioned in the Introduction the algorithm used in this simulation uses the Cluster analysis (K-means) to find small groups of clients to be analysed by the optimization software. Using the distance equal to the distance between the participants both in the origin and destination, we can identify groups of people with higher probability of being partners in a pool group. The following scheme shows how the integration between clustering and the optimization process works in determining groups of clients.
3. Preliminary results

The following figure is the result of the application of the simulation program for 20% of the population who is a single driver coming from the Borough of Sintra to Lisbon in the morning and returning in the evening (the typical Commuter trip) this represents 4882 people. The light dots represent the drivers that weren’t able to find a match, whilst the dark dots represent the different groups formed; it is easy to identify areas for which the probability of matching is higher.
Acknowledgements

A word of appreciation is due to Intergraph Portugal for the free software license of Geomedia Professional 5.2 and technical support. Thanks also to Bob Daniel from Dash Optimization for the free Xpress-MP and Xpress-SLP software license.

References

A DYNAMIC MODEL OF POINT-QUEUE PROCESS

Hilmi Berk CELIKOGLU, Mauro DELL’ORCO and Ergun GEDIZLIOGLU

Abstract. In this paper, a mesoscopic dynamic network loading model is proposed, based on discrete packets and taking into account the vehicles’ acceleration, to model the point-queue process on a highway node with multiple merging and diverging links. The model is run using theoretical input data to simulate point-queueing in over-saturation conditions. The results show that the model appears realistic in the representation of diverging link flow dynamics and is quite easy to calculate.

1. Introduction

Point-queueing and physical queueing are the two main assumptions that have been made in problems of dynamic network loading (DNL) in order to model link and network performances. The queue spillback can only be captured by physical-queue approach, which is more realistic and both studied and modelled in details by Newell [18]. So, the recent trend on traffic flow modelling for dynamic traffic assignment (DTA) is to propose models with physical-queue assumption. However capturing the effects of physical-queue in DNL modelling brings difficulties in obtaining an optimal solution of a DTA problem when a whole network is considered. An example is the calculation of route travel times by an averaging method, which is not consistent with respect to route flows. This leads to obtain convergent algorithms in DTA scheme. As an alternative, the point-queue assumption handles vehicles as points without physical lengths. The storage capacity of each link can be ignored. The queue spillback on a link can be simulated by assuming the existence of a buffer area in the initial node of the link, for the temporary storage of vehicles.

1 Technical University of Istanbul, Faculty of Civil Eng., Division of Transportation, Ayazaga Campus, 34469, Maslak, Istanbul, Turkey, hbcelikoglu@ins.itu.edu.tr
2 Technical University of Bari, Department of Highways and Transportation, Via Orabona 4, 70125, Bari, Italy, dellorco@poliba.it
3 Technical University of Istanbul, Faculty of Civil Eng., Division of Transportation, Ayazaga Campus, 34469, Maslak, Istanbul, Turkey, egedizlioglu@ins.itu.edu.tr
exceeding the maximum density. Therefore, all links can contain unconstrained number of vehicles and capacity constraint on a link can be applied without numerical and computational difficulties. Moreover, the outflow rate of a link is only affected by its own flow considering that the downstream links will always have sufficient storage capacities.

Since point-queues cannot traverse a junction horizontally, they can also be referred as vertical queues. The point-queue assumption is appropriate when the links are long enough and the storage capacities of the links are high enough. In the literature, point-queue assumption has been made in a varying structure of flow models adopting both exit-flow function approach \([16, 17, 5, 6, 14, 21, 22]\) and in travel time function approach \([2, 3, 13, 19, 23, 24, 7]\) to perform DNL.

In this paper, a mesoscopic dynamic node model (DNM) for network loading is proposed, based on discrete packets, to model the point-queue process on a freeway node with multiple merging and diverging links. The model is run using theoretical input data to simulate point-queueing in over-saturation conditions. General structure of the proposed model and formulation of its components, link model and node rules, are explained in the following section. The third section summarizes the validity of the proposed model. Finally, the last section concludes the paper with evaluating the results of the model runs.

2. Dynamic node model

The presented DNM has two components; a mesoscopic link model set with an exit link function formulation, general structure of which is previously introduced by Dell’Orco \([10]\), and an algorithm written with a set of node rules considering the constraints of conservation, flow splitting rates and non-negativities. First, the time-varying flows entering (inflows) to multiple merging links simultaneously are input to the mesoscopic link model. The exiting flows (outflows) from these merging links are computed regarding the link and flow characteristics. Then outflows of the merging links are input to a node as inflows. These conflicting flows are processed within the predefined splitting rates and characteristics of the diverging links, and then the nodal exiting flows are computed. Delays forming as a result of capacity constraint are represented with a conical delay function having desirable properties introduced by Spiess \([20]\).

2.1. Mesoscopic link model component

Mesoscopic models assume that vehicles are either grouped in discrete packets or spread within continuous packets. The packets are arranged in order of their entry time and the travel time is calculated, assuming that packets experience junction delays as a function of current flow on the link. Vehicles entering or leaving the link are respectively added or subtracted in the appropriate time intervals to produce new flow estimates. Flow propagation is generally studied with discrete packets when the mesoscopic approach is adopted. The link model proposed in this study is developed by both considering the over-saturation phenomenon and improving the computational efficiency on a previously proposed link model \([10]\). This model, also utilized in some other studies \([12, 11, 8]\), is set out with link exit function formulation, discretisation on time dimension, defining capacity constraint rules for over-saturated states and uniformly accelerated speed
assumption, which allows a realistic representation of outflow dynamics. Model has an iterative structure, which enables convergence to a target performance with the coded algorithm.

With the assumptions that:

- the vehicles of each packet are all located at the head of the packet;
- the speed is equal for all packets on the link;
- the movement of vehicles is uniformly accelerated;

a relation between speed and density exists and is valid.

2.1.1. Theoretical base of link model component

Regarding [10] the notations and the theoretical formulation of the proposed mesoscopic link model are summarized in the following.

Let \( P \) be the set of feasible paths on the network; the set of vehicles, leaving in the same interval \( j \) and following the same path \( p \) (\( p \in P \)) is called a “packet” \((j, p)\). Additionally, consider a generic link in which:

- \( m_{j, p} \): the total number of vehicles belonging to packet \((j, p)\);
- \( a'(t) \): the acceleration during the interval \([t, t+\Delta t]\) constant and common, in the reference period, for all vehicles on link \(i\);
- \( n_{i,j,p}(t) \): the number of vehicles in packet \((j, p)\) on link \(i\) at time \(t\);
- \( s'_{i,j,p}(t) \): the position on link \(i\) of the head of the packet \((j, p)\) at time \(t\). The value of \( s'_{i,j,p}(t) \) is 0 if packet \((j, p)\) is not on the link;
- \( n_i(t) \): the total number of vehicles on link \(i\) at time \(t\);
- \( k_i(t) \): the density on link \(i\) at time \(t\);
- \( N_i(t) \): the number of vehicles in exiting flow from link \(i\) at time \(t\);
- \( N_i'(t) \): the number of vehicles that exceeds the capacity and can not exit from link \(i\) at time \(t\);
- \( d_i \): the length of link \(i\);
- \( V'(t) \): the speed at time \(t\) on link \(i\), common to all vehicles on the link;
- \( w_i(t) \): the exiting flow from link \(i\) at time \(t\).

Speed is a function of average density \( k_i(t) \): \( V'(t) = V(k_i(t)) \), where \( k_i(t) = n_i(t)/d_i \). So, the total number of vehicles on a link can be calculated by Eq. 1.

\[
n'(t) = \sum_{p \in P} \sum_{j \in P} n_{j,p}(t)
\]  

Relying on to the existence of a valid relationship between speed and density, the variables \( V'(t), n_{i,p}(t), \) and \( s_{i,p}(t) \) can be derived from relationships given in Eqs. 2, 3, 4 and 5.

\[
s'_{i,j,p}(t) = s\left(V'(t-\Delta t) V'(t) s_{i,j,p}'(t-\Delta t)\right)
\]

\[
n_{i,j,p}(t) = n\left(s_{i,j,p}'(t) s_{i,j+1,p}'(t)\right)
\]

\[
N_i'(t) = N\left(N_i(t-\Delta t) n'(t)\right)
\]

\[
V'(t) = V\left(n'(t)\right)
\]
Here, \( n_{ij}^i(t) = 0 \) when \( j > t \); \( s \) and \( V \) are continuous functions. The outflow at time \( t \) on link \( i \), \( w^i(t) \), is also obtained as a function of \( N(t) \), given by Eq. 6.

\[
 w^i(t) = w\left( n^i(t) \right)
\]  

(6)

The speed can be assumed equal for all vehicles on the link, or assigned to each packet at the entry time. In the first case, since it would be \( V(t) = V(N(t)) \), the speed of each packet would also depend on the vehicle number behind the packet itself. However, the influence of this assumption on the model is lower as \( t \) approaches zero (\( t < 0 \)). It is worth noting that all existing mesoscopic models suffer from this drawback, since aggregate performances are used. Because of this approximation, the model based on the relationships given by Eq. 3 is now changed to Eq 7 (\( n_{ij}^i(t) \) is only a function of \( s_{ij}^i(t) \) and can only have two values)

\[
n^i_j(t) = \begin{cases} 0, & \text{if } s^i_j(t) = 0 \\ m^i_j, & \text{if } s^i_j(t) > 0 \end{cases}
\]  

(7)

The later expressions of the model are given through Eqs. 8-13.

\[
s^i_j(t) = s(\mathcal{V}^i(t - \Delta t), V^i(t), s^i_j(t - \Delta t))
\]  

(8)

\[
n^i_j(t + \Delta t) = \begin{cases} 0, & \text{if } s^i_j(t + \Delta t) = 0 \\ m^i_j, & \text{if } s^i_j(t + \Delta t) > 0 \end{cases}
\]  

(9)

\[
n^i(t + \Delta t) = \sum_{j=1}^{m} \sum_{p=1}^{n} n^i_j(t + \Delta t)
\]  

(10)

\[
N(t + \Delta t) = N^n(t) n^i(t + \Delta t)
\]  

(11)

\[
V^i(t + \Delta t) = V(n^i(t + \Delta t))
\]  

(12)

\[
w^i(t + \Delta t) = w(n^i(t + \Delta t))
\]  

(13)

Depending on the physical assumptions, the values of speed and space variables can be calculated by Eq. 14 and Eq. 15.

\[
V^i(t + \Delta t) = V^i(t) + \left( \frac{1}{2} a^i(t) \Delta t \right)
\]  

(14)

\[
s^i_j(t + \Delta t) = s^i_j(t) + V^i(t) \Delta t + \frac{1}{2} a^i(t) \Delta t^2
\]  

(15)

The model given through Eqs. 8-13 is a fixed-point problem with respect to the variable \( V^i(t+\Delta t) \), which can be resolved by the Method of Successive Averages. The speed \( V^i(t+\Delta t) \) is at first calculated through Eq. 14. Then, applying the model, the speed value is calculated by successive iterations as given in Eq. 16, where \( V^i_y(t+\Delta t) \) is the value of speed \( V^i(t+\Delta t) \) at iteration \( y \).

\[
V^i_y(t + \Delta t) = \left( \frac{1}{y} \cdot V^i(t \left( 1 + \left( \frac{k-1}{k} \cdot V^i(t + \Delta t) \right) \right) \right)
\]  

(16)

The iteration stops when the difference between two consecutive speed values is not greater than a fixed threshold. Then, the current value of acceleration is calculated as in Eq. 17 and used as input in successive calculations.

\[
a^i(t + \Delta t) = \frac{V^i(t + \Delta t) - V^i(t)}{\Delta t}
\]  

(17)
The capacity constraint for the exiting flows requires a comparison between the calculated outflow and the capacity of the link. The number of exiting vehicles from a link is updated with the relation given by Eq. 18 in the iterative process of the model:

\[
N^i(t + \Delta t) = \begin{cases} 
N^i(t) + n^i(t + \Delta t) & \text{if } N^i(t) + n^i(t + \Delta t) \leq C^i \cdot \Delta t \\
C^i \cdot \Delta t & \text{if } N^i(t) + n^i(t + \Delta t) > C^i \cdot \Delta t
\end{cases}
\]  

(18)

where \( C^i \) denotes the assigned capacity of the link \( i \).

### 2.2. Node component

The node component of the proposed DNL model is derived from the rules proposed in the work by Celikoglu [8] and also used in a recent study [9]. The variables, with their notations, of the node algorithm coded for the proposed DNM are given below.

- \( FW_k \): The set of links that are diverging from a node \( k \).
- \( BW_k \): The set of links that are merging to a node \( k \).
- \( i \): Generic link included in the set of merging links to node \( k \) (\( i \in BW_k \)).
- \( r \): Generic link included in the set of diverging links from node \( k \) (\( r \in FW_k \)).
- \( w^{ik}(t) \): The exiting flow from link \( i \) that enters node \( k \) at time \( t \) (\( i \in BW_k \)).
- \( u^{kr}(t) \): The total flow entering to link \( r \) at time \( t \), which exits from node \( k \) at time \( t \) (\( r \in FW_k \)).
- \( w^{ikr}(t) \): The partial flow demanding to enter link \( r \) at time \( t \), which exits from link \( i \) at time \( t \) (\( i \in BW_k \) and \( r \in FW_k \)).
- \( u^{ikr}(t) \): The partial flow supplied at link \( r \) at time \( t \), which exits from link \( i \) at time \( t \) (\( i \in BW_k \) and \( r \in FW_k \)).
- \( C^r \): The capacity of link \( r \) (\( r \in FW_k \)).
- \( PQ^r(t) \): The total number of vehicles that are stored in the point-queue at the entrance of link \( r \) at time \( t \) (\( r \in FW_k \)).
- \( G^r(t) \): The delay occurring due to the capacity constraint and present point-queue on link \( r \) at time \( t \) (\( r \in FW_k \)).
- \( \alpha^{ir} \): The partial flow splitting rate from a merging link \( i \) to a diverging link \( r \) (\( i \in BW_k \) and \( r \in FW_k \)).

A transportation network \( \Omega = (N, I) \) composed of a set of nodes \( k, k \in N \) and a set of directed links \( i \) and \( r \), that (\( i \in BW_k \) or \( r \in FW_k \)) \( \subseteq I \) is considered. Traffic originates at nodes \( o, o \in O \), \( O \subseteq N \) and is destined to nodes \( d, d \in D, D \subseteq N \). A set of paths \( p, p \in P \) connects \( o \)-\( d \) pairs. Since the model is proposed for a single node, the path flows are only used to define splitting rates. Traffic departs from origins in the interval \([0, T]\) and all traffic arrive at destination within the interval \([0, T']\) (\( T' > T \)). The total amounts of control variables, inflow \( u^i(t) \) and outflow \( w^i(t) \), are calculated as given in Eqs. 19 and 20 respectively.

\[
\begin{align*}
  u^i(t) &= \sum_{i \in BW_k} u^{ir}(t) \\
  w^i(t) &= \sum_{i \in BW_k} w^{ikr}(t)
\end{align*}
\]
The nodes rules component of the presented DNM contains another set of constraints considering the flow conservation principal, diverging link capacities and splitting rates at a node in addition to the present constraints previously defined for the link model component. The main difference of the proposed DNM in comparison to other models is that it respects capacity constraints regarding to splitting rules and consequently holds FIFO rule. For the link model component of integrated model structure has been set out with the point-queuing assumption, the point-queues and the delays calculated in the presence of these vertical queues are considered instead of the physical queues and the delays occurring as a result of over-saturation. A point-queue at the entrance of a diverging link is present only if at least one of the events given below exits:

- If the flow requiring to enter a diverging link exceeds this link’s capacity,
- If there is any amount of exceeding flow had by the computation for the previous time interval, depending on to time discretisation to have a solution for the formulation.

Delays due to flow states on each diverging link have been calculated with the conical delay function proposed by Spiess [20]. Note that the calculated delays are assumed to be solely dependent to the diverging link capacities. It is assumed that there becomes no delay as a result of flow conflicting. So, the partial flows exiting from merging links are considered to traverse junction area without any loss. This assumption is consistent, since the sample junction is unsignalized and is not an urban junction. Moreover, the relative increase or decrease on total delays calculated with (delay due to flow conflict)+(delay due to flow condition) do not have significant deviations in comparison to the case that delays due to flow conflict are not considered. The partial splitting rates from merging links to diverging ones are assumed to be derived from the path flows while coding the algorithm for DNM.

A constraint derived from diverging link capacity is that the total flow demanding to enter link r at time t can be equal at most to the capacity of link r:

\[ u^{kr}(t) \leq C_r; \quad \forall r \in FW_k \tag{21} \]

Regarding to flow conservation principal, the total flow attempting to enter all diverging links at time t can be at most equal to the amount of total flows exiting from the whole merging links at time t.

\[ \sum_{i \in BW_k} u^{ki}(t) \leq \sum_{i \in BW_k} w^i(t); \quad \forall i \in BW_k \text{ and } \forall r \in FW_k \tag{22} \]

Respecting to capacity constraint and above mentioned delay assumption, the total flow attempting to enter all diverging links can be determined from Eq. 23.

\[ \sum_{i \in BW_k} u^{ki}(t) = \begin{cases} \sum_{i \in BW_k} w^i(t) & \text{if } \sum_{i \in BW_k} w^i(t) \leq \sum_{i \in FW_k} C^i; \quad \forall i \in BW_k \text{ and } \forall r \in FW_k \\ \sum_{i \in BW_k} C^i & \text{if } \sum_{i \in BW_k} w^i(t) > \sum_{i \in FW_k} C^i; \quad \forall i \in BW_k \text{ and } \forall r \in FW_k \end{cases} \tag{23} \]

When the whole process to model the node is considered as the modelling horizon \([0, T]\), the inequality constraint given with Eq. 22 turns out to an equality constraint to hold, given by Eq. 24.

\[ \sum_{i \in BW_k} \sum_{r \in FW_k} u^{kr}(t) = \sum_{i \in BW_k} \sum_{r \in FW_k} w^{ik}(t); \quad t \in [0, T]; \quad \forall i \in BW_k \text{ and } \forall r \in FW_k \tag{24} \]

There are two kinds of split factors that can be used in system dynamics which are path based and link based. Since the path based split factors are considered in the network assignment context, the link based split rates are appropriate in modelling a single node and are used in this study.
Since the proposed model is for a single, link the split factors are calculated assuming that we have information on the currently transferred flow rates between BW_k and FW_k link sets, which are obtained by from the path flows. The proportion \( \alpha \) that is splitted from a merging link \( i \) to a diverging link \( r \) is calculated as given in Eq. 23. The constraint associated with this split factor inputs considering the conservation principal is given by Eq. 26. Considering Eq. 25, it is seen that the FIFO rule holds.

\[
\dot{a}^{ir} = \frac{u^{ir}(t)}{w^{ir}(t)} , \quad \dot{a}^{ir} \geq 0 \quad (25)
\]

\[
\sum_{i \in W_k} \dot{a}^{ir} = 1 \quad (26)
\]

When the capacity constraint on diverging links is considered, the relations given with Eqs. 25 and 26 change and \( \alpha \) becomes calculated for two different conditions as given in Eq. 27.

\[
\dot{a}^{ir} = \begin{cases} \frac{u^{ir}(t)}{w^{ir}(t)} , \quad \text{if } u^{ir}(t) = \sum_{i \in W_k} w^{ir}(t) \leq C' \\
C' \frac{w^{ir}(t)}{w^{ir}(t)} , \quad \text{if } u^{ir}(t) = \sum_{i \in W_k} w^{ir}(t) > C' \end{cases} \quad \text{and } \dot{a}^{ir} \geq 0 \quad (27)
\]

Constraint given in Eq. 26 is valid if the total flow entering to a diverging link does not exceed the capacity of this link. Otherwise, inequality constraint given by Eq 28 should hold.

\[
\sum_{i \in W_k} \dot{a}^{ir} < 1 \quad (28)
\]

Respecting to the point-queueing assumption that is made both in link model and for node rules, the amount of the flow that can not enter a diverging link due to capacity constraint is stored virtually in a vertical plane dimension at the entrance of this link. A possible point-queue (PQ), calculated in terms of number of vehicles, at the entrance of a diverging link can be determined by assuming that there is no initial queue when \( t=0 \), as given by Eq. 29.

\[
PQ'(t) = \begin{cases} 0, & \text{if } (t > 0) \lor \left( u^{ir}(t) + \frac{PQ(t \cdot \Delta t)}{\dot{A}t} \leq C' \right) \\
\left( u^{ir}(t) - C' \right) \dot{A}t + \frac{PQ(t \cdot \Delta t)}{\dot{A}t} & \text{if } u^{ir}(t) + \frac{PQ(t \cdot \Delta t)}{\dot{A}t} > C' \end{cases} \quad (29)
\]

The computation of the node flow rates uses a model which is described above, it is but of course is one of several possible approaches to this problem [15]. The formulation for the node model component [4] has been used by the authors in a complete analytical implementation of the model [1]. It is in fact the simplest way of solving the node problem while respecting the FIFO rule. It can be written as a problem of maximizing the total flow \( Z \) through node k as given in Eq. 33:

\[
\text{max} \quad Z = \sum_{i \in W_k} \sum_{i \in BW_k} \sum_{i \in FW_k} w^{ir}(t) \quad (30)
\]

subject to the constraints given through Eqs. 21-28. The optimization problem given by Eq. 30 is solved with simulation by removing the sum up for modelling horizon bound. Simulation process of the proposed model lasted as the inflows to merging links are wholly discharged from the entire node structure.
3. Validity of proposed DNL model

The validity of the proposed DNL model is tested on a hypothetical highway node. The sample node structure is set out with four merging links and three diverging links. For the proposed model is set out to model flow propagation at a single node, the flow splitting factors are assumed to be known with the information on path flows. Therefore, these factors are assumed to be constant in under-saturation condition. In over-saturation condition, the factors are updated upon the allowable flow rates regarding to capacity constraints. The time-varying inflows \( u(t) \) to merging links are generated theoretically to fit single-peak sinusoidal curve and the sample node is configured to simulate the over-saturation condition for each merging link with the capacity exceeding rates of 20%, 3%, 2%, and 8% for the first, second, third and the fourth merging link (ML) respectively.

The simulation times needed to discharge the dynamic loadings completely from merging links with the link model component computation of the proposed DNM are given in Table 1.

<table>
<thead>
<tr>
<th></th>
<th>ML1</th>
<th>ML2</th>
<th>ML3</th>
<th>ML4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simulation time (min)</td>
<td>68</td>
<td>63</td>
<td>63</td>
<td>64</td>
</tr>
<tr>
<td>CPU time (sec)</td>
<td>0.094</td>
<td>0.062</td>
<td>0.063</td>
<td>0.031</td>
</tr>
</tbody>
</table>

Table 1. Times needed for link model component simulations.

The time-varying outflows \( w(t) \) computed by the link model component of the proposed DNM are plotted with the corresponding inflows as shown in Figure 1.
The inflows $u'(t)$ that are entering to diverging links computed with DNM runs are shown in Figure 2.

![Figure 2. Inflows to diverging links computed with DNM within second configuration.](image1)

The point-queuings, $PQ(t)$, formed at the entrance of diverging links are shown in Figure 3.

![Figure 3. Point-queues at entrance of diverging links.](image2)

The maximum and the minimum simulation times for computing the merging link outflows, the total simulation times and the required total CPU times of DNM had after model run are 68 mins, 63 mins, 72 mins, and 4.625 secs respectively.

4. Conclusions

In this study, a dynamic network loading problem has been formulated using traffic flow relationships. The proposed dynamic node model is utilized to carry out a single node application.
Under the set node configuration it is seen that the outflows of the link model component \( w^{ik}(t) \) existed respecting to capacity constraints and the diagrams of these outflows seemed alike the sinusoidal inflow curves as plotted in Figure 1. The flows that are routed at the node \( w^{ik}(t) \) are used to calculate the inflows to diverging links \( u^{kr}(t) \) by considering the predefined path flows and capacity constraints. The capacity exceeding rates on the first, the second, and the third diverging link are 91%, 28%, and 58% respectively. Considering Figure 2, it is clearly seen that the flows requiring to enter the diverging links are above over-saturation rates, but it is also seen that capacity restraint is respected.

The results show that the model appears realistic in the representation of point-queuing process and diverging link flow dynamics, and is quite easy to calculate regarding to required processor time. The future extension of this study will be on the application of the proposed model to a general network.

References


A DYNAMIC NODE MODEL WITH A QUADRATIC TRAVEL TIME FUNCTION BASED LINK MODEL

Hilmi Berk CELIKOGLU

Abstract. In this paper, a dynamic node model (DNM) is proposed to model time-varying traffic flows at an unsignalized highway junction. The presented DNM has two components; a link model set with a quadratic travel time function of second order, and an algorithm written with a set of node rules considering the constraints of conservation, flow splitting rates and non-negativities. For a point queue assumption is made throughout the whole model structure, physical queue concept is not studied. Delays forming as a result of capacity constraint are represented with a conical delay function having desirable properties.

1. Introduction

Static structure of the previously proposed models fails to represent time-varying flows in the context of demand-supply interaction. Consequently, dynamic network loading (DNL) has been the efficient tool for application and evaluation of real-time control strategies in traffic assignment modelling. Many approaches on DNL, having a wide variety on both modelling concept and discretisation, have been proposed since late in 80ies. A main classification of these models can be made by considering the representative function used to calculate a control variable, which are preferred as the link outflow and the link travel time in most of the studies. Link exit function based formulations use an arbitrary function to calculate the exiting flows from a link and have the drawbacks of violating the first-in-first-out (FIFO) rule and generating unrealistic flow propagation inconsistent with the speed. Link flows and desirable properties can instead be described more accurately by models based on a link travel time formulation that is also adopted in this study. General structure of the proposed model and formulation of its components, link model and node rules, are explained in the following section. Finally, the last section concludes the paper with evaluating the results of the model runs.

1Technical University of Istanbul, Faculty of Civil Eng., Division of Transportation, Ayazaga Campus, 34469, Maslak, Istanbul, Turkey, hbcelikoglu@ins.itu.edu.tr
2. Dynamic node model

The presented dynamic node model (DNM) has two components; a link model set with a quadratic travel time function of second order, general structure of which is previously introduced by Carey and McCartney [6], and an algorithm written with a set of node rules considering the constraints of conservation, flow splitting rates and non-negativities. At the beginning, the time-varying flows entering (inflows) to multiple merging links simultaneously are input to the travel time formulation based link model. The exiting flows (outflows) from these merging links are computed regarding the link and flow characteristics. Then outflows of the merging links are input to a node as inflows. These conflicting flows are processed within the predefined splitting rates and characteristics of the diverging links, and then the nodal exiting flows are computed. Delays forming as a result of capacity constraint are represented with a conical delay function having desirable properties introduced by Spiess [13]. The proposed DNM has the integrated structure shown below in Fig. 1.

![Figure 1. Integrated structure of the proposed dynamic node model.](image-url)

2.1. Travel time function based link model component

In travel time based link models, the link travel time of a user arriving the entrance of a link at time \( t \) is introduced as \( \tau(t) \). Since the traffic flow is continuous and is resembled as a fluid, the user is treated just as a particle of this fluid. Each of the users arriving the entrance of a link have their own travelling time \( \tau \), which is the exact time to traverse a link. As the propagation of flows through a link is described by the relationships between
control variables such as link travel time $\tau(t)$, link inflow $u(t)$, link outflow $w(t)$ or the number of vehicles on the link $x(t)$ at each point in time, the travel time $\tau(t)$ for traffic entering a link at time $t$ can be expressed as given in Eq. 1, which is derived by following Ran et al. [11].

$$\dot{\alpha}(t) = T(u(t), x(t), w(t))$$  \hspace{1cm} (1)

In the former numerical examples [11], the “instantaneous flow-dependent running time” and “instantaneous queuing delay” are assumed to be non-negative and increasing functions of the variables $x$, $u$ and $x$, $w$ respectively. By simplifying the form proposed in the above mentioned numerical examples, the general form of a travel time function can be written as given in Eq. 2, where $\alpha$ is the free-flow travel time.

$$\dot{\alpha}(t) = \dot{\alpha} + f(u(t)) + g(x(t)) + h(w(t))$$  \hspace{1cm} (2)

Different cases of travel time formulation are used in [11, 9, 10, 12, 2, 3, 15] and elsewhere to set out a travel time function based model.

Within the DNL by using the hydrodynamic analogy to represent traffic, the flow $q(t)$ traversing a section of a road can be calculated as given in Eq. 3.

$$q(t) = \lim_{\Delta t \to 0} \frac{\Delta x}{\Delta t}$$  \hspace{1cm} (3)

Regarding Eq. 3 a DNL model should respect the conservation principal derived from continuity equations considering the control variables as given in Eq. 4.

$$x(t) = \int_0^t (u(\dot{u}) - w(\dot{u})) \, \dot{u} \, \frac{dt}{\dot{u}}$$  \hspace{1cm} (4)

Moreover, the traffic entering a link at time $t$ leaves at time $t+\tau(t)$. If the FIFO rule and conservation hold, the number of vehicles that enter the link by time $t$ must equal the number of vehicles that have left by time $t+\tau(t)$ as given in Eq. 5.

$$\int_0^t u(\dot{u}) \, \dot{u} \, \frac{dt}{\dot{u}} = \int_{t+\tau(t)}^t w(\dot{u}) \, \dot{u} \, \frac{dt}{\dot{u}}$$  \hspace{1cm} (5)

Differentiating Eq. 5 with respect to $t$ gives Eq. 6.

$$u(t) = (t + \dot{\alpha}(t)) \left( 1 + \frac{\dot{\alpha}(t)}{dt} \right)$$  \hspace{1cm} (6)

The relationship given by Eq. 7, first presented by Astarita [2, 3], is had by rearranging Eq. 6. Note that the denominator of Eq. 7 has to be positive to make a travel time formulation hold FIFO.

$$w(t + \dot{\alpha}(t)) = \frac{-u(t)}{1 + \frac{\dot{\alpha}(t)}{dt}}$$  \hspace{1cm} (7)

Note that dependent to continuity equation given in Eq. 4, the number of vehicles on a link at time $t+\Delta t$ can be calculated with Eq. 8, which is written by discretising the first derivative (wrt t) of Eq. 4.

$$n(t + \Delta t) = n(t) + (u(t) - w(t)) \cdot \Delta t$$  \hspace{1cm} (8)

In this study, we consider a travel time function, derived from Eq. 2, solely dependent on the number of vehicles on the link as given in Eq 9 (we set $n$ to 2).

$$\dot{\alpha}(t) = \dot{\alpha} + \dot{\alpha}(x(t)) \cdot n \geq 1, \hspace{0.5cm} \dot{\alpha} > 0$$  \hspace{1cm} (9)
The linear form of the travel time function given in Eq. 9 is widely used in most of the travel time formulation based DNL studies [2, 3, 14, 15, 12, 16, 1]. The formulation of the n-th order polynomial travel time function is first introduced and its properties are studied in the work by Carey and McCartney [6]. Taking $T$ and $U$ as travel time and inflow respectively, and substituting these into Eq. 9 gives Eq. 10, where $U\cdot T$ is $x(t)$.

$$\dot{A} - T + \hat{a}U^n T^n = 0$$

(10)

Setting the order of the polynomial to 2 gives a quadratic in $T$ (shown by Eq. 11), if the inflow is assumed to be constant (the numerical study is presented in [6].

$$\dot{A} - T + \hat{a}U^2 T^2 = 0$$

(11)

The capacity of flow $U=\sqrt{4\alpha\beta}$ occurs when the discriminant in (12) is zero. So, terms given in (13) can be drawn from the quadratic.

$$\lim_{U \to 0} [T_1] = \infty, \quad \lim_{U \to 0} [T_2] = \hat{a}$$

(13)

The numerical solution of the quadratic travel time formulation can fail to represent travel times, if the solution algorithm unintentionally converges to wrong travel time.

The link model component of the proposed model in this paper is formulated on travel time base by the series of equations given in (4), (7) and (9). Note that Eq. 4 and Eq. 7 are quite general and do not depend on the selected travel time function. From Eq. 4 and Eq. 9 (note that $n$ is set to 2), travel time dependent on number of vehicles on a link implies Eq. 14.

$$\frac{d\theta(t)}{dt} = 2\hat{a} \left( u(t) - w(t) \right)$$

(14)

Substituting (14) into (7) gives Eq. 15, denominator of which should be positive.

$$w(t + \theta(t)) = \frac{u(t)}{1 + 2\hat{a}(u(t) - w(t))}$$

(15)

The requirement of respecting the FIFO rule is first described in the study by Carey [5]. In the relevant cited papers, the necessary and sufficient conditions for FIFO rule behaviour is presented as given in Eq. 16 when a linear function for travel time calculation is assumed. Consistency of FIFO rule for quadratic travel time function based models can be derived as studied in [6].

$$\theta(t) = \frac{d\theta(t)}{dt} > -1$$

(16)

The travel time formulation given with equations (4), (7) and (9) has been studied extensively in some studies [2, 3, 14, 15, 12, 16, 1, 17, 6, 7] and elsewhere both to provide theoretical basis and application examples for the evolution of DNL concept.

### 2.2. Node component

The variables, with their notations, of the node algorithm coded for the proposed DNM are given below.
- \( \text{FW}_k \): The set of links that are diverging from a node \( k \).
- \( \text{BW}_k \): The set of links that are merging to a node \( k \).
- \( i \): Generic link included in the set of merging links to node \( k \) \((i \in \text{BW}_k)\).
- \( r \): Generic link included in the set of diverging links from node \( k \) \((r \in \text{FW}_k)\).
- \( w^i(t) \): The exiting flow from link \( i \) that enters node \( k \) at time \( t \) \((i \in \text{BW}_k)\).
- \( w^r(t) \): The total flow entering to link \( r \) at time \( t \), which exits from node \( k \) at time \( t \) \((r \in \text{FW}_k)\).
- \( u^i(t) \): The partial flow demanding to enter link \( r \) at time \( t \), which exits from link \( i \) at time \( t \) \((i \in \text{BW}_k \text{ ve } r \in \text{FW}_k)\).
- \( u^r(t) \): The partial flow supplied at link \( r \) at time \( t \), which exits from link \( i \) at time \( t \) \((i \in \text{BW}_k \text{ ve } r \in \text{FW}_k)\).
- \( C_r \): The capacity of link \( r \) \((r \in \text{FW}_k)\).
- \( \text{PQ}^r(t) \): The total number of vehicles that are stored in the point-queue at the entrance of link \( r \) at time \( t \) \((r \in \text{FW}_k)\).
- \( G^r(t) \): The delay occurring due to the capacity constraint and present point-queue on link \( r \) at time \( t \) \((r \in \text{FW}_k)\).
- \( \alpha^r \): The partial flow splitting rate from a merging link \( i \) to a diverging link \( r \) \((i \in \text{BW}_k \text{ ve } r \in \text{FW}_k)\).

A transportation network \( \Omega = (N, I) \) composed of a set of nodes \( k \), \( k \in \mathbb{N} \) and a set of directed links \( i \) and \( r \), that \((i \in \text{BW}_k \cup r \in \text{FW}_k) \subset I \) is considered. Traffic originates at nodes \( o \), \( o \in O \), \( O \subset \mathbb{N} \) and is destined to nodes \( d \), \( d \in D \), \( D \subset \mathbb{N} \). A set of paths \( p \), \( p \in P \) connects \( o \)-\( d \) pairs. Since the model is proposed for a single node, the path flows are only used to define splitting rates. Traffic departs from origins in the interval \([0, T]\) and all traffic arrive at destination within the interval \([0, T']\) \((T' > T)\). The total amounts of control variables, inflow \( u^i(t) \) and outflow \( w^r(t) \), are calculated as given in Eqs. 17 and 18 respectively.

\[
\begin{align*}
 u^i(t) &= \sum_{r \in \text{FW}_k} w^r(t) \\
 w^r(t) &= \sum_{i \in \text{BW}_k} w^r(t)
\end{align*}
\]  

(17) \hspace{0.5cm} (18)

The nodes rules component of the presented DNM contains another set of constraints considering the flow conservation principal, diverging link capacities and splitting rates at a node in addition to the present constraints previously defined for the link model component. The main difference of the proposed DNM in comparison to other models is that it respects capacity constraints regarding to splitting rules and consequently holds FIFO rule. For the link model component of integrated model structure has been set out with the point-queueing assumption, the point-queues and the delays calculated in the presence of these vertical queues are considered instead of the physical queues and the delays occurring as a result of over-saturation. A point-queue at the entrance of a diverging link is present only if at least one of the events given below exits:

- If the flow requiring to enter a diverging link exceeds this link’s capacity,
- If there is any amount of exceeding flow had by the computation for the previous time interval, depending on to time discretisation to have a solution for the formulation.

Delays due to flow states on each diverging link have been calculated with the conical delay function proposed by Spiess [13]. Note that the calculated delays are assumed to be
solely dependent to the diverging link capacities. It is assumed that there becomes no delay as a result of flow conflicting. So, the partial flows exiting from merging links are considered to traverse junction area without any loss. This assumption is consistent, since the sample junction is unsignalized and therefore no specific control action is seeked. Moreover, the relative increase or decrease on total delays calculated with (delay due to flow conflict)+(delay due to flow condition) do not have significant deviations in comparison to the case that delays due to flow conflict are not considered. The partial splitting rates from merging links to diverging ones are assumed to be derived from the path flows while coding the algorithm for DNM.

A constraint derived from diverging link capacity is that the total flow demanding to enter link \( r \) at time \( t \) can be equal at most to the capacity of link \( r \):

\[
    u^{kr}(t) \leq C^r; \quad \forall r \in \text{FW}_k
\]

(19)

Regarding to flow conservation principal, the total flow attempting to enter all diverging links at time \( t \) can be at most equal to the amount of total flows exiting from the whole merging links at time \( t \).

\[
    \sum_{r \in \text{FW}_k} u^{kr}(t) \leq \sum_{i \in \text{BW}_k} w^{ik}(t); \quad \forall i \in \text{BW}_k \text{ and } \forall r \in \text{FW}_k
\]

(20)

Respecting to capacity constraint and above mentioned delay assumption, the total flow attempting to enter all diverging links can be determined from Eq. 21.

\[
    \sum_{r \in \text{FW}_k} u^{kr}(t) = \begin{cases} 
    \sum_{i \in \text{BW}_k} w^{ik}(t); & \text{if } \sum_{i \in \text{BW}_k} w^{ik}(t) \leq \sum_{r \in \text{FW}_k} C^r, \\
    \sum_{r \in \text{FW}_k} C^r; & \text{if } \sum_{i \in \text{BW}_k} w^{ik}(t) > \sum_{r \in \text{FW}_k} C^r, 
\end{cases} \quad \forall i \in \text{BW}_k \text{ and } \forall r \in \text{FW}_k
\]

(21)

When the whole process to model the node is considered as the modelling horizon \([0, T]\), the inequality constraint given with Eq. 20 turns out to an equality constraint to hold, given by Eq. 22.

\[
    \sum_{i \in \text{BW}_k} \sum_{r \in \text{FW}_k} u^{kr}(t) = \sum_{i \in \text{BW}_k} \sum_{r \in \text{FW}_k} w^{ik}(t); \quad t \in [0, T], \forall i \in \text{BW}_k \text{ and } \forall r \in \text{FW}_k
\]

(22)

There are two kinds of split factors that can be used in system dynamics which are path based and link based. Since the path based split factors are considered in the network assignment context, the link based split rates are appropriate in modelling a single node and are used in this study.

Since the proposed model is for a single link and, hence, no path flow is computed, it is impossible to channelise the users that will exit from a merging link \( i \in \text{BW}_k \) and require to enter a diverging link \( r \in \text{FW}_k \) to another diverging link. Therefore, the splitting factors are processed as constant. The split factors are calculated assuming that we have information on the currently transferred flow volumes between \( \text{BW}_k \) and \( \text{FW}_k \) link sets, which are obtained by from the path flows. Flow rate \( \alpha^r \) that is splitted from a merging link \( i \) to a diverging link \( r \) is calculated as given in Eq. 23. The constraint associated with this split factor inputs considering the conservation principal is given by Eq. 24. Considering Eq. 23, it is seen that the FIFO rule holds.

\[
    \dot{\alpha}^r = \frac{u^{ri}(t)}{w^{ik}(t)}, \quad \dot{\alpha}^r \geq 0
\]

(23)
\[
\sum_{i=1}^{n} \hat{\lambda}_{ij} = 1
\]  

(24)

When the capacity constraint on diverging links is considered, the relations given with Eqs. 23 and 24 change and \( \alpha_i \) becomes calculated for two different conditions as given in Eq. 25.

\[
\hat{\alpha}_i = \begin{cases} 
\frac{u_i(t)}{w_i(t)} & \text{if } u_i(t) = \sum_{i=1}^{n} w_{ij}(t) \leq C_i \quad \text{and } \hat{\alpha}_i > 0 \\
\frac{C_i}{w_i(t)} & \text{if } u_i(t) = \sum_{i=1}^{n} w_{ij}(t) > C_i 
\end{cases}
\]  

(25)

Constraint given in Eq. 24 is valid if the total flow entering to a diverging link does not exceed the capacity of this link. Otherwise, inequality constraint given by Eq 26 should hold.

\[
\sum_{i=1}^{n} \hat{\lambda}_{ij} < 1
\]  

(26)

Respecting to the point-queue assumption that is made both in link model and for node rules, the amount of the flow that can not enter a diverging link due to capacity constraint is stored virtually in a vertical plane distance at the entrance of this link. A possible point-queue (PQ) at the entrance of a diverging link can be determined by assuming that there is no initial queue when \( t=0 \), as given by (27).

\[
PQ'(t) = \begin{cases} 
0, & \text{if } (t=0) \quad \text{if } u_i(t) + \frac{PQ'(t-\Delta t)}{At} \leq C_i \\
\left(\frac{u_i(t)}{C_i} - C_i\right) \frac{At}{PQ'(t-\Delta t)} & \text{if } u_i(t) + \frac{PQ'(t-\Delta t)}{At} > C_i
\end{cases}
\]  

(27)

In most of the traffic assignment methods, the effect of link capacity on travel times is specified by means of flow rate-delay functions, which is used to express the travel time on a link as a function of traffic volume. These functions are usually expressed as the product of the free-flow time multiplied by a normalized congestion function \( f(\omega) \):

\[
G'(u'(t)) = t_{\text{free-flow}} \cdot f(u'(t))
\]  

(28)

where the argument of delay function \( G'() \) of a link \( r \) is the ratio flow-volume/capacity \( u'(t) / C_i \) on link \( r \). Note that, dependent on the vertical queuing process, the volume is considered as inflow \( u'(t) \) requiring to enter a diverging link due to capacity constraint made in the proposed DNM. The conical congestion function used in this study to model delays on diverging links have the desired properties of a delay function and have the form given in Eq. 29 [13]:

\[
f\left(\frac{u'(t)}{C_i}\right) = 2 + \sqrt{a^2 \left(1 - \frac{u'(t)}{C_i}\right)^2 + b^2 - a \left(1 - \frac{u'(t)}{C_i}\right)} - b
\]  

(29)

where, \( b \) is given as \( b=(2a-1)/(2a-2) \) and \( a \) is any number larger than 1. Considering the necessity of point-queue assumption made in the proposed DNM, delay due to capacity constraint at a diverging link entrance at time \( t \) is represented by the congestion value \( f'(t) \) as given in Eq. 30 (note that there is no delay when there is no point queue).

\[
f'(t) = \begin{cases} 
0, & \text{if } PQ'(t) = 0 \\
f(u'(t)) & \text{if } PQ'(t) > 0
\end{cases} \quad \forall t \in FW, \text{ and } t \in [0, T]
\]  

(30)
Note that \( f(t) \) is the congestion value and delay at time \( t \) can be calculated by multiplying this value with free-flow travel time as given in Eq. 28.

The computation of the node flow rates uses a model which is described above, it is but of course is one of several possible approaches to this problem [8]. The formulation for the node model component [4] has been used by the authors in a complete analytical implementation of the model [1]. It is in fact the simplest way of solving the node problem while respecting the FIFO rule. It can be written as a problem of maximizing the total flow \( Z \) through node \( k \) as given in Eq. 31:

\[
\max Z = \sum_{i=1}^{n} \sum_{e \in \mathcal{E}_i} \sum_{w \in \mathcal{W}_k} w^{ae}(t)
\]

subject to the constraints given through Eqs. 19-26. The principal weakness of this formulation is that \( \alpha \) are fixed to two values, one for under-saturation and one for over-saturation. As a result, the excess supply can not be re-assigned. Therefore, the optimization problem given by Eq. 31 is solved with simulation by removing the sum up and letting the dynamic loadings discharged within the modelling horizon. Simulation process of the proposed model lasted as the inflows to merging links are wholly discharged from the entire node structure.

3. Conclusions

In this paper a proposition has been made to evaluate the performance of a quadratic travel time function link based DNL by integration to the node concept. The integrated structure of a modelling approach, dynamic node modelling, for intersection representation is, hence, set out. Comparisons are had with the same DNM adopting a linear travel time function based link model instead. The proposed DNM provides realistic results in representing outflow dynamics. In addition, its accuracy in representing the fundamental laws of traffic flow is somewhat better than the other considered model. The proposed algorithm is capable of generating accurate results even in over-saturated conditions respecting to assigned capacity constraints. The computational load of the model seems rather heavy, due to iterative calculation. The model is promising overall and its extension to a general network and its validation with real data will be matters of future development, combined with study and analysis of more effective link model components.

References


CONTAINER STOWAGE OPTIMIZATION WITH CRANE UTILIZATION BALANCING: MODELLING AND HEURISTICS

Masoud Hamedi1, Haghani, Ali1,2, Yang, Saini1

1. Introduction

Because of higher competition among ports in recent times, improving the operational efficiency of ports has become an important issue in containership operations. One of the measures of performance is the berthing time at a port, which is determined by arrangement of containers both within the container terminal and on the containership. The berthing time of a containership is mainly composed of unloading and loading time of the containers. Containers in a containership are usually put into stacks. A container is only accessible if it is on the top of the stack (Last In First Out). The ship visits several ports during a voyage and containers are loaded and unloaded at each port. The task of determining a good container arrangement to minimize the number of re-handlings while maintaining ship’s stability is called stowage planning, which is an everyday problem faced by ship planners.

Containers are loaded to and unloaded from the containership using quay cranes. The problem of allocating quay cranes to the ships and to the ship’s sections is known as crane split. In some cases up to five quay cranes might be allocated to a ship. The technical requirements determine the range in which each quay crane can operate. The ship’s dwell time is determined by the time that the latest crane finishes its job. Since the distribution of the containers over the bays affects crane utilization and overall ship berthing time, crane split and stowage problem are interrelated. Given the configuration of cranes at each visiting port, the stowage planning must take into account the utilization of quay cranes as well as the reduction of unnecessary shifts to minimize the total time at ports over the voyage. Integration of the stowage plan and the crane split results in a more efficient working instruction which increases port utilization.

1 Department of Civil and Environmental Engineering, University of Maryland. College Park, MD 20742. Email: masoud@umd.edu, snyang@eng.umd.edu, haghani@umd.edu
2 Corresponding author Ph: (301) 405-1963 - Email: haghani@umd.edu
This paper focuses on the optimization of containership stowage planning problem with crane utilization balancing. Extensive literature review is preformed. Different solution approaches are explored and summarized.

2. Mathematical model

Each container ship has a cellular structure and each cell is identified using three indices: bay-row-tier. This address is a system of numerical coordinates relating to length, width and height of the container ship. Assume the weight, size, type and destinations of containers to be loaded/unloaded at each port are known. Given the cellular structure of the container ship, the sequence of the ports to be visited and the assignment of quay cranes at each port, a mathematical model can be formulated that minimizes the ship’s berthing time at all ports by minimizing the number of re-handlings and maximizing crane utilization. The stability of the ship and technical constraints of the cranes are taken into consideration.

![Cellular structure of a containership](image)

A 0-1 integer programming model is proposed to minimize the number of unnecessary shifts and maximize crane utilization at each port while satisfying the ship’s stability constraints. Since each unload operation is associated with a load operation the objective function minimizes total number of container handlings. The constraints are to make sure that real world operational requirements such as stability requirements, requirements regarding handling different size containers, unload/load requirement in intermediate ports and crane utilization requirements are met.
3. Heuristics

Although the model is solved to the optimality for several examples, commercial solvers can only handle very small instances of the model. Due to NP-Completeness nature of the problem, a genetic algorithm based approach is implemented to solve real world size problems. Instead of encoding the complete stowage plan into chromosomes which turns out to be very inefficient in terms of both memory consumption and speed, a compact encoding is implemented. In this new compact assignment strategy, grouping policy and sequence of the containers to be loaded at each port are encoded into genes which together form relatively short chromosomes. In the evaluation process the chromosomes are rewarded for lowering the total time spent at all ports and penalized for violating technical and operational constraints. A parallel implementation of the proposed genetic algorithm is implemented as well.

4. Numerical example

Several examples are presented to illustrate the performance of mathematical models and heuristics. For the first group of numerical examples, mathematical models are generated for five hypothetical containerships with approximately the same capacity (20-21 Twenty-feet Equivalent Units) but different structure (number of bays, rows and tiers). Optimal solutions are obtained for all hypothetical containerships by using CPLEX solver. It can be observed that the running time increases exponentially with the increase of tier numbers. The proposed heuristic procedure generates the optimal solution for the same group of problem within very short computational times.

The second group of examples investigates the effect of crane split on the solution. It is assumed that two cranes are available at each port, and each crane can handle one container per unit time. One handling can be loading a container at its origin, unloading at destination or shifting the position at an intermediate port. Two scenarios are compared in this example. In the first scenario the optimization of crane utilization is not enforced while in the second scenario workload is distributed among two cranes. The results shows with balanced crane utilization, the total berth time will decrease by about 8% although the total number of handlings may slightly increase. For large scale problems, the heuristics can generate good solutions within short computation time while the existing optimization software cannot provide any solution in reasonable time.

5. Conclusion

This paper presents an optimization model which focuses on the containership stowage planning problem during a voyage. By minimizing the total number of container handlings and balancing the crane utilization at each port along a voyage, the berthing time can be minimized.
to improve the operational efficiency of ports and ships. Different container sizes, types and the stability requirements of ships are taken into consideration as well. Optimal stowage plan can be obtained by solving the model with commercial optimization software.

The number of variables in the mathematical model grows exponentially with the height of the stack and the number of ports to be visited. In large real world problems, an optimal stowage plan for a containership carrying thousands of containers is too computational expensive to be obtained using existing commercial solvers. Therefore a genetic algorithm based heuristic is developed to produce near optimal solutions in a timely manner. The case study shows the heuristic can provide quality solution and the computational time is much shorter than those of the existing commercial solvers. Extended research is in progress by the authors to develop lower bound for the heuristic and to accommodate more real operational requirements.

References


[12] www.containerhandbuch.de
Abstract. In this paper we study the 2-period probabilistic TSP. In this problem, we have urgent nodes that must be visited immediately and regular nodes that must be visited either immediately or the day after. The objective is to minimize the average long-run delivery costs, knowing the probabilities governing the requests.

The problem is cast as a Markov Process with Costs, which has an exponential number of states in the number of nodes in the network. To overcome this drawback, we also suggest a second Markov approach (Aggregate Model) with a reduced number of states, only polynomial in the number of nodes.

The important relations between the optimal solutions of the original and the aggregate models will be discussed. The viability of the proposed methodology is shown by applying the procedure to a numerical example.

1. Introduction

This paper investigates a specific version of the TSP problem, named the 2-period probabilistic TSP (2-P-P-TSP). In this problem, we have urgent nodes that must be visited immediately and regular nodes that must be visited either immediately or the day after. The objective is to minimize the average long-run delivery costs, knowing the probabilities governing the requests.

The 2-P-P-TSP is motivated by a real-world application concerning blood delivery (see Doerner et al. [3]). The Austrian Red Cross (ARC), a non-profit organization, is in charge of delivering blood to hospitals on their request. In current operations, the ARC is obliged to fulfil any order within the following day. This policy leads to high delivery costs and many extra working hours for its drivers. To reduce costs through higher flexibility the ARC is interested in changing policy, in particular is investigating the possibility of providing two different types of service: one which
delivers the blood within one day and the other within two days. Obviously, these services would have different prices. Under such a policy, the ARC will be confronted with two different types of requests each day $t$: urgent requests from hospitals that want blood delivered within the same day, and regular requests from a set, say $H_r(t)$, of hospitals that allow blood delivering within the following day. The ARC has to decide if it is better to serve hospitals in $H_r(t)$ immediately or the day after in order to minimize the average daily delivery costs. These decisions should depend on the location of hospitals that have to be served and the foreseen blood orders for the following days. Each hospital has a known probability to have an urgent, a regular or no request at all.

The problem can be viewed as a combination of the probabilistic TSP and the two-period TSP. In the two-period TSP (see Butler et al. [1]) the problem is to identify two tours with a combined distance that is minimized and such that:

1. some specified nodes have to be visited by both tours;
2. the remaining nodes are visited exactly once by only one of the tours.

In this paper, we begin with casting the 2P-P-TSP as a Markov decision process to exploit the good theoretical properties of MDP, see the textbook by Puterman [7] for a comprehensive review. From here on, this model is named Exact Model. Since in real applications the number of different states may be prohibitively large, we also propose an aggregate MDP that is an approximation of the exact one but contains a manageable number of states, herein referred to as Aggregate Model. As we will see, the number of states in the Exact Model grows exponentially with $n$ (the number of nodes in the network) and only polynomially in the Aggregate Model. The issue of formalizing MDPs with a smaller number of states is not new in the literature, e.g. see White [8], Whitt [9] and Porteus [6]. However, these results are valid for discounted MDP and rely on the theory of monotone contraction operators (see Denardo [2]). Therefore, all these results cannot be applied to our model since it minimizes the average undiscounted cost. Although the mentioned results cannot be applied to our specific problem, we will show the important relations between the optimal solutions of the Exact and Aggregate models. The 2P-P-TSP is obviously NP-hard since it requires the solution of (many) TSP problems. Moreover, even if an oracle would provide the solution of any TSP in constant time, to find an optimal policy, we would still need to associate an optimal decision to each possible state.

2. Exact Markov Model

The problem described above can be cast as a Markov Process with Costs. A Markov decision process is a discrete-stage sequential decision making problem in a stochastic environment. The key feature of such a process is that the transition probability from the current state to that at the next decision epoch depends only on the current state and not on earlier states and actions of the process. This model, at least in principle, can be solved using an approach originally proposed by Howard [4] for Markov Processes with Rewards. This approach is an iterative procedure and each iteration consists of two phases. In the first phase a linear system of $N$ equalities in
$N$ variables is solved, where $N$ is the number of different states. In the second phase, for each single state, the “best” decision for that state is found using parameters computed in the first phase but irrespective of the decisions to be taken in all other states. Usually the number of iterations required to reach optimality is extremely small. Given a network with $n$ nodes, the state of the network in any given day is described by a vector with $n$ components where the value of each component reflects the status of the corresponding node. The possible values are 0, $u$ or $r$. Value 0 means that there is no request for blood; value $u$ means that there is an urgent request of blood (needs to be satisfied within 1 day); value $r$ means that there is a regular request of blood, i.e., a request that can be satisfied either the same day or the day after. Consequently, the overall number of states grows exponentially with the number of nodes $n$, and is given by $N = 3^n$. Already for $n = 19$, $N$ is greater than a billion. For this reason the exact approach may become impractical for real size instances of the problem.

3. Aggregate Markov Model

Here, we propose an Aggregate Model (AM) whose number of states grows only polynomially with the number of nodes in the network. A state in AM is characterized by a couple of non-negative integers $(n_u; n_r)$ such that:

$$n_u + n_r \leq n$$

where $n_u$ counts the number of nodes with urgent requests and $n_r$ counts those with regular requests. Indeed, the number of states in AM is $N_{AM} = (n + 1)(n + 2)/2$ which grows only quadratically with $n$. Compare this with $N_{EM} = 3^n$ in the Exact Model (exponential growth): the reduction in the number of states can be dramatic. AM introduces an equivalence relation on EM: EM states $i$ and $j$ are equivalent if $n_u(i) = n_u(j)$ and $n_r(i) = n_r(j)$, i.e., if they have the same number of nodes with urgent requests and the same number of nodes with regular requests. AM can therefore be viewed as the quotient space of EM with respect to the given equivalence relation. Loosely speaking, the states in AM can be viewed as clusters of states from EM. For instance, AM state $(2; 1)$ in a network with 10 nodes corresponds to the 360 EM states having exactly 2 urgent and 1 regular requests. Some good properties relate solutions of the Exact Model to solutions of the Aggregate Model. For instance, provided we are able to evaluate the expected cost of adopting a decision $\Delta$ in state $I$ of the aggregate model, we are able to compute the upper bound of the optimal $g^*$ by finding the optimal solution of AM that requires information only about a polynomial number of states. Moreover, if we are not able to efficiently compute the exact value of such a cost, but we might still be able to obtain a good approximation of it from historical records or by using subjective we may still obtain an upper bound. We also give some sufficient conditions under which the optimal EM policy is AM-constant, i.e., the number of regular requests satisfied by the decision induced by the policy is the same for all EM states $i$ belonging to the same AM-state $I$; and therefore the
optimal solution of AM, provides not only an upper bound, but also the true optimal value of $g^*$. 

4. Numerical Example

As an example, we consider the network depicted in Figure 1, consisting of a central depot $D$ and 5 possible clients, numbered $1, 2, \ldots, 5$. This example is a toy-problem and is not representative of the true complexity of real instances. However, the small dimensions of this example allow explicitly computing and comparing the optimal solutions for both the models (Exact and Aggregate). Note that, a network with just 20 nodes, not much bigger than this toy problem, will have $320 = 3,486,784,401$ different states, making it practically impossible to find the exact solution.

![Figure 1. Network example.](image)

Once we have evaluated all the costs, for each possible state and decision, and all transition probabilities, for each couple of states, we applied the iterative procedure recalled above and obtained the optimal strategy.

5. Results and Conclusions

On the provided numerical example, we showed that the cost of the optimal policy in the AM model is a tight upper bound for the cost of the EM optimal policy. Indeed the percentage deviation between the two average long-run delivery costs is 3.3%. A general correspondence exists between the optimal decisions suggested by the two models. Both models - EM and AM - have the same basic optimal decisions which differ in very few states. Obviously, in the aggregate model we do not use any piece of information on the configuration (location) of the nodes, thus making impossible to take advantage of those situations where serving an additional regular node does not procure any extra cost as it is in the exact model.

The two models can be combined in a hybrid procedure, where the first stage decision is optimally taken considering the information about the current state and
approximating future costs using those obtained from the Aggregate Model. Hence, in each day, after having observed the state in which the system is, we find the best decision to be taken that day, combining detailed exact information about the current day with aggregate information about the future. Notice that, we only need to solve the computationally easy AM problem once and then a simple problem every day.

The computational results confirm the viability of our approach.

As a final remark, let us note that the applicability of our approach could be even wider. Indeed, it could be used to compute the recourse action in an "a priori" optimisation model like the one developed by Jaillet [5].

References


CONFLICT RESOLUTION STRATEGIES IN REAL-TIME TRAIN TRAFFIC MANAGEMENT

Andrea D’ARIANO1, Dario PACCIARELLI2, Marco PRANZO3

Abstract. The train conflict resolution problem consists in rescheduling train operations in real-time, to face unexpected events that make infeasible the off-line generated timetable. In order to face the expected growth of transport demand in the next years, several new traffic control policies have been proposed and analyzed both to generate timetable and to effectively manage the traffic in real-time. In this paper, we use a detailed optimization model to analyze one such policy, called green wave, which consists in avoiding train speed profile modifications in open corridors. To this aim, we use a detailed model for conflict resolution, based on the alternative graph formulation, and a branch and bound algorithm for solving the problem at optimality. We compare optimal solutions computed when trains are allowed or not to change their speed profile in open corridors. A computational study on a small and complicated rail network has been carried out, based on a bottleneck area of the Dutch railway network.

1. Introduction

The problem faced by railway managers when rescheduling trains in real-time, is called conflict detection and resolution problem. As observed by Szigel [7], this problem is similar to the job shop scheduling problem. However, there are significant additional constraints. In this paper, we use the alternative graph formulation of Mascis and Pacciarelli [5] to carefully model the additional constraints of the problem, and to reschedule train movements optimally when timetable disruptions require to partially modifying the timetable.

The conflict resolution problem has been analyzed with various techniques. For example, Dorfman and Medanic [3] propose a discrete-event model for scheduling trains on a single line and a greedy strategy to obtain sub-optimal schedules. Dessouky et al. [2]

1 Delft University of Technology, Stevinweg 1, 2600G Delft, The Netherlands, a.dariano@tudelft.nl
2 Università degli Studi Roma Tre, Via della vasca navale 79, 00146 Roma, Italy, pacciarelli@dia.uniroma3.it
3 Università degli Studi di Siena, Via Roma 56, 53100 Siena, Italy, pranzo@dii.unisi.it
develop a MILP formulation and a branch and bound procedure for determining the optimal dispatching times for trains traveling in complex rail networks. The largest problem they can solve within 2 hours of computation is with 14 trains.

In this paper, we use a branch and bound procedure based on the alternative graph formulation to investigate the effectiveness of different railway traffic management strategies. We discuss in particular a strategy based on dwell time extensions and running time modifications, and the green wave strategy, which consists in letting the trains waiting only within the stations. With the green wave strategy a train, after leaving a station, should travel at the planned speed profile until the next station. In other words, while the general control strategy allows modifying both running times and dwell times, the green wave strategy permits only dwell time modifications, i.e. it permits solving train conflicts only at stations. We model the green wave restrictions as additional constraints in the alternative graph model of the problem (no-wait constraint). In our computational experiments, we compare the effects of two strategies in order to quantify the increase of trains delay.

The conflict resolution strategies are analyzed in terms of maximum secondary delay minimization, i.e. the additional delay caused by the resolution of conflicts. In order to make the comparison independent from the particular algorithm, we use the branch and bound procedure [1] to compare the optimal solutions to the conflict resolution problems, for different scenarios and for varying the time margins in the timetable.

A computational study on a small and complicated rail network has been carried out, based on the Schiphol bottleneck area of the Dutch railway network [4]. With specific reference to the location of waiting times, our tests show that the use of advanced systems for conflict resolution is important to obtain benefits in terms of delay minimization.

The paper is organized as follows. In Section 2 we introduce the notation and the alternative graph formulation and in Section 3 we illustrate the effects of different conflict resolutions strategies on a real test case.

2. Train scheduling formulation

In this section the mathematical notation used to model the train scheduling problem is introduced. In particular, we model the sequence of rail segments traversed by a train, called block sections, as a set of machines in a job shop scheduling problem, where trains correspond to jobs. The traversing of a block section by a train is called an operation $o$, and requires an occupation time $p$. Since at most one train at a time can traverse the same block section, the combinatorial structure of the train scheduling problem is similar to that of blocking job shop scheduling problem [6]. Blocking constraints represent the fact that a job, after finishing processing on a resource, cannot leave it until the subsequent resource becomes available for processing. Moreover, a train is not allowed to depart before its minimum departure time, and it is considered late if arriving at a station after its maximum arrival time. We define the delay of a train at a station as the difference between its actual arrival time, observed in real-time, and its maximum arrival time in the timetable.

In this paper, we address the real-time problem of facing unplanned events which make the planned timetable infeasible. In particular, we focus on the minimization of secondary delays, which are computed by subtracting to the train arrival times those delays that cannot be recovered by rescheduling trains movements. These are called primary delays, and are
obtained by computing the earliest possible arrival time of each train at each station/exit point, when the train travels at maximum speed and disregarding the presence of other trains. We aim at generating train schedules minimizing the maximum secondary delay of all trains.

In Figure 1, we show the alternative graph formulation for a small instance with three trains. At time $t=0$ there are three trains in the network. Trains A and B are passenger trains with a stop at the station. Train C is a freight train going from block section 5 to block section 8.

Figure 1: A railway network and the alternative graph formulation

At the bottom of Figure 1 the alternative graph for this example is reported. Each node of the alternative graph is denoted by the pair (train, block section). Each alternative pair of arcs is associated to the usage of a common block section. In particular, trains A and B share block sections 3, 4 and 8. Train C shares only the block section 8 with trains A and B. For each pair of trains sharing a block section, a sequencing decision must be taken. The pair of possible precedence between the two trains is called a pair of alternative arcs, and represents the two possible sequencing decisions to be taken on the precedence relation between the trains. Alternative arcs are depicted in dotted lines in figure. Arcs (0, A8) and (0, B8) represent the minimum departure time from the station for trains A and B, respectively. Finally, arcs (AStop, n), (A*, n), (BStop, n), (B*, n) and (C*, n) represent the modified due dates of the trains at the station and at the exit of the network, i.e. at the end of section 8. The length of all horizontal and positive arcs equals the occupation time of the trains on the different block sections. Arcs with negative weight between two adjacent operations of the same job represent the constraint that a train is not allowed to change its planned speed profile in a corridor (no-wait constraints). These arcs are removed when the green wave requirement is relaxed. The conflict resolution problem consists in choosing
one arc from each alternative pair, such that the resulting graph has no positive length cycles and such that the length of the longest path from 0 to $n$ is minimum.

3. Test case

In this section we use our branch and bound procedure [1] to solve the conflict resolution problem. Our experiments are carried out on the bottleneck area of the Dutch railways, shown in Figure 2, which includes the underground station of Schiphol, beneath the international airport of Amsterdam.

![Figure 2: Schiphol bottleneck area](image)

We study the network by using one hour of the provisional Dutch timetable for 2007 [4] with 54 running trains. We simulate different traffic conditions for ten timetable perturbations, obtained by letting trains enter the network late with respect to the timetable. The entrance delays are generated according to a Gaussian distribution and a random maximum delay varying from 300 and 1000 seconds. The number of delayed trains varies from 7 up to 27.

<table>
<thead>
<tr>
<th>Green Wave</th>
<th>Time Margins</th>
<th>Time</th>
<th>Max Out</th>
<th>Max Hfd</th>
<th>Max Shl</th>
<th>Avg Out</th>
<th>Avg Hfd</th>
<th>Avg Shl</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>Yes</td>
<td>5,41</td>
<td>62</td>
<td>22,5</td>
<td>76,5</td>
<td>3,76</td>
<td>2,74</td>
<td>4,40</td>
</tr>
<tr>
<td>No</td>
<td>Yes</td>
<td>1,25</td>
<td>62</td>
<td>22,5</td>
<td>76,5</td>
<td>3,36</td>
<td>2,61</td>
<td>3,83</td>
</tr>
<tr>
<td>Yes</td>
<td>No</td>
<td>5,38</td>
<td>100,7</td>
<td>87,6</td>
<td>96,4</td>
<td>15,29</td>
<td>11,62</td>
<td>10,00</td>
</tr>
<tr>
<td>No</td>
<td>No</td>
<td>1,26</td>
<td>99,4</td>
<td>84,9</td>
<td>95</td>
<td>14,21</td>
<td>10,54</td>
<td>8,78</td>
</tr>
</tbody>
</table>

Table1: Average results on 10 test cases

In Table 1 we show the comparisons between allowing or not trains to vary their running times in a corridor (Green Wave: Yes or No, respectively) according to the timetable with or without recovery and buffer times (Time Margins: Yes or No, respectively). On the overall set of 40 runs the branch and bound procedure is always able to attain an optimal
solution in few seconds. Each entry of the table reports the average results on 10 instances. Column Time indicates the computation time in seconds on a Pentium M 1.6 GHz processor. The value Max [respectively, Avg] is the maximum [average] secondary delay measured in seconds at platforms Hoofddorp (Hfd), Schiphol (Shl) and at exit stations (Out), respectively.

From the results in Table 1 we can draw several conclusions. Despite each instance is modelled with more than 8000 alternative pairs, the branch and bound algorithm is able to generate in few seconds the optimal solutions for all the instances. Moreover, the green wave strategy produces less efficient results in terms of delay minimization: even though the maximum secondary delay does not increase, the average secondary delay increases of about 10% with respect to the strategy without green wave. However, the availability of time margins is more critical from the viewpoint of delay minimization. The maximum secondary delay increases more than 30% with respect to the case with margins, and the increase is more than 240% for the average delay.

On the traffic control side, these results demonstrate the real possibility of automating the process of real-time railway traffic management, still commonly carried out by human operators. The branch and bound method can also be an efficient instrument to test different real-time control strategies.

References


Abstract. The main aim of this paper is to analyse the mobility system by taking into account the current issues and the possible tools which can be used to improve the efficiency of transport services and to recover the sustainable urban mobility culture. Therefore, this paper provides, on one side, a framework of the current situation of transport by highlighting the strategic role of new tools and, on the other side, detects the action lines which may be adopted in order to make the transport activity more sustainable.

1. The European framework

An explosive growth in urban population and in car use has been characterised the past fifty years. Nowadays many households have also moved to suburbs with the consequence of a greater car dependence, longer journeys, increased transport costs and higher environmental impacts. Urban decentralisation has greatly increased distances, which are not covered normally and feasibly with public transport. Current collective public transport systems have enormous difficulties in facing users’ needs, as collective transport better serve the urban areas rather than suburbs. So people are forced to use private modes, because there are not other possible means. Therefore, the further consequence is that current levels of mobility have negative effects on society, the economy and the environment.

Nevertheless the major part of EU citizens live in urban areas and benefit from the presence of public transport, although not completely adequate to users’ needs. This is due to the fact that there is a sort of mismatch between the current form of cities and users’ needs, on one side, and existing forms of public transport supply, on the other side.

Over the years there have been an extraordinary large number of analyses of the problem. To date no solution has been found, better: solutions have been suggested by the

1 Università degli studi di Bari – Dipartimento di Scienze economiche, Via Camillo Rosalba n. 53, 70124-Bari-Italy, e.venezia@dse.uniba.it
2 University of Hong Kong - Department of Architecture HKU, Pokfulam Road, Hong Kong, justyna@arch.hku.hk
literature but with scarce implementation by operators and politicians. Probably this is due to inadequate channels of communication between researchers, planning officials and policy makers. In this light, it is urgent to face the question on how to implement integrated strategies for sustainable development and sustainable urban mobility. We need to have in mind an urban transport system for the future to satisfy future needs for flexible passenger transport, respecting – at the same time – the urban environment. Although forms of public transport available today still reflect past needs, it is important to underline that nowadays changes in culture, social live, technology bring about changes in cities and in their transport system. Therefore, there is an urgent need to find solutions which are, on one side, environmentally acceptable and, on the other side, compatible with the current and future needs of the new structure of cities. In other words, as there is a strong link between public transport and sustainable development, what has to be done is to ensure people’s mobility in the long run without endangering the satisfaction of the future needs of today’s generations and of the coming one. In order to reach some results in terms of sustainable development, it is necessary to find a better balance between the individual mobility and public transport.

1.1. Public transport

Public transport is losing ground across Europe with the exception of cases where there have been strong investments and the transport service has been perceived by users as comfortable, reliable and safe. But, generally, in EU member states, people tend to include cars into their travel patterns and it will be difficult for them to switch back to public transport.

Throughout the European Union improvements have being undertaken which cover information provision, quality of service, planning priority for public transport, and land use planning within an integrated policy approach. Some of these improvements are the effects of changes recorded in the regulatory framework which, in turn, reflect options undertaken in different countries.

A recent report of Earchimede Consulting [1] shows some structural and performance indicators which currently characterise the public transport in some European countries, among which Italy is present. In particular, Italy has the lowest local-level railway transport rate in comparison with the other analysed countries. Another important aspect is given by the kilometres covered by public operators in comparison with the private ones in different countries. Further, the prevalence of public operators in Italy compared with the full deregulation existing in the UK. The performance indicators are even worse for Italy, as its last place finish in profitability proves: the gross operative profit on production value is equal to 6.5% against a European average of 10%. The other negative aspects are linked to the very high public funding per km and the level of traffic revenues. It is thus clear that a very critical situation exists: low service quality, very high labour cost and low average speed in urban areas. This last aspect is particularly penalising, on the one side, costs, and on the other side, the potential attraction of buses which is negatively affected by the absence of mobility policies. By augmenting bus-only lanes, for example, the commercial speed could be increased and this could attract a portion of bus demand currently unsatisfied.
Well-developed and efficient public transport systems can therefore play an important role in reducing the time spent on travel, thereby freeing up more time for citizens to devote to professional or leisure activities. However there still are a lot of improvements to be made, especially in the light of the fact that time spent on public transport is still perceived by passengers as lost, particularly during waiting times and at interchanges.

Finally, by considering the satisfaction degree of transport users, it emerges that nations where there is a strong public support also supply a worse qualitative service. Then we can conclude that countries characterised by a mature liberalisation process have the best performance both in economic and in qualitative (satisfaction) terms. We believe that in certain situations radical measures might be needed – although these require courage and ambition – to push a quick reform process of public transport. This brings to a general improvement of the service and better economic results for operators.

1.2. Tools of sustainable mobility

Demand management can make economies more effective, reducing environmental transport negative impacts, and improving quality of life. This implies the use of pricing, planning, market and political levers. So policy makers could put into force sustainable mobility policies. They could favour private collective transport, the increase of the occupancy rate of cars, the implementation of the shared property of cars, and so on.

At this point we can wonder if new alternatives to public transport can be implemented. We think that there is a sensitivity to sustainable development, but the start up seems to be quite slow. The reasons are linked to a question of style and behaviour of Italians, as well as of other Europeans, who have a strong “affection” to cars.

So far, although many of the main drivers of transport demand are not directly subject to transport policy control, this does not imply that transport policy tools are without impact. The practical evidence suggests that prices, speeds, and quality of transport alternatives offered, spatial planning and regulation are powerful instruments which have a large effect on the volume of travel, especially in combination, and when sustained for long enough to enable people and firms to modify their behaviour, which needs several years.

We believe that demand management is sometimes resisted as restrictive or unfair. But it is a necessary condition for improving mobility and environment. Failure to implement the opportunities for managing demand will undermine the value for money and effectiveness of infrastructure improvement, and lead to increasing congestion and environmental damage.

2. The Asian situation and the Hong Kong case

Urban infrastructure in Hong Kong, especially urban transport, has been added to cities over time as a major element in the promotion of urban efficiencies and safety, facilitating exchange of goods, and therefore improved city economics. Today, however, urban transport is increasingly seen as a threat to urban health, safety, economic efficiency and quality of life [2].
Hong Kong occupies a relatively small area of about 1100 square kilometers. With a population of 6.8 million, the average densities of Hong Kong are about 600 people per hectare. The real situation is, however, very different. With only 176 square kilometers occupied and the remainder unbuildable (as it is too steep or unstable) or left as country park, the local densities of Hong Kong are typically between 1000 to 5000 people per hectare. This puts Hong Kong at the top of the list of the densest places in the world. In short, Hong Kong is a dense and compact city; one of the very few examples of the truly compact city in the world. Its form, in theory should promote "walking city", or "public transport city", based on tram or railway system. Even if Hong Kong public transport is one of the most profitable in the world [3], current planning policies privilege roads and road dependent movement and support this with significant government subsidies. Car density on the road is already extremely high, one of the highest in the world. In December 2000 a total of 1904 km of roads were used by 580,000 licensed vehicles or 305 cars per kilometer of road. This number is achieved even though car ownership and car use in Hong Kong is very low, with only 49 cars per 1000 population. This is changing; recent trends indicate that the number of cars is rapidly increasing. If this statistic is considered along side the government remarks that lower residential densities are desirable for health reasons, significant concerns must be raised of the dangers this places on the compact city form as it is replaced by urban sprawl. Can reduced densities coexist with effective urban transport?

In order to answer this question let us first look at Hong Kong transport in greater detail. In Hong Kong 90% of the population rely on public transport. Average amount of journeys per resident annually is around 485. Compared with the same data for Atlanta, where the figure is only 11 passenger journeys per resident, per year, the figure is very high. The demand on public transport is obviously very high, which in turn makes it possible for the prices to be relatively low and services very frequent, compared with the rest of the world. All in all Hong Kong transport is efficient, affordable, accessible, and offers variety of choice to the public. There four major modes of transport: vehicle, rail, water transport and pedestrian movement.

All the above, will be discussed in greater detail in the following sections.

2.1. Road vehicle transport

The road public transport includes five privately owned franchise bus companies. The road public transport also includes mini buses and taxis. Hong Kong dependency on public transport one of the highest in the world.

The largest carrying mode of transport in Hong Kong is franchised buses. In 2002 franchised buses carried 4.33 million passengers a day.

Apart from franchised buses, which are mainly double-decker vehicles, capable of carrying substantial number of passengers, there are also two types of minibuses. The green minibuses have exclusive rights on fixed routes, which are determined by Transport Department. The red minibuses are free to operate nearly anywhere within Hong Kong with exception to the areas where special prohibitions apply. The number of minibuses is lower than franchise buses. In 2001 there were 2360 green minibuses and 1983 red minibuses on the roads, offering services to nearly all parts of Hong Kong. Thus, there are a large number of bus networks operating with larger buses on major routes supplemented by smaller buses servicing a wide network thus reaching most parts of the territory.
In addition to these public systems, many schools, hotels, hospitals, clinics, and some residential developments have their own bus services that operate under strict restrictions and regulations. They provide very desirable services and contribute substantially to reduction in traffic congestion.

10% of public transport journeys are journeys by taxis. Taxis are plentiful in Hong Kong and they are available everywhere. Taxis are legislated by the regions in which they may operate, namely urban, rural and island, each differentiated by colour; there are red, green and blue taxis. In 2001 in total there were 18,130 taxis operating in Hong Kong, approximately three taxis per 1000 people.

The car ownership in Hong Kong is comparatively low for such a wealthy community with only 49 cars per 1000 people. In the early 1970s both Hong Kong and Singapore, in response to rise in car ownership and an upsurge in traffic, introduced methods of restraining car ownership. The policy of making driving expensive has worked well in Hong Kong. High taxes, related to car ownership, high petrol prices, limited and expensive parking facilities, all contributed to keeping car ownership to at very low levels. This in turn leads to a very heavy demand for public transport facilities.

2.2. Rail transport

Rail is the second most popular form of transport. In 2002 railways carried 3.52 million passengers per day. Railways are the most environmentally friendly form of mass transport.

Urban rail system in Hong Kong is one of the most heavily use in the world. By 1993, 50% of the entire population of Hong Kong lived within 500m of Mass Transit Railway. KCRC, with recently extended lines system provides extensive service for population living in New Territories.

The other mode of light rail transport in Hong Kong is tram. Hong Kong Tramways Limited, a member of the Wharf Group of companies, has been operating tram service in Hong Kong since 1904. Today there are six tram routes (totally 30km). The Hongkong Tramways Limited provides alternative, cheap service for those who are not in a hurry. It is very popular with the tourists and elderly.

The Peak Tram Hong Kong's famous funicular railway is the city's most popular tourist attraction (http://www.thepeak.com.hk/tram/location.html).

2.3. Water transport

Water transport is always attractive to its users; however, in most cities around the world, it often proves to be ineffective. The role of ferries in Hong Kong has been shrinking drastically ever since the opening of the first cross harbour tunnel. From 1972 to 2002 the lost of market share went down from 13.9% to 1.3%; and from 649000 passenger trip daily to 150000. There are two major ferry companies: Star Ferry and New World First Ferry Ltd. Star Ferry provide cross harbour service only and it is very popular with tourists as well as people traveling in between Hong Kong Island and Kowloon Peninsula. It is very pleasant, enjoyable and attractive way to travel; an alternative to traveling underground.
2.4. Pedestrians

A lot emphasis has been given to separate vehicle and pedestrian movement. Great amount of elevated walking system has been provided. Without them the number of pedestrian on the street level would have been impossible to accommodate. By providing reasonable fast and safe access these elevated walkway system definitely improve pedestrian movement, but they are only addressing one problem, which is to move people fast. These pedestrian motorways are similar to fast roads, or motorways and they have very little to do with urban structure, city live, or civic spaces.

But there are some successful and interesting examples of pedestrian infrastructure as well. One of them is the Central escalator. The main aim of the escalator was to reduce vehicle traffic movements. When it was initially built it was expected to carry only about 30,000 people every day. In April 2004, it was reported that 350,000 people use the escalators every day. The success of the escalator led to creation of another similar project. There are more proposals for escalators around Hong Kong, if realized these project will substantially reduce the traffic other areas.

3. Conclusions

Should Europeans learn from Hong Kong experience? We think yes. Nevertheless, firstly, political courage is needed in the development of urban transport policies in European countries. Further, a concerted effort is required to manage the growth of motor vehicle use and encourage use of public transport systems or alternative tools. For example car restraint measures, including pedestrianisation of city centers, also contribute to reducing traffic congestion, especially when coupled with parking management initiatives. Through comprehensive policy packages which include effective pricing and communication measure, urban transport system can follow a sustainable course of development. So the Hong Kong case can be a benchmark example because there public transport functions well, providing an efficient, reliable, frequent, and affordable service. It offers choice in terms of modes of transport and in terms of how much money one is prepare to spend. However the question remains if the existing infrastructure can cope with forever increasing demand.

References


[1] EARCHIMEDE, STRATEGY CONSULTANTS (2005), La resa dei conti, Rapporto sul trasporto pubblico locale: situazione attuale e prospettive evolutive.

AUCTION BASED CONGESTION PRICING: THE BASIC IDEA

Dušan TEODOROVIĆ 1,2, Konstantinos TRIANTIS 2, Yueqin Zhao 2, Snežana MLADENOVIĆ 2

Abstract. Planners, engineers, and economists have introduced various demand management methods in an attempt to reduce the fast growing traffic congestion. Demand management strategies force transportation network users to travel and use transportation facilities more during off-peak hours. In this paper, a new demand management concept - Auctions Based Congestion Pricing is proposed and modeled.

1. Introduction

Planners, engineers, and economists have introduced various demand management methods in an attempt to reduce the fast growing traffic congestion (park-and-ride, high occupancy vehicle lanes (HOV), HOT lanes, ridesharing and transit use, road pricing, parking pricing). Transportation demand management is a common term for various activities that advocate a decrease in the demand for existing transportation systems. Demand management actions also force transportation network users to travel and use transportation facilities more during off-peak hours. Some of the demand management strategies could advance the transportation choices accessible to users. Some other demand management strategies generate changes in departure time, route choice, destination or mode choice. Successfully planned and implemented demand management strategies can result in significant toll revenues, decrease in total number of vehicle trips, decrease in total number of vehicle trips during peak periods, increase in the number of vehicle trips during off-peak periods, increase in ridesharing, rise in public transit ridership, and in some cases increased cycling, walking, and tele work. It should be noted that some of the demand management strategies have been already successfully implemented. In this paper, a new demand management concept Auctions Based Congestion Pricing is proposed and modeled. The paper is organized in the following way. Section 2 explains the basic principles of auctions, Section

1 Faculty of Transport and Traffic Engineering, University of Belgrade, Vojvode Stepe 305, 11000 Belgrade, Serbia, duteodor@vt.edu
2Virginia Polytechnic Institute and State University, 7054 Haycock Road, VA 22043, USA, triantis@vt.edu

688
3 presents the model of the problem considered, and Sections 4 and 5 the proposed solution approaches and conclusions respectively.

2. Auctions

The Auction represents a market-based procedure in which an item or a collection of items is sold on the basis of bids (prices offered by the auction participants for the item being auctioned). The auction participants are also called bidders, or agents. Post stamps, old coins, paintings, old automobiles are typical items that are frequently auctioned. Recently, many services started also to be auctioned. For example, the Federal Communications Commission (FCC) created great application of Game Theory in 1996. The auction participants played 112 rounds with open bids while competing for rights to operate wireless personal communications services. Nevertheless, depending on the items being auctioned auctions could be performed online or offline. Most frequently auctions are conducted with respect to one item (service). In this case, there are multiple auction participants (buyers) trying to buy the auctioned item. In the case of so-called reverse auction there are multiple sellers (manufacturers, operators) trying to sell the product (service) to one buyer. The English auction is a sequential auction in which bidders try to beat the current bid. Any new bid must increase the current bid by a previously defined increase. The English auction finishes when no bidder is ready to beat the last offered bid. The winner of the English auction is the bidder who offered the last bid. He/she pays the amount bid.

On the other hand, Dutch auction starts with the initial price much higher than any bidder is willing to pay. The price is decreased little by little until some of the bidders express his/her willingness to pay. The auction is then finished and the winner pays the amount he/she indicated. The price never goes below the minimum price set up by the auctioneer. This auction type is a widespread market mechanism for selling flowers in Holland.

There are a few other forms of auctions. First price auction is an auction where the winner is the auction participant who submits the highest bid. He/she pays the amount bid. In a Vickrey auction, all auction participants submit bids simultaneously. The highest bidder is the winner. In the end, the winner pays the second-highest bid offered. Finally, in combinatorial auctions bidders compete to buy many different but related items. Every auction participant makes one or more bids for any of the possible item combinations. If not all requested resources are assigned to bidder, the partial acquisition of the resources has much less value for the bidder.

The winner determination problem in the case of combinatorial auctions could be defined in the following way: Identify the winning set of bids that maximizes revenue of the auctioneer. It has been shown that winner determination problem is a NP-hard problem. In reality, auction participants make bids only for limited number of item combinations. Auctioneer frequently prescriptions maximum number of bids per auction participant in order to decrease problem dimensions. Combinatorial auction mechanisms have been proposed as a convenient tool for allocation of resources in public services (electromagnetic spectra, time slots for landing at airports).
3. **Auction based congestion pricing: Downtown time slot auction**

The basic ideas of the Auction Based Congestion Pricing concept is that all drivers that want to enter downtown (Figure 1) have to *participate in an auction*. The operator (traffic authority) is the auctioneer who makes the decision whether to accept or reject particular bids sent by the drivers.

![Figure 1: Drivers’ Requests.](image)

Let us introduce the following notation:

- \( i \) - agent index
- \( j \) - bid index
- \( I \) - total number of agents (bidders)
- \( J(i) \) – the total number of bids made by the agent \( a_i \)

Let us divide the considered time interval (for example one week) into smaller time intervals (the width of small time interval could be for example 15 minutes). We denote by \( T \) the total number of small time intervals. We use small time intervals to measure the duration of the visit to downtown. For example, vehicle stays in downtown three time intervals, five time intervals, eight time intervals, etc. Driver (Agent \( a_i \)) makes one or more bids for one or more down town visits during observed time period (one week) (see Figure 2).

![Collection of visits](image)

**Figure 2. Collection of visits.**
We call the set of requested time intervals - *collection of visits*. Let us denote by $c_{ij}$ the $j$-th collection of visits created by the agent $a_i$. The agent $a_i$ offers to the auctioneer (operator of the system) $m_{ij}$ monetary units for right to stay in downtown during time intervals defined in the collection of visits $c_{ij}$.

Let us note small time interval $t$. We define parameters $d_{ij}(t)$ in the following way:

$$d_{ij}(t) = \begin{cases} 
1, & \text{if collection } c_{ij} \text{ created by the agent } a_i \text{ contains small time interval } t \\
0, & \text{otherwise} 
\end{cases}$$

(1)

Let us introduce variables $x_{ij}$ defined in the following way:

$$x_{ij}(t) = \begin{cases} 
1, & \text{if the auctioneer accepts collection } c_{ij} \text{ requested by the agent } a_i \\
0, & \text{otherwise} 
\end{cases}$$

(2)

The Winner Selection Problem (WSP) in the case of downtown time slot auctions could be defined in the following way: Discover the best set of collection of visits that should be accepted by auctioneer. The WSP is expressed mathematically in the following way:

Maximize

$$R = \sum_{i=1}^{I} \sum_{j=1}^{J(i)} m_{ij} x_{ij}$$

(3)

subject to:

$$\sum_{i=1}^{I} \sum_{j=1}^{J(i)} d_{ij}(t)x_{ij} \leq D_{\text{max}}, \quad t = 1, 2, ..., T$$

(4)

$$\sum_{j=1}^{J(i)} x_{ij} \leq 1, \quad i = 1, 2, ..., I$$

(5)

Objective function represents the total revenue $R$ of the auctioneer that should be maximized. We denote by $D_{\text{max}}$ the maximum allowed number of cars that could be present in downtown at any time point in time. The total number of cars that are present in a downtown area during small time interval $t$ equals $\sum_{i=1}^{I} \sum_{j=1}^{J(i)} d_{ij}(t)x_{ij}$. This number must be less than or equal to the maximum allowed number of cars in downtown $D_{\text{max}}$ (equation (4)). Constraint (3) indicates that maximum one bid of any agent that could be accepted. The WSP is a difficult combinatorial optimization problem.

The model described by the relations (3)-(5) is related to the *One-Shot Auction*. *Multiple Round Auctions* are the auctions in which bidders can bid for combinations of items and modify their bids in response to bids from the other agents. Multiple Round Auction is by its nature an *Iterative Combinatorial Auction*. Iterative Combinatorial Auction would enable drivers to avoid peak hours, switch to other time slots, and pay smaller amount for right to enter downtown, or change the mode of transportation, based on the information they got from the previous round. In the case of a One-Shot Auction drivers would not have a chance to change their decisions despite the fact that some of the available time slots could be “acceptable” for some of the drivers. Costs, implementation
issues, and potential benefits of the One-Shot Auction versus Multiple Round Auction will be highlighted in more details in future research.

4. Solving the downtown time slot auction problem

4.1. A “Greedy” heuristic algorithm

“Greedy” heuristic algorithms build the solution of the studied problem in a step-by-step procedure. In every step of the procedure the value is assigned to one of the variables in order to maximally improve the objective function value. In every step, the greedy algorithm is looking for the best current solution with no consideration of future costs or consequences. Greedy algorithms use local information available in every step. In the problem considered, a simple greedy algorithm could be formulated in the following way. The total number of time intervals $D_{ij}$ in a collection of visits $c_{ij}$ created by the agent $a_i$ equals:

$$D_{ij} = \sum_{t=1}^{r} d_{ij}(t)$$  

(6)

The ratio $a_{ij}$ between the number of monetary units $m_{ij}$ and the total number of requested time intervals $D_{ij}$ equals:

$$a_{ij} = \frac{m_{ij}}{D_{ij}}$$  

(7)

The ratio $a_{ij}$ represents the offered unit price per one requested time interval. The list of calculated ratios $a_{ij}$ is created and sorted in a descending order. Allocation of available time slots could be done according to the formed list. Available time slots must be updated all the time.

4.2. Metaheuristics approaches

Metaheuristics can be defined as general combinatorial optimization techniques. These techniques are designed to solve many different combinatorial optimization problems. The developed metaheuristics are based on local search techniques (Simulated Annealing, Tabu Search, etc.), or on population search techniques (Genetic Algorithms). The problem considered in this paper will be solved by various metaheuristics approaches.

5. Conclusion

The new demand management concept - Auctions Based Congestion Pricing is proposed and modeled in the paper. The basic principles of auctions are described in the paper. The
downtown time slot auction model is proposed, as well as potential solution approaches. The most important issues to be addressed in future research are one-shot auction versus multiple round auction, payment determination (first price versus second price), information flows between bidders and the auctioneer, bid language, public acceptance, etc.

References


DEALING WITH UNCERTAINTIES IN TRANSPORT POLICYMAKING: A NEW PARADIGM AND APPROACH

J.W.G.M. van der Pas¹, D.B. Agusdinata, W.E. Walker, and V.A.W.J. Marchau

Abstract. Intelligent Transportation Systems (ITS) seem highly promising in reaching policy goals. Policymakers, however, are confronted with major uncertainties regarding ITS implementation (e.g. future transportation demand, ITS costs and benefits, technological developments, etc.) Current approaches are too limited to handle these uncertainties in a proper way. In this paper innovative approaches for handling uncertainties regarding ITS implementation are presented and applied to specific ITS examples.

1. Introduction

Policymaking is about the future. As the future is uncertain, transport policymakers must take uncertainty into account in their policymaking. A scan of the proceedings of the 13th mini-Euro conference “Handling Uncertainty in Transportation Systems” [1] provides a variety of examples of uncertainties involved in transport policymaking (e.g., uncertainties about future transportation demand, modal choices of transport users, and route choice behavior). But an in-depth analysis shows that these examples are quite narrow and usually apply to only one type of uncertainty or one type of uncertainty at a time. This limited handling of uncertainty might lead to policy failure because not all aspects of uncertainty have been considered.

In this paper we establish an overview of how policymakers have dealt with these uncertainties in the past and suggest a new paradigm and approaches for dealing with uncertainties in policymaking. We focus on uncertainties in the transportation domain and on the case of Intelligent Speed Adaptation (ISA) systems in vehicles in particular.

¹ Address correspondence to: Delft University of Technology, Faculty of Technology, Policy and Management. E-mail: J.vanderpas@tbd.tudelft.nl. Address: Jaffalaan 5, 2628 BX Delft, Tel.: +31 15 2783479, Netherlands
ISA systems can be described as systems that “know” the prevailing speed limit on a road and are capable of using that information to inform the driver when the speed limit is being exceeded or even limiting the vehicle’s speed to comply with the speed limit [2, 3]. Several studies have shown that ISA has a high potential to decrease the number of road casualties and fatalities [3]. However, from the perspective of implementing ISA in the real world this potential becomes questionable. We chose the case of ISA for several reasons. First it is a typical case of using technology as solution for a policy problem as to an alternative for traditional measures such as more roads, stricter police control. Second it involves multiple uncertainties (e.g. cause of implementation, technical reliability, user acceptance). Therefore, the insights from ISA implementation can also be useful for other types of transport policies.

In Sec. 2, we introduce a systems view of policymaking and a categorization of uncertainty, and discuss how uncertainties were dealt with in the past in transport policymaking. Sec. 3 describes three new methods for dealing with uncertainties involved in policymaking for ISA and our research into dealing with ISA implementation uncertainties using these methods.

2. **Introducing a system analytic view on policymaking**

In order to categorize the various types of uncertainty, we first introduce an analytical framework called the policy analysis framework [4]. Policymaking requires an integrated systems view of the various factors influencing the performance of a system, their consequences for system performance, and a way of valuing these consequences in order to choose a policy to implement. In other words, policymaking concerns making choices regarding a system in order to change the system outcomes in a desired way (see Figure 1). The elements in this framework can be assembled in a structure labeled ‘XPIORV’, where:

- **X** = External forces: factors that are beyond the influence of policymakers in the particular policy domain (e.g., economic factors in the transportation policy domain).
- **P** = Policies: instruments that are used by policymakers to influence the behavior of the system in order to help achieve their objectives (e.g. road pricing, built new roads, fuel taxes).
- **I** = Internal factors: factors inside the system that are influenced by the external forces and policies (e.g. vehicle use, throughput, driving behavior).
- **O** = Outcomes of interest: measurable system outcomes that are related to the objectives of policymakers and stakeholders (e.g. traffic safety, congestion, emissions and fuel use).
- **R** = Relationships: the functional, behavioral, or causal linkages among the external forces, policies, and internal factors that produce the outcomes of interest. Hence; **O** = **R**(X, I, P) (e.g. the relationship between speed and fatalities)
- **V** = Valuation of outcomes of interest by policymakers and stakeholders (reflecting their goals, objectives, and preferences).
Given the policy analysis framework, there are different locations in the model-based policy analysis process where uncertainty can manifest itself [5]. First, there are uncertainties related to the external forces driving changes in the system (e.g. future economic growth and technological development). Next, there are uncertainties about the system model itself called system model uncertainty. System model uncertainty refers to the selection of the relevant system components and the identification of the relationships among these components (e.g., the functional form of the relationships). Both the uncertainties about external forces and the system model lead to uncertainties about the estimated outcomes. In addition, there are great uncertainties about the current and future valuation of the outcomes by the stakeholders [5, 6].

2.2 Policymaking approaches for dealing with uncertainty

In general, we can distinguish three broad categories of dealing with uncertainty (see also Table 1):

- **Ignore (passive) or reduce (active) uncertainties.** This category involves a broad variety of approaches, varying from passively ignoring the uncertainty or to investing in research and trying to reduce the uncertainty [7, 8]. The basic assumption of this approach is that the future will more or less the same as the present.

- **Robust or “what-if”**. In the 1950s, RAND developed scenario techniques allowing policymakers to do “what if” analysis. Different scenarios for the future are sketched and the outcomes are quantified with the help of models. Decision-makers use these scenarios to “probe for weaknesses in proposed plans”[9]. The central assumption of
this paradigm is that the future can be predicted well enough to identify policies that will produce favorable outcomes in one or more specific plausible future worlds. [10]

The first and the second approach mentioned above are commonly used in transportation policymaking. The last one however is more scarcely used, and is associated with a new paradigm:

- **Adaptive approaches, building in flexible options.** The underlying assumption in these approaches is that the future cannot be predicted. Adaptive approaches allow policymakers to create policies that learn over time, as the uncertainties about the future are reduced.

<table>
<thead>
<tr>
<th>Policymaking approach</th>
<th>Basic assumption</th>
<th>Analytical methods (not exhaustive)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ignoring (passive) and/or reducing (active)</td>
<td>Future will be the same (or an extrapolation) of the present</td>
<td>Impact assessment, Multi-criteria analysis/Cost benefit analysis, Preference modelling</td>
</tr>
<tr>
<td>Robust, what-if</td>
<td>Future can be predicted well-enough to identify specific plausible future worlds</td>
<td>Scenario analysis, scenario building, Backcasting, Bouncecasting, Exploratory modeling</td>
</tr>
<tr>
<td>Adaptive strategies</td>
<td>Future cannot predicted</td>
<td>Real options, Exploratory modeling</td>
</tr>
</tbody>
</table>

Table 1: Paradigms, approaches and analytical methods for dealing with uncertainty

### 2.3. Analytical approaches used in transport policy research

Until now, uncertainties in transport policymaking have been primarily dealt with by using traditional methods. We reach this conclusion by scanning the proceedings of the 13th Mini-Euro Conference [1] and by surveying the ISA literature which will be described below.

#### 2.3.1 13th Mini-Euro Conference

In depth analysis of the uncertainties that are addressed in the conference proceedings of the 13th Mini-Euro Conference [1] reveals a focus on model uncertainty. A lot of research addresses uncertainties involved in modeling road accidents, transportation demand (long and short-term), transportation planning problems, route choice behavior (or finding optimal routes in networks), modal choice behavior, and travel time prediction. Most of the efforts are focused on improving models to better reflect reality. The main focus of research is on reducing these uncertainties using fuzzy logic, neural networks, and other modeling techniques. This approach fits in the category of “reducing” uncertainty; the focus is on making better models by doing more research and using new and different modeling techniques. Uncertainties regarding the external forces (or model inputs) are often addressed by scenario analysis. A number studies in the Bari proceedings (2002) looked at uncertainties in model outputs that are used as inputs to other transport models (e.g.,
uncertainties in the output of demographic models that are used as input to travel demand models, or uncertainties in the output of land-use models that are used as input for travel demand models).

2.3.2. ISA literature review

In most ISA studies, external variables (X) are not taken into account at all or only to a limited extent, implying outcomes under the assumption that for instance future transport demand will not change or follow historical trends. Next, most studies use a too simplistic model structure (I, R) in their analysis for describing the relationship between speed and traffic safety. For example Carsten & Fowkes [11] use a single initial speed profile (i.e. normal). Ma & Andreasson [12] use a single distribution for driver reaction time and a single functional relationship between probability of death and collision speed. It has however been shown that the exact relationship between speed and traffic safety is uncertain, as this depends on various factors (e.g. lane-width, traffic density, junction density) [13], the way in which crucial stakeholders for ISA implementation will value outcomes of interest is scarcely studied [14], this is important for policymakers to know because it influences the acceptance of their policies and crucial stakeholders might be a threat or opportunity for the success of their policies.

Hence, traditional approaches seem inappropriate to handle uncertainties associated with ISA implementation. Therefore, recently, we started to (further) develop and apply innovative methods in the field of adaptive policymaking.

3. Dealing with uncertainties in the case of ISA implementation

We now describe the results of three different methods we have applied for dealing with uncertainties in ISA implementation. Because of space limitations, we restrict our discussion to a brief explanation of each method, the location in the policy analysis framework (figure 1) of the uncertainty it addresses, and initial findings for the case of ISA implementation.

3.1 Bouncecasting for dealing with uncertainties in the implementation of ISA

Bouncecasting is an approach developed by RAND that combines two methods to dealing with uncertainty -- forecasting and backcasting -- in order to obtain the benefits of both [2]. First, several scenarios are developed, then a game is played involving policymakers and all relevant stakeholders. The aim of the game is to have each stakeholder explore how to exploit the advantages and avoid the disadvantages of each of the scenarios. This process is iterative and the scenarios are adjusted in between.

In 2005, a bouncecasting exercise was performed for dealing with the uncertainties in ISA implementation [15]. Students of the Delft University of Technology performed this exercise, producing results regarding both the bouncecasting methodology and ISA
implementation. The bouncecasting exercise allows policymakers to deal with many different locations of uncertainty in one exercise. Our initial experience suggests that bouncecasting is a promising approach. It provides policymakers with insights into the perceptions of other (crucial) stakeholders, uncertainties in and relevance of external influences, advantages and disadvantages of different states of the future policy domain, and threats and opportunities in the world of today. It is found out for example that the issue of liability (i.e. who pays when accident happens) dominates the concern of key stakeholders. To minimize this threat, a scheme of risk sharing among drivers, manufactures, and insurance companies has been proposed as one potential solution.

3.2 Adaptive policymaking for dealing with uncertainties in the implementation of ISA

The adaptive approach [16] departs from the ‘predict then act’ approach in which all information needed for policymaking must first be predicted, often using best guess estimates of scenarios and states of the system. The normal practice is that the best estimates of each policymaking support activity are used as inputs for designing and choosing a policy. In contrast, an adaptive approach does not require that all uncertainties are resolved before the policy can be implemented. In fact, it takes relevant conflicting results of policymaking support activities as its inputs. It allows policymakers to deal with multiple (massive) uncertainties.

In the adaptive approach, initial policy actions are chosen based on overall objectives. The policy is protected against uncertainties or vulnerabilities with mitigating and hedging actions at the time of initial implementation and by corrective and defensive actions as knowledge is gained during implementation. Adaptive policies include a method of dealing with uncertainties through use of monitoring, whereby signals are given about the potential failure of the policy that trigger implementation of corrective or defensive actions. The approach seems most useful when there are multiple uncertainties involved and the uncertainties are very large, both of which apply to the case of ISA implementation.

Initial results are promising. As the success of ISA implementation depends on user acceptance, for instance, an option of scaling down when acceptance dwindles should be considered as a corrective action. The decrease of acceptance can be known by monitoring an increase in traffic mean speed.

3.3 Exploratory modeling for dealing with uncertainties in the implementation of ISA

Exploratory modeling (EM) is a tool developed by Bankes [9]. It uses computational experiments to help policymakers reason about situations involving massive uncertainties [17, 18]. EM enables the handling of uncertainties with respect to external variables (X), the state of the system (I, R), and the valuation system (V). It can, therefore, be considered as sensitivity analysis in a broad and deep sense. Some of the insights we have gained by using EM are described below:
• *Pattern of system performance*: what factors matters in which circumstances.

• *Performance robustness*: insights into the boundary between policy success and policy failure (i.e., factors that can cause policies to fail).

• *Change in mental models*: a shift from optimization to robustness as criteria policy choice criterion requires a change in the way policymakers perceive what works, what does not, and by what mechanisms (i.e. their mental models).

For instance, it is shown that an ISA policy targeting both young drivers (with speeding behavior) and older driver (with normally distributed speed profile), has to have a minimum 50 % acceptance level to succeed. Having this critical level, a threshold value of 0.2 km/hr of increase in mean traffic speed (i.e. a sign of dwindling acceptance) has been established.

These improvements in policymaking must be balanced by cost or resources needed to perform EM. Computing time is no longer an issue. The large numbers of computer runs needed by EM can be performed quickly. The challenge is to identify patterns of performance in the output from these runs. This might take considerable efforts and resources (e.g. in developing a pattern search algorithm).

4 Conclusion

Transportation/ ISA implementation uncertainties are usually dealt with in traditional ways. This paper presents new approaches to dealing with uncertainty that view uncertainty as an opportunity instead of a threat. Bouncecasting, exploratory modeling, and adaptive policymaking seem to be promising approaches to support policymakers implementing policies in the face of massive uncertainties.

Acknowledgements: This research is supported by the Next Generation Infrastructures Foundation.

References


Providing Public Participation in Transportation Decision Using SPAN (Social Participatory Allocative Networks) Method

Murat Karacasu 1 Safak Bilgiç 1

Abstract. Transportation has been a very important subject for people for a long time. People have used various transportation systems for their transportation necessity in their daily life. Some problems have occurred in urban transportation in the course of time. Public transportation is generally supplied by local authorities in cities. In many developing countries local authorities are insufficient about public transportation. Local authorities try to find a solution for transportation problems. But, generally local authorities make decisions by themselves. Whereas, public participation should be supplied in decisions which are related to transportation. This is necessity for democratic life. We emphasize public participation in public transportation. Also decisions should be supported by models. Public participation was evaluated by SPAN (Social Participatory Allocative Networks) Method. Data which are gotten from privatization public survey (performed in Eskisehir, Turkey) were used for model.

1. Introduction

Transportation investments contain big scope and huge cost. Therefore, any decision made about transportation is very important for public. They include economic, environmental, social, political, and technological aspects. Public transportation has been generally dealt by local authorities. However, while performing these duties, local authorities face some difficulties because of various reasons namely financial constraints. The Transportation Equity Act (TEA-21) formally requires public involvement in transportation planning from all state Departments of Transportation. Furthermore, TEA-21 requires not only disseminating information to the public, but also soliciting and considering public opinion in forming transportation policy [1]. Decision making is to select the one of alternatives under some constraints [2]. Resources which are used for

1 Eskisehir Osmangazi University, Eskisehir, Turkey, muratk@ogu.edu.tr
investment are limited. Therefore managers should be careful about decision making. Also public oversight should be considered in decision making process. A voting about public participation was practiced in Edinburgh like London. The residents of Edinburgh have voted against the introduction of congestion charging in the city by a margin of about three to one. A set of policy agencies, transportation experts and green advocacy groups were consulted in this road pricing process [3]. Soh and Yuen’s paper explores various dimensions of public participation in urban planning in Singapore: the changing nature and attitudes towards citizen participation, the players and methods of participation, and the means of evaluating their impact [4]. Davidoff and others emphasize the need for public participation in the planning process [5, 6, 7]. Also Bickerstaff, Tolley, Walker point out that need of public participation in point of inclusivity, transparency, interactivity and continuity [8].

On the other hand, privatization is a transportation planning decision and transportation investment. As in some transportation sectors privatization is thought to be a solution for the problems. The privatization about public transportation in some countries succeeded; however, in some others expected results were not achieved. Local authorities, seeing these unsuccessful results, have been trying to re-gain the task for the privatization [9]. In 1980s important developments were practiced in Britain [10]. It was promoted by the government as a way of introducing competition, improving efficiency, providing innovations, increasing user choice and reducing subsidy requirements [11]. Britain has gained by far the most extensive experience of bus and coach deregulation and privatization within Europe. This experience has been mixed, with both positive and negative outcomes in different respects [12, 13].

In this study, aim is supplying public participation for decisions about public transportation. For this aim, a public survey was realized in Eskisehir, Turkey.

2. Model and Application

2.1. Data

There are four groups (experts, users, administrators and civil organizations) for public participation. Academicians who are related to transportation and specialist which work at various organizations participated to survey as expert. Academicians were selected from all universities in Turkey. The number of experts is 53. People who live in Eskisehir took part in survey as public. 830 people and students answered the question in their houses and at bus stops. Civil organizations (CO) were made up of representatives from 8 organizations which related to transportation one by one. Administrators who are employed by municipality and private firms from Eskisehir and other cities were chosen as transportation administrator. Alternatives consist of publicly owned bus \((A_1)\) and privately owned bus \((A_2)\). There are two different types of public bus system, one run by municipal authorities and one run by private agencies. The data were collected from four groups, public, transportation administration, civil organizations and experts in Turkey. Data collected from these groups were provided as rational numbers based on the stated preferences. There are 33 criterions in the data set. Public, transportation administration, civil organizations
have 14, 13 and 6 criterions respectively. Public survey was carried out as two steps. In first step, each group determined criterions which are important for them about public transportation. That is, each group answered the questions which are related to them. This case is very important for reliability of public survey. For example, public is interested in comfort or ticket fee, not fuel cost. Thus, criterions of each group are different. But experts concern all criterions because of their experience backlog. In second step, participators answered the survey questions which are connected with them. Criterions are given for each group in Table 1.

Survey composes of three parts. Firstly, participator determined weights of criterions than preference ratio for publicly owned bus and privately owned bus. In third step, participator gave participate ratio each other. The values of weight refer importance degree of criterion. Criterions were determined by each group one by one. Than, participator made a choice between publicly owned bus (A₁) and privately owned bus (A₂) for each criterion. They gave a point to alternatives. Participator shared out 100 points to two alternatives. In last step participator determined importance level of their decision.

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Public (n=14)</th>
<th>Transportation administration (n=13)</th>
<th>Civil organizations (n=6)</th>
<th>Experts (n=33)</th>
</tr>
</thead>
<tbody>
<tr>
<td>p₁</td>
<td>ticket fee</td>
<td>fuel cost</td>
<td>influence to car usage</td>
<td>all criteria</td>
</tr>
<tr>
<td>p₂</td>
<td>comfort</td>
<td>maintenance cost</td>
<td>influence to public</td>
<td>(14+13+6)</td>
</tr>
<tr>
<td>p₃</td>
<td>vehicle cleanness</td>
<td>personal cost</td>
<td>public pleasure</td>
<td></td>
</tr>
<tr>
<td>p₄</td>
<td>payment type</td>
<td>amortization</td>
<td>environmental effect</td>
<td></td>
</tr>
<tr>
<td>p₅</td>
<td>fullness</td>
<td>accident cost</td>
<td>easiness for disabled</td>
<td></td>
</tr>
<tr>
<td>p₆</td>
<td>sitting probability</td>
<td>insurance cost</td>
<td>benefit for country economy</td>
<td></td>
</tr>
<tr>
<td>p₇</td>
<td>personal behavior</td>
<td>benefit ratio</td>
<td></td>
<td></td>
</tr>
<tr>
<td>p₈</td>
<td>standards of vehicles</td>
<td>fleet productivity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>p₉</td>
<td>stop conditions</td>
<td>decision elasticity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>p₁₀</td>
<td>service reliability</td>
<td>future investment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>p₁₁</td>
<td>time schedule</td>
<td>standards of vehicles</td>
<td></td>
<td></td>
</tr>
<tr>
<td>p₁₂</td>
<td>service frequency</td>
<td>elasticity in vehicle choice</td>
<td></td>
<td></td>
</tr>
<tr>
<td>p₁₃</td>
<td>trip time</td>
<td>traffic confidence</td>
<td></td>
<td></td>
</tr>
<tr>
<td>p₁₄</td>
<td>accident confidence</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table I Criteria for each group
2.2. SPAN (Social Participatory Allocative Networks)

SPAN is a process developed by MacKinnon. This method has a voting system. A basic concept of the SPAN process is to permit each group member to distribute his original points among both group members and options. In this way, individual differences in judgmental abilities are taken into account [14, 15]. The steps for using SPAN are:
1. Each group has equally distributed point. This is generally 100 points.
2. The participators distribute their points among other participants. Points are added up for each group.
3. Each group distributed points which are collected by them again.
4. The terminal point is reached when the cumulative final points among the options are equal or close to (less than 0.1 percent) the total number of original points. The parcel located at the options, unlike those located among participators, will not be redistributed.
5. The optimum with the greatest number of points is selected as the best alternative in cumulative final points.

2.3. Application of SPAN

Combining individual decisions and accumulating at one point is supplied by SPAN method. Application results are given using SPAN at Table 2.
1. Participation ratio is determined for each group (experts, users, administrators and civil organizations) one by one. These data were gotten from public survey as average and can be seen first section of Table 2.

<table>
<thead>
<tr>
<th>Points Matrix</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Expert</td>
<td>0.28</td>
<td>0.27</td>
<td>0.24</td>
</tr>
<tr>
<td>User</td>
<td>0.27</td>
<td>0.33</td>
<td>0.19</td>
</tr>
<tr>
<td>Administrator</td>
<td>0.26</td>
<td>0.24</td>
<td>0.34</td>
</tr>
<tr>
<td>CO</td>
<td>0.23</td>
<td>0.30</td>
<td>0.15</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Participation ratio for Group and others</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Group</td>
<td>0.72</td>
<td>0.68</td>
<td>0.67</td>
</tr>
<tr>
<td>Other</td>
<td>0.28</td>
<td>0.32</td>
<td>0.33</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Proportion ratio of alternatives</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>A₁</td>
<td>55</td>
<td>46</td>
<td>45</td>
</tr>
<tr>
<td>A₂</td>
<td>45</td>
<td>54</td>
<td>55</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Preference distribution</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>A₁</td>
<td>15.40</td>
<td>14.72</td>
<td>14.85</td>
</tr>
<tr>
<td>A₂</td>
<td>12.60</td>
<td>17.28</td>
<td>18.15</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Standardization of Points Matrix</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Expert</td>
<td>0</td>
<td>27.20</td>
<td>26.13</td>
</tr>
<tr>
<td>User</td>
<td>26.64</td>
<td>0</td>
<td>24.12</td>
</tr>
<tr>
<td>Administrator</td>
<td>23.76</td>
<td>19.72</td>
<td>0</td>
</tr>
<tr>
<td>CO</td>
<td>21.60</td>
<td>21.08</td>
<td>16.75</td>
</tr>
</tbody>
</table>
### Results of iterations

<table>
<thead>
<tr>
<th>Alternatives</th>
<th>A₁</th>
<th>A₂</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>60.7000</td>
<td>64.3000</td>
</tr>
<tr>
<td>2</td>
<td>41.6669</td>
<td>43.8901</td>
</tr>
<tr>
<td>3</td>
<td>28.7223</td>
<td>30.3006</td>
</tr>
<tr>
<td>4</td>
<td>19.7682</td>
<td>20.8479</td>
</tr>
<tr>
<td>5</td>
<td>13.6134</td>
<td>14.3572</td>
</tr>
<tr>
<td>6</td>
<td>9.3729</td>
<td>9.8853</td>
</tr>
<tr>
<td>7</td>
<td>6.4538</td>
<td>6.8064</td>
</tr>
<tr>
<td>8</td>
<td>4.4437</td>
<td>4.6866</td>
</tr>
<tr>
<td>9</td>
<td>3.0597</td>
<td>3.2269</td>
</tr>
<tr>
<td>10</td>
<td>2.1067</td>
<td>2.2219</td>
</tr>
<tr>
<td>11</td>
<td>1.4506</td>
<td>1.5299</td>
</tr>
<tr>
<td>12</td>
<td>0.9988</td>
<td>1.0534</td>
</tr>
<tr>
<td>13</td>
<td>0.6877</td>
<td>0.7253</td>
</tr>
<tr>
<td>14</td>
<td>0.4735</td>
<td>0.4994</td>
</tr>
<tr>
<td>15</td>
<td>0.3260</td>
<td>0.3439</td>
</tr>
<tr>
<td>16</td>
<td>0.2245</td>
<td>0.2368</td>
</tr>
<tr>
<td>17</td>
<td>0.1546</td>
<td>0.1630</td>
</tr>
<tr>
<td>18</td>
<td>0.1064</td>
<td>0.1122</td>
</tr>
<tr>
<td>19</td>
<td>0.0733</td>
<td>0.0773</td>
</tr>
<tr>
<td>20</td>
<td>0.0505</td>
<td>0.0532</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>194.4535</td>
<td>205.3173</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Preference ratio</th>
<th>A₁</th>
<th>A₂</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>0.48</strong></td>
<td><strong>0.52</strong></td>
</tr>
</tbody>
</table>

**Table 2. Application of SPAN**

2. Points which are one group and total of other three groups for are given at second section of Table 2.
3. Alternatives consist of publicly owned bus (A₁) and privately owned bus (A₂). Proportion ratios for alternatives obtained from public survey.
4. After step 3, steps of SPAN will start. Elements of preference distribution section are obtained by multiplying elements of 2 and 3 sections.
5. Elements of point’s matrix are standardized in this step. After this, distribution iterations are realized in section 6.

According to results which are given on Table 2, groups preferred privately owned buses but preference ratio is close to each other. All results are given briefly but discussed as detailed in conference.

### 3. Conclusions

All countries have a problem with transportation, because transportation plans are not well documented considering the future. People create problems, and then they try to find a
solution for these problems. Local authorities are generally insufficient about public transportation in many cities. In this study, we try to emphasize necessity of public participation for decisions which are about transportation. Privatization is a decision about transportation. We used data which are gotten from public survey. The public survey carried out in Eskisehir, Turkey. Than decisions of stakeholders (public, transportation administrator, experts and civil organizations) was combined using SPAN method. Results were given as briefly.

References

Abstract. Emergency response administrators often face the difficult task of locating a limited number of ambulances in a manner that will yield the best service to a constituent population. Demand for ambulances varies spatially and temporally by day of the week and time of the day. To maintain or improve coverage increasingly EMS operators are practicing dynamic redeployment strategies by relocating the fleet throughout the day. However, frequent relocations are causing crew fatigue and increasing attrition. We formulate a probabilistic model to determine the locations for a fleet of ambulances during their shift. The objective is to minimize the number of relocations while maintaining coverage requirements. We introduce a search algorithm to solve the model and test the model using real data from Mecklenburg County, NC.

1. Background

The goal of emergency medical services (EMS) is to reduce mortality, disability, and suffering in persons (Brotcorne et al., 2003; Goldberg, 2004). EMS administrators and managers often face the difficult task of locating a limited number of ambulances in a way that will yield the best service to a constituent population. Typically, calls originating from a population center are assumed to be covered if they can be reached within a time threshold. This notion of coverage has been widely accepted and is written into the EMS Act of 1973, which requires that in urban areas 95 percent of requests be reached in 10 minutes, and in rural areas, calls should be reached in 30 minutes or less (Ball & Lin, 1993).

The literature on ambulance location problems is rich and diverse. A recent review by Brotcorne, Laporte, and Semet (2003) is an excellent source that traces the pertinent

1 The University of North Carolina at Charlotte, Business Information Systems and Operations Management, Department, Charlotte, NC 28223, USA, hrajagop@uncc, saydam@uncc.edu
developments in this domain. There are also reviews by Owen and Daskin (1998), Marianov and ReVelle (1995) and Shilling, Jayraman, and Barkhi (1993) that examine and classify early models. In the literature models are typically classified as deterministic or probabilistic and furthermore static or dynamic. Deterministic models are prescriptive in nature and used to find the optimal location of ambulances to minimize or maximize an objective.

Probabilistic models acknowledge the possibility that a given ambulance may not be available when it is called. These types of models provide a way to model uncertainty by either using a queuing framework or via a mathematical programming approach. The models based on a mathematical programming approach employ simplifying assumptions such as all units operate independently, yet have the same busy probability. For example server independence and system-wide server busy probability is a common assumption used in Daskin’s maximum expected coverage location problem (MEXCLP) (1983) and ReVelle and Hogan’s maximum availability location problem (MALP) (1989). In their second MALP formulation (MALPII), ReVelle and Hogan allow server busy probabilities to be different in various neighborhoods, sectors of a given region but not location specific. The Probabilistic Location Set Covering Problem (PLSCP) formulated by ReVelle and Hogan (1989) minimized the number of servers needed to guarantee coverage for a city. This model like MALPII uses neighborhood server busy probabilities. Later Marianov and ReVelle (1994) extended PLSCP using the assumption of neighborhood probabilities in MALPII to formulate Queuing Probabilistic Location Set Covering Problem (Q-PLSCP). They model each neighborhood as a multi-server loss system and calculate the neighborhood busy probabilities \( a \text{ priori} \) and then use it as an input into the system.

True probabilistic models are based on queuing theory (Larson, 1974) or simulation (Zaki et al., 1997) and they are by definition descriptive. Typically, they are employed to evaluate the vehicle busy probabilities and other performance metrics of a given allocation of ambulances. However, in the ambulance location literature prescriptive (optimization) models which incorporate some probabilistic aspects of the problem domain are also classified as probabilistic models (Owen & Daskin, 1998; Brotcorne et al., 2003). The hypercube model and its various extensions have been found particularly useful in determining performance of EMS systems (Larson, 1974, 1981; Batte et al., 1989; ReVelle, 1989; Burwell et al., 1993; Saydam et al., 1994; Daskin, 1995; Chan, 2001; Goldberg, 2004; Takeda et al., 2005). However, hypercube is computationally expensive. For \( m \) servers the number of simultaneous equations to solve would be \( 2^m \). For fleet sizes of 10 or more this approach would be computationally impossible to solve with the present day computers. To solve this problem Larson developed an approximation to the hypercube problem (1975) which would require solution only \( m \) simultaneous nonlinear equations for \( m \) servers. One of the assumptions used in Larson’s approximation is that service times are exponentially distributed and identical for all servers, independent of the customers they are serving. Jarvis generalized Larson’s approximation (1985) for loss systems (zero queue) by allowing service time distributions to be of a general type and may depend on both server and customer.

Redeployment models however look at operational level decisions managers make daily or even hour by hour in an attempt to relocate ambulances in response to demand fluctuations by time and space. There have been two earlier papers on relocation in the EMS literature (Repede & Bernardo, 1994; Gendreau et al., 2001). Repede and Bernardo (1994) extended MEXCLP for multiple time intervals to capture the temporal variations in
demand and unit busy probabilities, hence, called their model TIMEXCLP. Their application of TIMEXCLP to Louisville, Kentucky resulted in an increase of coverage while the average response time decreased by 36%. The most recent and comprehensive dynamic relocation model is developed Gendreau et al. (2001). The objective of their dynamic double standard model at time $t$ (DDSM) is to maximize backup coverage while minimizing relocation costs. There are several important considerations incorporated into this model. While the primary objective is to maximize the proportion of calls covered by at least two vehicles within a distance threshold, the model penalizes repeated relocation of the same vehicle, long round trips, and long trips. The model’s input parameters are updated each time a call is received and DDSM is solved. To solve this complex model, particularly at short time intervals, Gendreau, Laporte and Semet developed a fast tabu search heuristic implemented on eight parallel Sun Ultra workstations. Using real data from the Island of Montreal, their tests showed that the algorithm was able to generate new redeployment strategies for 95% of all cases. Furthermore, comparisons with CPLEX generated exact solutions showed that the worst case departure from optimality was only 2%.

Apart from these studies (Repede & Bernardo, 1994; Gendreau et al., 2001), there has been very little work done concerning the periodic relocation (redeployment) of ambulances in an environment where demand and the location and quantity of available ambulances are changing. Brotcorne, LaPorte and Semet (2003) predict that current and future advances in this field are likely to be in probabilistic location models, dynamic redeployment models, and fast heuristics designed to solve generally large scale problem instances.

In this paper, inspired by Gendreau et al. (2001) we formulate the minimum redeployment location (MRL) model which tries to determine the minimum number of redeployments for a given set of ambulances that meet or exceed a predetermined expected coverage requirement in response to changing demand. While Gendreau et al. (2001) approach the problem from a real-time perspective, our approach is for shift planning. We determine ambulances posts for each time interval where there is a significant change in demand in volume and space. We also enhance the realism of our model by computing server specific busy probabilities using Jarvis’ hypercube approximation (1985) which is similar to Galvão, Chiyoshi and Morabito (2005). This paper is organized as follows. Section 2 presents the minimum redeployment coverage problem. Section 3 details the search algorithm. Section 4 presents the numerical experiments and conclusions and suggestions for future work are discussed in Section 5.

## 2. The model

Let $J_t$ be the set of nodes which contains servers at time $t$, $D_j$ be the set of nodes within the maximum relocation distance of node $j$, $y_{kj}$ be the number of servers covering node $k$ at time $t$, $\rho_{ij}$ be the busy probability for server $i$ at time $t$, $m$ be the number of servers that are deployed, and $\xi^+_js$ is the number of redeployments moved out of node $j$ after time $t$. 

710
period \( t \) and \( \xi_{ij}^- \) is the number of redeployments moved into node \( j \) after time period \( t \).

Also let,

\[
x_{ij,t} = \begin{cases} 
1 & \text{if server } i \text{ is in node } j \text{ at time } t \\
0 & \text{if not}
\end{cases}
\]

Minimize:

\[
\sum_{t=1}^{T} \sum_{j=1}^{n} (\xi_{ij}^+ + \xi_{ij}^-)
\]

Subject to:

\[
x_{ij,t} - x_{ij,t+1} = \xi_{ij}^+ - \xi_{ij}^- \quad \forall i, j \in J_t, t
\]

\[
x_{ij,t} = \sum_{k \in D_j} x_{ik,t+1} \quad \forall i, j \in J_t, t
\]

\[
\sum_{i=1}^{m} \sum_{j=1}^{n} a_{ik,t} x_{ij,t} \geq y_{k,t} \quad \forall k, t
\]

\[
\sum_{k=1}^{m} \sum_{i=1}^{n} \sum_{j=1}^{m} Q(m, p_i, i - 1) k_{ik,t} y_{k,t} \left(1 - p_{i,t}\right) \prod_{i=4}^{i-4} p_{i,t} \geq c_t \quad \forall t
\]

\[
\sum_{j=1}^{m} x_{ij,t} = m \forall t
\]

\[
y_{k,t} \geq 0 \quad \forall i, t
\]

\[
x_{ij,t} \in \{0, 1\}
\]

The objective function (1) counts the number of redeployments (trips). Constraint (2) enumerates the number of redeployments with \( \xi_{ij}^+ \) storing the number of redeployments moved out of node \( j \) after time period \( t \) and \( \xi_{ij}^- \) being the number of redeployments moved into node \( j \) after time period \( t \). The LHS of constraint (3) is the set of current locations of ambulances during \( t \). Similar to Gendreau et al. (Gendreau et al., 2001) we also constrain the distance an ambulance can be redeployed since it is not practical to move ambulances across large distances. The RHS of constraint (3) ensures that each server can only be moved within the maximum allowed relocation distance and for each unique ambulance-location pair, the set of possible relocations are dynamically developed in the set defined as \( D_j \). Thus, constraint (3) ensures that the new locations for the next time period, \( t+1 \), are all within the maximum relocation distance threshold. Constraint (4) ensures that the coverage distances are enforced and constraint (5) ensures that the expected coverage will never go below the predetermined limit. Constraint (6) ensures that the number of servers in the system is fixed and constraints (7) and (8) enforce the integrality constraints.
3. The algorithm

There have been various attempts to identify near-optimal solutions for location problems through the use of meta-heuristic search methods (Michalewicz, 1999; Michalewicz & Fogel, 2000) and more recently, evolutionary algorithms (Beasley, 1993; Benati & Laporte, 1994; Saydam et al., 1994; Gendreau et al., 1997, 2001; Aytug & Saydam, 2002; Brotcorne et al., 2002; Jaramillo et al., 2002; Saydam & Aytug, 2003; Galvao et al., 2005). To solve MRL, we developed a Comprehensive Search Algorithm (CSA). The CSA algorithm works with the assumption that the fleet size for the shift is predetermined and cannot be changed. To find a set of server locations that can meet or exceed the minimum coverage requirements for \( t = 1 \) we use EMEXCLP as detailed by Galvão, Chiyoshi and Morabito (2005) which can be solved by any one of the meta-heuristic search methods such as tabu search (Glover & Laguna, 1997), evolutionary algorithm (Holland, 1975), simulated annealing (Kirkpatrick et al., 1983) or using hill climbing with random restarts as shown by Rajagopalan, Vergara, Saydam and Xiao (2005). For this problem, we use hill climbing with random restart to get the initial set of locations for the first time interval.

The data structure for the algorithm is a vector which is of size \( m+1 \) where \( m \) is the number of servers. The first index of the vector contains the coverage provided by the set of server locations and each index after that gives the location of each server. Once the best locations for the first time interval are found CSA uses these locations to determine the minimum number of redeployments needed to ensure the coverage requirements for each of the subsequent time intervals. CSA first checks to see if the current locations which ensured minimum coverage in the previous time interval still continue to ensure minimum coverage. If it does not then CSA starts to search for the minimum relocations which will ensure the required coverage.

The algorithm starts with the first server searches for the location within maximum relation distance which gives the greatest increase (gain) in coverage. Once this location has been found, the location is then stored in memory with the new coverage. With the first server in the original location, we search for the “best” location for the second server and this location and the coverage increase is stored in memory. This is done for all \( m \) servers and the server which results in the maximum increase coverage is moved to its new location which is within maximum relocation distance. This results in \( m \) searches where \( m \) is the number of servers. Once the server is selected and moved, that server is placed in a tabu list and cannot be moved again. If the new configuration of servers does not meet the minimum coverage requirement, the algorithm again searches among \( m-1 \) servers for the best server to relocate to a new node and this continues until the minimum coverage requirement is met. The worst case scenario for this algorithm would be \( m \) factorial searches. The algorithm also terminates if it finds that it cannot improve the current coverage by moving any of the available servers. Therefore the worst case scenario will occur only if at every pass the algorithm finds a relocation of one server which improves its current solution but does not reach the minimum coverage requirements. Clearly, CSA is a variation of the Steepest Ascent Algorithm for this problem as we search for the server in each pass which when relocated gives the maximum gain in coverage. We believe that the Steepest Ascent Algorithm is well suited for this model because we are trying to minimize the number of redeployments, and therefore we need to move the fewest number of ambulances.
while trying to get the maximum increase in coverage. Details of the algorithm are available from the authors.

4. Experiments and results

We test our model, MRL, on two important metrics: (1) quality of solutions, and (2) CPU times. The quality of solution is measured by the number of redeployments and the expected coverage. If any of the solutions prescribed by our model does not meet the coverage requirements then we declare that solution as suboptimal or unacceptable. The experiments are conducted on a Dell PC Pentium IV 2.4 MHz with 512 MB RAM. the ISA is coded in Java (jdk 1.4). We applied our model to Mecklenburg County, NC, a region spread over approximately 540 square miles with a 2004 population of 801,137. The county’s call demand distribution is known to fluctuate significantly by day of the week and time of the day which provides an opportunity for our research to be applied to real-life data thus increasing the chance of being implemented by interested parties and agencies. In 2004 the county received a total of 77,292 calls of which 61,630 of them were classified as emergency triggering ambulance dispatches. For this experiment we used data from one day of the week (Monday) for the entire year. There were 8,742 ambulance dispatches in the year 2004 for all Mondays distributed over the entire region. For this experiment we chose eight, 3-hour time intervals for our model. For each time period, the fleet size is 20 and maximum allowable relocation distance is 6 miles. Coverage radius is 8 miles and 95% of all calls must be covered. For this run, results show that Mecklenburg County only needs to redeploy two ambulances at 6 am in the morning and then four ambulances at 9 am in morning. During the other times the existing locations are robust enough to account for fluctuations in demand and do not need any redeployments. It is important to note that for periods 3 and 4 without redeployments the expected coverage would have dropped below 90%. Work in progress: Sensitivity of the solutions to maximum redeployment distance and analysis of other days>>

5. Conclusions

We formulated a new model for dynamic environments with the objective of minimizing the number of redeployments required to meet mandated coverage requirements given a set of ambulances. We developed a complete search algorithm (CSA) based on the steepest descent algorithm to solve the new model. Preliminary results are encouraging. Remaining work include fine tuning the algorithm, conducting more tests, and applying the model to all seven day’s data from Mecklenburg County, developing Pareto curves for increasing the maximum travel distance versus number of relocations while meeting the coverage requirement.

References available from the authors upon request.
DYNAMIC SET PARTITIONING APPLICATIONS: A HEURISTIC APPROACH

Luca COSLOVICH\textsuperscript{1}, Raffaele PESENTI\textsuperscript{2}, Walter UKOVICH\textsuperscript{1}

Abstract. In this paper we consider large-scale set partitioning problems. Our main purpose is to solve set partitioning problems originating from particular real-world dynamic applications. More precisely, we seek how to include a dynamic request in a previously found ("static") solution. We have developed a fast heuristic algorithm that exploits both the static solution and the static algorithm that provided it in order to build another solution including the new request. In our computational tests, we have compared our “dynamic” solutions with the static ones obtained assuming a complete a priori knowledge of all the requests that have to be carried out.

1. Introduction

In this paper we consider large-scale set partitioning problems. Set partitioning is a fundamental model in combinatorial optimization: main applications include truck delivery management, vehicle scheduling, bus driver scheduling, and airline crew scheduling. It is also well known that set partitioning is an NP-hard problem [8].

The tackled real-world scenario originates from vehicle routing and scheduling. Starting from a feasible solution of the set partitioning problem, obtained with a static heuristic, we try to deal with another request (\textit{i.e.}, transportation order) that shows up at the last minute. This is a core problem faced by fleet managers, since it is in general impossible to run the static algorithm once more, due to the time taken to solve the new instance of the set partitioning problem. Hence, it is of basic importance to try to deal with the new request in a dynamic fashion, in order not to have to reject it.

We have developed and implemented a fast algorithm that finds an overall solution including the new request; it exploits the knowledge of the solution previously obtained and makes use of the static algorithm, as well.

\textsuperscript{1} DEEI - University of Trieste, Via A. Valerio 10, I-34127 Trieste (Italy), \{coslovic, ukovich\}@units.it
\textsuperscript{2} DINFO - University of Palermo, Viale delle Scienze, I-90128 Palermo (Italy), pesenti@unipa.it
In our computational tests, we have compared the dynamic solutions with the static solutions obtained assuming a complete a priori knowledge of all the requests that have to be executed. The real-world set partitioning instances used in the numerical tests originate from container-trucking industry real cases [5].

The paper is organized as follows. In Section 2, the set partitioning problem is formally defined and some of its applications are mentioned. In Section 3, the adopted static solution approach is outlined and, in Section 4, our dynamic solution algorithm is detailed. In progress computational results are presented in Section 5. In the last section, some conclusions and our ongoing developments are summarized.

2. The Set Partitioning Problem (SPP)

2.1. Definition

First, we give a formal description of this problem. Let \( M \) be a non-empty and finite set and let \( F \) be the set of feasible subsets of \( M \). Associated with each element \( j \) of \( F \) is a cost \( c_j \). Let \( A \) be a \( 0-1 \) matrix with a row for each element in \( M \) and a column for every characteristic vector of a feasible subset in \( F \). The problem is to find a collection of elements of \( F \) (i.e., of columns of \( A \)) giving a partition of \( M \) such that the cost sum of these elements is minimal.

2.2. Applications

Many real life problems, such as vehicle routing and airline crew scheduling, can be formulated as SPPs. The SPP has been studied extensively (see Balas and Padberg [1] for a survey of some of its applications and solution methods). Although the best known application of the SPP is airline crew scheduling [10], several other applications exist, including truck delivery management [9] and vehicle scheduling [2]. Moreover, some of the largest real-world SPP instances are generated by vehicle scheduling problems (see, e.g., [2]).

3. The static solution approach

In order to tackle large-scale SPPs, we have implemented a heuristic algorithm (in this work, the matrix \( A \) of the largest tested instances presents hundreds of thousands columns). In particular, we have adopted a Lagrangian relaxation based heuristic that exploits the subgradient method, a well-known approach used to find good Lagrangian multipliers (see [7] for an introduction to Lagrangian relaxation). We have implemented such a heuristic solution approach since it has proven to be quite effective in tackling large-scale covering and partitioning problems [4]. The implemented SPP algorithm is an adaptation of the Set Covering method presented in [3]. Although we have implemented only one heuristic algorithm, it is important to note that different static solution approaches might differently affect our dynamic solutions, as explained in the next section.
4. The dynamic solution approach

In the following, our dynamic solution algorithm is detailed, in five main steps.

Step 1. Build the new SPP instance, including the last minute request. The new matrix $A$ presents one more row and, in general, several more columns. As a matter of fact, the new request may be matched with other ones, in order to build new feasible pairings (i.e., columns). Without loss of generality, let the new row be a vector whose first $|F|$ elements are equal to zero and whose last $k$ elements are equal to one. That is, the last $k$ columns represent the new pairings that have been built with the last minute request.

Step 2. Include the $k$ new columns in the static solution, one at a time, and proceed after each inclusion to the next step. Note that exactly one of those columns will belong to the dynamic solution.

Step 3. Each time a new column is included in the current solution, in general, some of the columns belonging to the static solution have to be removed, in order to maintain feasibility. Nevertheless, unfortunately, this removal may uncover some rows and yield again to an infeasible solution.

Step 4. Cover all the uncovered rows by solving a reduced SPP. In particular, the columns of matrix $A$, in the reduced SPP, must be such that they do not cover rows which are already covered, in the solution we are constructing. In practical applications, the reduced SPP can be solved by the static algorithm, due to the small size of the reduced instances. Note that a feasible solution always exists, since a pairing including only one row is a feasible one. Go to Step 2 until every new column has been considered.

Step 5. Select the best solution found.

5. Numerical results

In this in progress work, we have tested 10 real-world SPP instances originating from the container-trucking industry. Such instances have been provided by an Italian transportation company that operates in Italy and in the neighboring countries. Its fleet consists of hundreds of tractors and semi-trailers capable of handling all types of containers. Its main customers are shipping lines, maritime agencies, service companies, industrial firms as well as private individuals [6].

The considered real-world SPP instances model a container transportation problem. The rows of matrix $A$ represent the transportation orders to be executed in a given time horizon (namely, a day) and the columns represent the so-called pairings, i.e., the sequences of orders that can be carried out by a single vehicle.

The numerical tests have been performed on a Pentium 4 laptop with 1.6 GHz and 256 MB of memory. The implemented algorithms have been coded in ANSI C.

Table 1 reports the main characteristics of the tested real-world SPP instances. As it can be seen, the matrix $A$ of the largest instance presents more than 680 thousands columns. In particular, the most significant instances are the ones whose matrix $A$ presents more than a hundred thousand columns. The variety of the instances depends on the considered day and on the number of transportation orders that have to be carried out. The density of the unitary elements of matrix $A$ is always less then 1%.
<table>
<thead>
<tr>
<th>Instance</th>
<th>Algorithm</th>
<th>Time (ms)</th>
<th>Gap %</th>
<th>Cost</th>
<th>LB</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-a</td>
<td>Static</td>
<td>2,363</td>
<td>1.01</td>
<td>33,624,375</td>
<td>33,289,737</td>
</tr>
<tr>
<td>1-b</td>
<td>Static</td>
<td>2,843</td>
<td>1.09</td>
<td>33,013,602</td>
<td>32,658,310</td>
</tr>
<tr>
<td>1-b</td>
<td>Dynamic</td>
<td>50</td>
<td>1.38</td>
<td>33,108,957</td>
<td>32,658,310</td>
</tr>
<tr>
<td>2-a</td>
<td>Static</td>
<td>823,873</td>
<td>3.00</td>
<td>77,919,835</td>
<td>75,651,441</td>
</tr>
<tr>
<td>2-b</td>
<td>Static</td>
<td>798,818</td>
<td>2.93</td>
<td>76,791,582</td>
<td>74,605,709</td>
</tr>
<tr>
<td>2-b</td>
<td>Dynamic</td>
<td>2,664</td>
<td>3.35</td>
<td>77,102,146</td>
<td>74,605,709</td>
</tr>
<tr>
<td>3-a</td>
<td>Static</td>
<td>123,276</td>
<td>1.92</td>
<td>104,606,240</td>
<td>102,632,989</td>
</tr>
<tr>
<td>3-b</td>
<td>Static</td>
<td>121,764</td>
<td>2.48</td>
<td>103,639,486</td>
<td>101,155,465</td>
</tr>
<tr>
<td>3-b</td>
<td>Dynamic</td>
<td>750</td>
<td>2.38</td>
<td>103,539,443</td>
<td>101,155,465</td>
</tr>
<tr>
<td>4-a</td>
<td>Static</td>
<td>69,540</td>
<td>1.20</td>
<td>53,974,000</td>
<td>53,334,418</td>
</tr>
<tr>
<td>4-b</td>
<td>Static</td>
<td>77,862</td>
<td>1.85</td>
<td>52,944,074</td>
<td>51,984,473</td>
</tr>
<tr>
<td>4-b</td>
<td>Dynamic</td>
<td>270</td>
<td>1.73</td>
<td>52,885,529</td>
<td>51,984,473</td>
</tr>
<tr>
<td>5-a</td>
<td>Static</td>
<td>12,858</td>
<td>0.75</td>
<td>73,705,420</td>
<td>73,153,215</td>
</tr>
<tr>
<td>5-b</td>
<td>Static</td>
<td>14,300</td>
<td>0.58</td>
<td>72,421,156</td>
<td>72,003,717</td>
</tr>
<tr>
<td>5-b</td>
<td>Dynamic</td>
<td>180</td>
<td>0.64</td>
<td>72,465,975</td>
<td>72,003,717</td>
</tr>
<tr>
<td>6-a</td>
<td>Static</td>
<td>1,074,625</td>
<td>4.87</td>
<td>75,900,288</td>
<td>72,372,180</td>
</tr>
<tr>
<td>6-b</td>
<td>Static</td>
<td>1,048,157</td>
<td>6.40</td>
<td>74,622,141</td>
<td>70,117,927</td>
</tr>
<tr>
<td>6-b</td>
<td>Dynamic</td>
<td>8,270</td>
<td>6.92</td>
<td>74,970,021</td>
<td>70,117,927</td>
</tr>
<tr>
<td>7-a</td>
<td>Static</td>
<td>60,415</td>
<td>2.34</td>
<td>72,836,202</td>
<td>71,174,174</td>
</tr>
<tr>
<td>7-b</td>
<td>Static</td>
<td>70,962</td>
<td>2.99</td>
<td>71,421,416</td>
<td>69,345,414</td>
</tr>
<tr>
<td>7-b</td>
<td>Dynamic</td>
<td>400</td>
<td>3.45</td>
<td>71,736,236</td>
<td>69,345,414</td>
</tr>
<tr>
<td>8-a</td>
<td>Static</td>
<td>1,055,919</td>
<td>5.45</td>
<td>86,346,209</td>
<td>81,883,092</td>
</tr>
<tr>
<td>8-b</td>
<td>Static</td>
<td>1,055,618</td>
<td>5.33</td>
<td>85,402,633</td>
<td>81,082,944</td>
</tr>
<tr>
<td>8-b</td>
<td>Dynamic</td>
<td>10,595</td>
<td>5.33</td>
<td>85,406,546</td>
<td>81,082,944</td>
</tr>
<tr>
<td>9-a</td>
<td>Static</td>
<td>160,290</td>
<td>2.04</td>
<td>85,510,688</td>
<td>83,800,454</td>
</tr>
<tr>
<td>9-b</td>
<td>Static</td>
<td>150,958</td>
<td>2.92</td>
<td>84,373,435</td>
<td>81,980,386</td>
</tr>
<tr>
<td>9-b</td>
<td>Dynamic</td>
<td>620</td>
<td>2.95</td>
<td>84,401,917</td>
<td>81,980,386</td>
</tr>
<tr>
<td>10-a</td>
<td>Static</td>
<td>414,004</td>
<td>3.00</td>
<td>71,080,092</td>
<td>69,010,539</td>
</tr>
<tr>
<td>10-b</td>
<td>Static</td>
<td>794,872</td>
<td>2.98</td>
<td>69,564,423</td>
<td>67,553,405</td>
</tr>
<tr>
<td>10-b</td>
<td>Dynamic</td>
<td>1,743</td>
<td>4.12</td>
<td>70,335,875</td>
<td>67,553,405</td>
</tr>
</tbody>
</table>

Table 1. Characteristics of the tested real-world SPP instances.

Table 2. Computational results.
In Table 2 our in progress numerical results are summarized. In the first column we report the tested instance. The suffix “a” indicates a static instance whereas the suffix “b” indicates the dynamic one (i.e., the one with the last minute row: see Step 1 of our dynamic algorithm in Section 4). The second column reports the type of algorithm used in the test. Computation times are displayed in the third column. The gap between the solution found and the best computed lower bound is reported in the fourth column, in percentage. The cost, expressed in generic cost units, of the solution found and of the best computed lower bound are reported in the last two columns, respectively.

As it can be seen in Table 2, the computation times of our dynamic algorithm are operatively acceptable, also in a dynamic context: less than 11 seconds for the biggest tested instance (8-b). Note that, for that instance, the static algorithm uses more than one thousand seconds, that is about 17 minutes (actually, it really terminates for a time excess, since 1,000 seconds is the adopted time limit in this test).

As it can be expected, the gaps of the solution values obtained with the dynamic algorithm are, in general, slightly worse then those provided by the static one. Nevertheless, they keep operatively good: the worst gap increase is 1.14% (10-b). Moreover, note that in some cases (3-b and 4-b) the gap in the dynamic case even improves the static one. This is due to the non-deterministic behavior of the static algorithm: two runs, with the same parameter settings, may yield to slightly different solution values.

6. Summary and ongoing developments

In this paper we have considered large-scale set partitioning problems originating from particular real-world dynamic applications. We have developed and implemented a fast algorithm able to include a dynamic request in a previously found static solution. Our heuristic algorithm exploits both the static solution and the static algorithm that provided it.

We have tested so far 10 real-world SPP instances originating from the container-trucking industry. Such instances have been provided by an Italian transportation company. In our numerical tests, we have compared the dynamic solutions with the static ones obtained assuming a complete a priori knowledge of all the requests that have to be carried out.

The numerical results indicate that the developed algorithm is very efficient with respect to the static one, also in a dynamic context, and the quality of the solutions keep acceptable for practical applications.

Our ongoing developments are mainly devoted to perform more extensive numerical tests and a relative statistical analysis. We also intend to use different static algorithms and compare the obtained dynamic solutions.

References


Abstract. Evacuation modeling is a multi-stage process, in the macro sense including demand generation and traffic assignment modules. Both these modules and sub-modules have inherent human behavior and interaction uncertainties, since the main actors are the evacuees. Besides the human action under emergency/panic conditions, during natural disasters there are additional uncertainties related with the disaster type. This study approaches the evacuation problem from a probabilistic point of view, addressing the sources of uncertainty mainly, demand and network capacities. A dynamic traffic assignment model with probabilistic constraints is proposed to assess network-wide time-dependent traffic conditions under these uncertainties.

1. Introduction

Evacuation modeling is a multi-stage process, in the macro sense including demand generation and traffic assignment modules. Especially demand generation part requires heavy human behavioral insights. Many surveys have been conducted to shed light to human decision process however there are hardly a few factors that are commonly accepted to have effect. Among the agreed factors, quantifying the degree of influence is another major challenge. Even assuming that actual demand can be predicted, issue of human behavior is still important while modeling the evacuee/driver compliance with the authorities. Most evacuation performance studies are based on system optimal assignment assuming that the evacuees will follow the proposed routes. The natural circumstances (storm track, intensity etc.) make this issue further complicated. Hurricane surge flooding a
road, or a landslide can significantly reduce the network capacity, and consequently the outcome of the traffic assignment. On the other hand, the impact of flood risk can be considered as part of the demand generation process. The first part of this paper will briefly go over the uncertainties in demand generation and traffic assignment, with a brief review of the impacts of demand generation model choice on the evacuation performance. Then, a system optimal dynamic traffic assignment model with probabilistic capacity constraints will be presented and applied to a simple network. Finally, conclusions and future work will be presented.

2. Uncertainty in Evacuation Demand

A general travel demand forecasting process for hurricane evacuations was first described by Lewis [7]. He approached the problem using the traditional urban travel demand forecasting methodology. State-of-practice in hurricane evacuation travel demand modeling has two main steps: 1) the estimation of total evacuation demand, and 2) the estimation of departure times [6]. ‘Participation rates’ are the most common method for estimating the total evacuation demand. While determining these rates, evacuation behavior is considered homogeneous in geographic subdivisions of the study area and they are assumed to vary among various geographic subdivisions (evacuation zones) depending on the severity of the storm and flood risk. Participation rates are generally established subjectively based on past behavior under different storm conditions [6].

Regarding the factors affecting evacuation decision, the survey results obtained in different regions are often found to identify different set of parameters as significant. Furthermore, even two studies based on the same region may give different results. Baker [5] after studying 26 hurricane evacuations, identified the five most important variables in hurricane evacuation: 1) Risk Level (Hazardousness) of the area, 2) Actions by public authorities, 3) Housing type, 4) Prior perception of personal risk, 5) Storm specific threat factors. These are some of the major factors that are agreed by most of the researchers in this area to affect the evacuation behavior. However, if the factors are further investigated, it can be seen that apart from the housing type, all the other factors are inherently probabilistic. Evacuees may not agree with this “risky” region definitions and leave their homes before/after the presumed times. That causes an artificial unexpected bottleneck in the road network, delaying the evacuation of regions that are “really” in danger. This phenomenon is mentioned in the literature under the name “Shadow Evacuation”. Hurricanes make the landfall on the shores, and tourist population can be considerable high. Not being familiar with the area changes the decision mechanisms that are used for the resident population. Storm specific factors, such as track and intensity, are already uncertain which leads to uncertain evacuation order timing.

2.1.1. Effect of Demand Model Choice on Network Evacuation Performance

It is important to mention the importance of the evacuation demand model choice in the context of evacuation modeling. Literature points out several different models which all claim to capture the uncertainties in hurricane demand, however choice of the model for an
evacuation study may change the final outcome considerably. Ozbay et. al. [9], and Ozbay and Yazici [8] study the effects of demand models. It can be briefly said that the studies conduct a sensitivity analysis of demand generation models for studying the network-wide performance of traffic under evacuation conditions. In Ozbay et. al. [9], New Jersey Cape May County network is analyzed using a deterministic multiple origin-single destination System Optimal Dynamic Traffic Assignment (SO-DTA) for different demand patterns. The possible outcome of capacity decrease in the network is also studied by applying a time-dependent reduction of overall network capacity. In [9] it is shown that the choice of demand model affects the overall performance parameters. Besides, within the same model, changes in participation rates (PR), capacity reduction of the network links also has considerable effects. Shadow evacuation, which refers to the evacuees that have already evacuated before the evacuation order, is found not to have a considerable effect. In the other study by Ozbay and Yazici [8], focusing on behavioral response curves (Sigmoid, or S-Curves), which are the most popular model used in all official studies by Army Corps of Engineers [4], similar results are obtained. Mathematical formula of behavioral response curves is as follows:

\[
P(t) = \frac{1}{1 + \exp(-\alpha(t-H))}
\]

where \( P(t) \) is the cumulative percentage of the total trips generated at time \( t \). The “\( \alpha \)” parameter represents the response of the public to the disaster and alters the slope of the cumulative traffic-loading curve. \( H \) is the half loading time; the time at which half of the vehicles in the system have been loaded onto the highway network. \( H \) defines the midpoint of the loading curve and can be varied by the planner according to disaster characteristics. Again employing SODTA, it is shown that there is major difference in the performance of S-curves depending on its parameters. Slow-Medium-Fast response curves do not necessarily imply a similarly increasing network performance and the results are not linearly dependent on the response curves. There can be a wide difference between a slow and a medium response curve, especially in case of a network capacity reduction. The important results concerning the current work are that the performance results can change dramatically with the capacity reductions in the network. Thus, evacuation studies must also consider the reliability of the network capacity. The section three proposes a mathematical formulation for capturing the probabilistic nature of the capacities.

3. Traffic Assignment

Both user equilibrium and system optimal assignments are considered in the literature for modeling evacuation. However, system optimal assignment theoretically suits evacuation process better since public good is of concern and system optimal is known to produce optimal results for the society. Here the optimal result is considered to be minimum clearance time for evacuees and minimum overall delay vehicles. However, as mentioned before, the performance of system optimal assignment heavily relies on cooperation of the evacuees with a central authority. In other words, the impact of the human behavior has to be considered as part of the traffic assignment. Even when system optimal assignment is considered to work as intended, there are also other important factors that need to be taken into account such as, risk exposure and the time-dependent change of roadway capacity.
Risk exposure is an important parameter that should be considered in all processes dealing with human safety. The proposed models can assure some overall system safety. However, each agent in the model should be exposed to equal amount of risk exposure. In hurricane evacuation, this can be the period that the evacuees are still on their way to their destination. Although system optimal assignment algorithms minimize the overall time that those evacuees are considered to be unsafe, it neglects the change in the track or speed of the hurricane, which may make the landfall before the estimated times. As recent storm cases (such as in hurricane Rita, 2005) have shown us, the track and speed of the hurricane can change rapidly. That’s why this phenomenon must be incorporated in an evacuation study for reliable risk exposure estimates.

Evacuation network capacity is another parameter that is open to disruptions by outside factors. Storm surge is simply water that is pushed toward the shore by the storm. This can cause severe flooding and roadway capacity reductions in coastal areas, particularly when the storm tide coincides with the normal high tides. There are tools to evaluate the threat from storm surge, such as SLOSH (Sea, Lake and Overland Surges from Hurricanes). SLOSH is a computerized model run by the National Hurricane Center (NHC) to estimate storm surge heights and winds. Hence, a pre-analysis of the flooding probabilities can be incorporated while deciding about the optimal routes to evacuate the public. This study proposes a probabilistic DTA scheme to address this unique and challenging problem.

3.1. System Optimal Dynamic Traffic Assignment with Probabilistic Capacity Constraints

In this section, a single origin single destination, linear dynamic traffic assignment model, proposed by Ziliaskopoulos [1] is extended with probabilistic capacity constraints. The proposed model uses cell transmission model, proposed by Daganzo [2],[3], for modeling the traffic flow and a linear optimization module is employed for finding the system optimum travel time for all vehicles. The cell flow capacities are assumed to be probabilistic, with a user specified discrete distribution. The extended stochastic programming problem, re-written based on LP standard form, is given simply in (2).

\[
\begin{align*}
\min & \sum_i \sum_j c_{ij} x_{ij} \\
\text{s.t.} & A_{d} \mu = b_{d}, \\
& A \mu \leq b, \\
& P(\mu \geq Q) \geq p, \\
& x_{ij} \geq 0, y_{ij} \geq 0, \forall (i, j) \in \xi, \forall t \in T
\end{align*}
\]

where \( \mu = [x_\mu, y_\mu]^T \) is the vector of system states, \( p \) is the value that determines reliability of the assignments and \( Q \) is the random capacity. The problem representation differs from the common LP formulation with the flow constraint, which is probabilistic. For the solution of this stochastic LP, P-Level Efficient Points (pLEP) method proposed by Prekopa [10] is employed.
3.1.1. P-Level Efficient Points

**Definition:** A point $z \in \mathbb{Z}_p$ is called a pLEP of a probability distribution function $F$, if $F(z) \geq p$ and there is no $y \leq z, y \neq z$ such that $F(y) \geq p$, where $p \in (0, 1)$.

Basically, pLEPs provide discretized set of points, which give the lower bound of a specific probability distribution. pLEPs are used in the deterministic equivalent of the probabilistic constraint and assure that the constraint will satisfy the given reliability level. For a scalar random variable $\xi$ and for every $p \in (0, 1]$, there is exactly one $p$-efficient point, namely, the smallest $z$ such that $F(z) \geq p$. Figure 1 shows the pLEP points for a two dimensional case. To enumerate the $p$-efficient points for a multidimensional discrete probability distribution, Prekopa [11] proposes a recursive algorithm. After finding pLEP points, the resulting LP is easily solved by conventional methods.

A straightforward way for solution is to find all $p$-efficient points and to process all LP problems. Let $v^{(i)}$ is the optimal solution to the $i^{th}$ LP problem with constraint $\mathbf{c}^T \mathbf{v} \geq z^{(i)}$. If $c^T v^{(i)} = \min_{y \neq z^{(i)}} c^T v^{(i)}$, then $v^{(i)}$ is the optimal solution. However, for high-dimensional random vectors the number of pLEPs can be very large and enumeration of those points may not be efficient. For this kind of situations, optimality bounds, or using a dual problem solution as a starting point can be used to decrease the number of pLEPs. For our problem, the network is designed to be tractable with straightforward enumeration of pLEPs, however a detailed discussion on dealing avoiding numerous pLEP enumeration can be found in [12]. Details of using pLEPs for stochastic LP problems can be found in [11].

3.2. Numerical Example

For the analysis, 3 different cases are tested for a simple network.

1. Deterministic case: All flow capacities are assigned equal to the full capacity (reproduction of Ziliaskopoulos’ paper for different network and demand). Results of this case are used as a base to show the changes in the traffic assignment with probabilistic constraints.
2. Stochastic Case-1: Flow capacity probabilities are time independent (Capacity change probabilities are the same during whole evacuation).

3. Stochastic Case-2: Flow capacity probabilities are time dependent representing changing flood risk as time evolves.

The road network shown in Figure 2 is used for the analysis. The numerical values for the demand and network parameters (e.g. link capacities, number of lanes) are shown in Table 1. For those values, a cell is chosen to be 500 ft long as assumed in [1], then relevant flow and physical capacities are given accordingly. Network is loaded for first 3 time steps (each assumed to be 10 seconds) with 120, 180 and 120 vehicles respectively. It takes around 25 time steps for all vehicles to leave the network. Table 2 shows the available lane probabilities for the flood risk cells. Note that there are 2 different rows for cell #8. The bottom row represents the increase of probabilities for less available number of lanes.

![Figure 2 Cell Representation of the Network](image)

**Table 1 Cell Physical Properties**

<table>
<thead>
<tr>
<th>Cell #</th>
<th># of Lanes</th>
<th>Max Flow (veh/τ/ln)</th>
<th>Physical Capacity, (N_i) (veh / L<strong>1</strong> at τ)</th>
<th>Cell Length (ft)</th>
<th>Speed (ft/sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>2</td>
<td>12</td>
<td>20</td>
<td>500</td>
<td>50</td>
</tr>
<tr>
<td>6, 8, 9, 10</td>
<td>4</td>
<td>24</td>
<td>40</td>
<td>500</td>
<td>50</td>
</tr>
<tr>
<td>2, 3, 4, 5, 11, 12, 13</td>
<td>6</td>
<td>36</td>
<td>60</td>
<td>500</td>
<td>50</td>
</tr>
</tbody>
</table>

*τ: Time interval = 10 seconds, **L: Cell length = 500 ft

It should be noted that physical characteristics of the network and demand are presented in this paper are over-simplified to be able to better present the properties of the proposed approach. However, without loss of generality, the network can be easily extended to have a more realistic representation of a real evacuation process, e.g. realistic demand, a large network with longer cell lengths, probabilistic capacities for all cells with some additional computational effort. Vehicles are loaded onto network with numbers just sufficiently large to make the network congested so that existing capacity is fully utilized and the effect of capacity changes are easily visible. Likewise, loading is performed with a short time span to be able to track and represent the network flow changes.

All the cells on Route-3, which represent a major highway, are assigned to have more capacity than other cells in the network. Route-1 can be assumed to be a flood-risk road. It has less capacity but on the other hand, is shorter than other two routes. Route-2 can be assumed to be an alternative to Route-3. Route-2 has a smaller capacity than Route-3, and has a level of capacity loss probability. In other words, Routes-1 and 2 both have probabilistic capacities whereas Route-3 has fixed capacity. For simplicity, only two cells
(Cell # 7 and #8) in the network are assigned random capacity. The capacity decreases are assigned based on available lanes basis (lanes which are not flooded or blocked), e.g. 1,2… or all lanes are being used with predetermined probabilities. This kind of discrete approach is more appropriate in terms of practical concerns. A possible flood or incident on the roadway may block/flood all the lane(s) depending on the situation. In other words, a lane is either available, or unavailable. For calculating pLEPS, joint capacity probability distribution of cell #7 and #8 are found by assuming that they are independent. Although this assumption sounds counter-intuitive first, it does not make much sense to discuss the dependency of those probabilities since for a real life application, those probabilities will be determined by SLOSH etc. In other words, for real life applications, the joint probabilities are just inputs to the model and do not need any additional mathematical treatment.

Table 2 Probabilities Assigned for Available Number of Lanes

<table>
<thead>
<tr>
<th>Cell #</th>
<th>Available Number of Lanes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>0.4</td>
</tr>
<tr>
<td>8</td>
<td>0.1 (0.5*)</td>
</tr>
</tbody>
</table>

* Shows the increased probabilities for approaching hurricane

In the analysis, for each case the volume that uses specific route is recorded to detect the changes in the assignment. Different probability levels are set in equation (1), which can be interpreted as reliability of the forecasted capacity differences. As probability level (p) increases, the capacity constraints become less strict and flood risk is given less attention. Table 3 shows the percentage of traffic carried through corresponding routes, for our 3 different cases and different probability levels.

Table 3 Overall Traffic Flow Percentage on Each Route for Different Scenarios

<table>
<thead>
<tr>
<th>Routes</th>
<th>Case-1</th>
<th>Case-2</th>
<th>Case-3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>p=0.75</td>
<td>p=0.9</td>
<td>p=0.75</td>
</tr>
<tr>
<td>Route-1</td>
<td>44%</td>
<td>51%</td>
<td>41%</td>
</tr>
<tr>
<td>Route-2</td>
<td>24%</td>
<td>17%</td>
<td>26%</td>
</tr>
<tr>
<td>Route-3</td>
<td>32%</td>
<td>32%</td>
<td>33%</td>
</tr>
</tbody>
</table>

In the deterministic case, major flow is routed through Route-1, which is shorter than Route-3. Since Route-2 has the same length as Route-1, flow is close to Route-3 although the capacity is one third of Route-3. For the second case, cell #7 (Route-2) is assigned a higher probability for capacity loss compared to cell #8 (Route-1). There are two sets of percentages because of multiple optima. First solution is obtained with capacity loss in cell #7, thus the reduced flow on Route-2 is diverted to Route-1, which is less likely to lose capacity. For the second solution, cell #8 is assigned less capacity thus traffic is diverted to Route-2. For the 3rd case, the Route-1 capacity loss probability is dramatically increased after time step-8, which is basically the time that first loaded vehicles reach the destination. For strict case (p=0.75), less risk is taken and the traffic is diverted to Route-3, which is longer but safer. For p=0.9, the assignment does not regard the capacity loss probabilities as much as p=0.75 case and part of flow through Route-3 is directed towards Route-2. Overall, it can be said that the assignments through the cells with random capacity change considerably, affecting also Route-3 assignment, which is deterministic.
4. Conclusions

In this study, demand uncertainty in evacuation is discussed in terms of the changes in network performance along with the impact of different demand models. Traffic assignment during evacuation is also discussed, and a simple numerical example to emphasize the importance of incorporating probabilistic road capacity reduction phenomena is presented. It is shown that SO DTA assignment with probabilistic flow capacity constraints can lead to significantly changed flows over the network. Although the presented network is very simple, the results can be used as a first step to show the importance of stochastic considerations for a real life case. Remembering the problems faced by drivers during the hurricane, it is recommended to incorporate the impact of possible capacity reductions in evacuation modeling.

5. References


VALIDATION AND COMPARISON OF CHOICE MODELS

Giulio Erberto CANTARELLA¹, Stefano DE LUCA¹, Viviana FEDELE²

Abstract. In this paper a procedure for choice model validation and comparison is presented. Several benchmarking indicators are proposed within a general procedure, proposed to be adopted as common practice.

1. Introduction

It is widely recognized that a relevant part of transportation system analysis regards travel demand. The basic tool to carry out such an analysis is choice modeling. A systematic framework was first introduced in the well known book by Domencich and McFadden (1975). Since then, most effort in literature has been devoted over the years to choice model formulations, reviews in the books by Ben-Akiva and Lerman, 1985, Cascetta, 2001 and Train, 2003. On the other hand, less attention has been paid to methods for choice model validation and comparison against real data. It can be well the case that a very sophisticated choice model actually only slightly out-performs a simpler one. In this paper, after a formal introduction to specification and calibration of choice models against a sample of observed choices, a procedure for choice model validation and comparison will be presented. Several benchmarking indicators are proposed within a general procedure, proposed to be adopted as common practice, and resumed in the table below.

¹ Dept. of Civil Engineering – University of Salerno, Via Ponte don Melillo, 84084 Fisciano (SA) (Italy), g.cantarella@unisa.it
² Dept. of Civil Engineering – University of Salerno, sdeluca@unisa.it
³ University of Reggio Calabria, viviana.fedele@tin.it
The calibration and the validation have been carried out to four-mode choice on a real-size application, the comparisons concern two different modeling approaches: random utility theory vs. fuzzy utility theory. The above indicators have been computed both for calibration and hold-out samples.

2. **Calibration of a choice model: presentation of results**

Parameters of the utility function, $\beta$, as well as those of the choice function, $\gamma$, can be calibrated trying to reproduce a sample of observations of user choices (disaggregate) and/or demand flows (aggregate).

It is good practice to breakdown the available data set into a calibration sample and a hold-out one to be used for validation and comparison. The size of calibration sample can be obtained by progressively increasing the number of observations, randomly sampled out of the available data, and calibrating the choice model until obtained choice shares are close to the observed ones (up to an error threshold). At this aim a MultiNominal Logit can be used as benchmarking choice model.

If parameters of perceived utility distribution do not depend on utility attributes, a scale parameter cannot be identified, $\gamma_1$, and results of calibration are the ratios $\beta_j/\gamma_1$. Hence the values of utility parameters should be better presented as ratios w.r.t. one chosen as reference, $\beta_1$ (usually regarding monetary cost or LoS). This way comparison of utility parameters among different models is made independent of the scale parameter value.
From the above, results of calibration should be presented as in the proposed list below:

1. calibration and hold-out sample size, # parameters, # attributes,
2. scale parameter $\gamma_1$ definition, for example $\gamma_1 = \theta = (\sqrt{6}/\pi)\sigma$ for MNL,
3. reference utility parameter with its unit: $\beta_1 / \gamma_1$
4. other utility parameters with their units: $\beta_j / \beta_1 \quad j > 1$
5. other utility parameters, for example regarding non-linearity
6. choice function parameters if any (apart the scale parameter, $\gamma_1$): $\gamma_i \quad i > 1$

3. Validation and comparison of choice models

It should be noted that even though the calibrated values of the model parameters are optimal, the resulting model may yet perform poorly. The first analysis deals with the values of utility parameters, regarding their signs and possibly their statistical significance. Then, the effectiveness of the model resulting from the calibration stages must be explicitly analyzed, that is, the validation stage should be carried out, before the model can be effectively used to generalize. Validation may be carried out through some benchmarking indicators applied both to the calibration and the hold-out sample.

_Generate indicators_:

- Simulated vs. observed shares for each choice alternative
- Simulated vs. observed choice fractions for each user

  - **MSE** $= \sum_i \sum_k (p^\text{sim}_{k,i} - p^\text{obs}_{k,i})^2 / N_{\text{users}} \geq 0$ mean square error between the user observed choice fractions and the simulated ones, over the number of users in the sample, $N_{\text{users}}$. SD is the corresponding standard deviation, which represents how the predictions are dispersed if compared with the choices observed.

  - **MAE** $= \sum_i \sum_k |p^\text{sim}_{k,i} - p^\text{obs}_{k,i}| / N_{\text{users}} \geq 0$ mean absolute error.

  - **FF** $= \sum_i p^\text{sim}_i / N_{\text{users}} \in [0,1]$, with FF = 1, when the model perfectly simulates the choice actually made by each user, say with $p^\text{sim}_i = 1$. It results $\text{MAE} = 2 \times (1 - \text{FF})$.

- Log-Likelihood tests

All the above indicators should be compared with a reference value obtained assuming all simulated choice fractions equal to one over the number of choice alternatives.
Analysis of clearness of predictions

Commonly this analysis is carried out through the %right indicator. This index, very often reported, is rather meaningless if the number of alternatives is greater than two.

A really effective analysis can be carried out through the below indicators:

- \( \%_{\text{clearly right}}(t) \) percentage of users in the sample whose observed choices are given a probability greater than threshold \( t \) by the model (50% for \( t = 60\% \) in figure below);

- \( \%_{\text{clearly wrong}}(t) \) percentage of users in the sample for whom the model gives a probability greater than threshold \( t \) to a choice alternative different of the observed one (16% for \( t = 60\% \) in figure below);

- \( \%_{\text{unclear}}(t) = 100 - (\%_{\text{clearly right}} + \%_{\text{clearly wrong}}) \) percentage of users for whom the model does not give a probability greater than threshold \( t \) to any choice (34% for \( t = 60\% \) in figure below);

These indicators can be plotted against threshold in the range 0.5 - 1.0.

Analysis of ASA role for mode choice models

When modeling transportation mode choices it is common practice to have an alternative specific attribute, \( ASA_k \), or modal constant for each alternative \( k \). ASA’s may play a relevant role to obtain significant utility parameters, but the greater they are the poorer the model is expected to perform when applied to other choice scenarios.

Two types of indicators can be used to analyze the role of ASA’s:

- \( ASA_{\text{effect}} \) can be measured by any distance between choice probabilities computed with or without ASA’s.

- \( ASA_{\text{impact},k} \) can be defined for each choice alternative \( k \), but the one taken as reference, usually \( k = 1 \). Once the utility parameters have been calibrated, the frequency distribution of difference \( (v_k - ASA_{k,1}) - v_1 \) over the users in the sample is defined. Then, the relative number of users with a value of \( |(v_k - ASA_{k,1}) - v_1| \) less than or equal to \( |ASA_{k,1}| \) is computed.
References


DECOMPOSITION METHODOLOGY TO SOLVE TAXI PLANNING

Ángel G. MARÍN

Abstract. The Taxi Planning problem studies the flight ground movements from the aircraft gate to the runway and vice versa. It is modelled using a binary multicommodity flow network and the flow capacity constraints and other side constraints. New decomposition methodologies have been adapted to take advantage of the model structures. The computational tests have been run using actual data from Madrid-Barajas airport. They have been aimed at comparing Lagrangean decomposition with Branch and Bound.

1. Introduction

Taxi Planning (TP) may be considered under different planning horizons, at operative level (assumed in this paper) short planning periods (PP) are considered with planning times of 15 to 30 minutes, so the computational time is critical and must be less than 1 minute.

The airport traffic may be classified as Departure Traffic (DT) and Arrival Traffic (AT). DT is the flight traffic from the aircraft gates to a designated runway access point. AT is the flight traffic from the runway exits to an assigned stand.

[4] was the first paper with the TP formulation in terms of a binary multicommodity network flow model. They defined the previous formulation and developed Branch and Bound (B&B) and Fix and Relax (F&R) to be applied to TP. In [2] Marín redefines new approaches to TP model and present computational results using real data from Madrid-Barajas airport. [3] developed TP and studied the computational results comparing B&B and F&R. In the text it is frequently mentioned that some aspects of the formulation are not included here, because the complete TP model formulation would require a lot of pages to be described in detail, so a reduced but basic formulation of the decomposition adaptation to TP model is showed.
2. Taxi planning model

TP is modelled as a multicommodity binary network flow model using a time-space network to represent the conflicts between aircrafts and the congestion generated when several flights try to use the limited resource capacity of an airport.

A directed graph $G (N, A)$ is used to define TP model. $N$ represents the set of airport “nodes” and $A = \{(i, j), \forall i, j \in N\}$ the set of “arcs”. A node $i \in N$ can be a parking, a holding area, a junction or intersection of two or more taxiways, or a runway header or exit gate. An arc $(i, j) \in A$ connecting nodes $i, j$, typically represents a physical taxiway, but in TP they are also used to represent entrance- and exit-ways to-from a stand, and to represent intermediate segments of a taxiway where an aircraft may need to stop.

In order to take account the time within the planning period (PP), each $t \in T$ represents a time mark or “period” of equal length “$t_p$”. Typically in steps (“$t_p$”) of 30 seconds in-between periods. In that case, $t = 0$ would represent the current time and $t = 60$ would represent 30 minutes afterwards.

The TP model is based on the definition of a time space graph, $G^* (N^*, A^*)$, where $N^*$ is obtained replicating $|T|$ times the previous node set, $\{(n, t) : \forall n \in N, \forall t \in T\}$ and $A^*$ are the links defined between them, $\{(i, t), (j, t'), \forall i, j \in N, \forall t, t' \in T\}$.

The TP optimal variables are binary. They are referred to as the space-time link set:

$E^w_{it} = 1$, if the aircraft “$w$” uses a wait link ($A^w$) defined by the wait node “$i$” in the period “$t$”; and 0, otherwise.

$X^w_{ijt} = 1$, if the aircraft “$w$” uses a moving link ($A^M$) departing from node “$i$” to the node “$j$” at period “$t$”; and 0, otherwise.

Each flight “$w$” is defined by an origin node “$o(w)$”, a destination node “$d(w)$”, and an origin time “$t(o(w))$”. In the time space graph the origin is a single node “$\{o(w), t(o(w))\}$”, and the destination is a set of space time nodes defined by: the set of destination nodes “$d(w)$” at different periods and the intermediate nodes between the origin and the destination at the end of the PP, whether the flight cannot arrive at its destination during the PP. For this second group of flights it must be estimated the time needed ($r^w_{i}$) for the aircraft “$w$” to arrive at its destination from the node “$i$”, that is the node where the aircraft is located at the end of the PP. $r^w_{i}$ is obtained using an external shortest path algorithm which is run prior to TP.

TP minimizes the total routing time for all the flights. The objective function, $F$, is defined by the time periods spent to route all the aircrafts:

$$z_{TP} = F(X, E) = \sum_{w \in W} \sum_{t \in t(w)} \lambda^w \left( \sum_{i, j, t \in A^w} t_{ij} X^w_{ijt} + \sum_{i, t \in A^w} E^w_{it} \right) + \sum_{w \in W} \sum_{i \in N^w} r^w_{i} E^w_{iT}$$
where $\lambda^w$ is the priority of the aircraft “$w$”, $W^W$ is the wait node set, and $W$ is the aircraft set. Priority is used if some aircrafts have preference in the event of conflict or congestion.

The flow conservation constraints at nodes are:

$$E_{it}^w + \sum_j X_{(j,i),t-1:t+1}^w = E_{it+1}^w + \sum_j X_{(i,j),t+1}^w, \forall w, i, t$$

Consider $j_i$ as the time it takes to get from node $j$ to an adjacent node $i$. The above equations establish that, in order for aircraft $w$ to be allowed to wait at node $i$ in period $t + 1$, or to start moving from $i$ to $j$ in period $t + 1$, the aircraft must have been waiting at that node $i$ in the current period $t$, or moving from some adjacent node $j$ to $i$ in period $t - j_i + 1$ (to guarantee that it arrives at node $i$ in period $t$).

These two examples show how the different variables can be combined in order to create a consistent set of restrictions that represents the physical and logical constraints of the problem. Although the purpose of this paper is not to specify all the model formulation, the following list describes the rest of equations:

- Initial conditions: Aircraft origin and time at origin
- Final conditions: Runway for departing flights, and stand for arriving flights. These are elastic conditions, because some aircraft will be en-route at the end of a given planning time.
- Overtaking: Avoid aircraft overtaking while taxiing
- Capacity conditions: Enforce limits at holding areas and taxiways
- Runways: Ensure runway is not used by two aircraft simultaneously, and avoid conflicts if more than one runway header exist
- Stands: Avoid the arrival of landing traffic to a stand and/or stand exit area if it is still occupied by a departing flight
- Other logical conditions: Avoid the use of stand areas for “waiting” or “transit” purposes by any other traffic than the aircraft departing and/or arriving at that stand; persist on immediate pre-specified route for aircraft on the taxiways at the beginning of the planning period; etc.

The formulation of the TP may be shortened by calling “$a$” to the space time links, and calling $x_a^w$ to the binary flow variables (These variables are equivalent to the previously defined: $X_{ij,t}^w$ or $E_{it}^w$). Taking this into account, the following variable vectors defined by $x^w = \{x_a^w\}$ and $x = \{x_a^w\}$ are used to define the short TP formulation:

$$TP: \quad z_{TP}(x) = \min_{x_a^w \in \{0,1\}} \sum_{a \in A} \sum_{t \in T} c_a x_a^w$$

s.t.

$$A^w x^w = b^w, \forall w \in W \quad \text{(Node flow conservation)}$$

$$\sum_{w \in W^w} x_a^w \leq q_a, \forall a \in A \quad \text{(Link capacity)}$$

735
Where the “Node flow conservation constraints” are equivalent to the previous flow conservation ones, and the “Link capacity constraints” are the remaining constraint mentioned above. The TP may be considered decomposing the constraints into easy and difficult ones. The easy constraints are the node flow conservation constraints, which are separable by nodes and flights. The difficult constraints are the link capacity ones; they are only separable by link.

3. Taxi Planning: Lagrangian Decomposition.

TP may be solved by Branch and Bound (B&B), but it is not efficient in terms of computational time given the large number of binary variables of the model and the necessity to solve it very fast, so other methodologies have been developed. Fix and Relax is other methodology considered to solve TP, in effect it is in general about a 20 percent more efficient than B&B, but it is not enough for the velocity needed to solve TP in its context. Marín (2005).

To solve TP, different Lagrangian Relaxation and Decomposition options have been previously analyzed. As result of this analysis (which is not included here for space limitations), the Lagrangean Decomposition (LD) has been the best relaxation methodology chosen.

The LD introduces a new variable, which is a copy of the existing one, and the corresponding constraint that establishes the equality between both copies of the variables (variable splitting). LD is used to group the constraints into two groups depending on the copy of the variable used to define them, so TP may be described by copies (x and y) of the original optimisation constraints in the following way:

\[
\begin{align*}
    z_{tp} &= \min_{x, y} \left( \gamma_{xsm} \sum_{w} \sum_{a} c_w x^w + (1 - \gamma_{xsm}) \sum_{w} \sum_{a} c_w y^w \right) \\
    \text{subject to} & \quad A^x x^w = b^w, \quad \forall w \in W \\
                     & \quad \sum_{w} y^w_a \leq q_a, \quad \forall a \in A^* \\
                     & \quad x^w = y^w_a, \quad \forall w \in W, \forall a \in A
\end{align*}
\]

Where \( \gamma_{xsm} \) is the weight of the “x” sub model (XSM) objective function term; it can take values between \((0,1)\), and the weight of “y” sub model (YSM) objective function term is \((1 - \gamma_{xsm})\). The last constraints are the variable splitting.

The constraints are grouped on XSM and YSM defining submodels with some relaxed constraints. The master model (MM) is the Lagrangian dual of TP and it takes decisions to update the Lagrange multipliers of the variable splitting constraint.
The multipliers are applied at each iteration to redefine the submodels, and the submodel solutions are used by the MM to update the multipliers, defining a cycle, which terminates with an optimal or an $\varepsilon$-optimal solution.

The TP may be also defined using the TP Lagrangian Dual Problem (TP_LDP), which is defined as follows:

\[
\text{TP_LDP: maximize } \quad L(\lambda) \\
\text{subject to: } \quad \sum_{\forall \in W} A^w x^w = b^w \quad \forall w \in W
\]

Where $L(\lambda)$ is the sum of two terms: $L_x(\lambda)$ and $L_y(\lambda)$. The first is the objective function of the SM, where the flow node conservation constraints are considered, $XSM(\lambda)$. The second is the objective function, where link capacity and the take-off separation constraints are considered, $YSM(\lambda)$.

\[
\begin{align*}
XSM(\lambda): & \quad L_x(\lambda) = \min_{\forall \in [0,1], \forall \in A} \sum_{w \in W} \sum_{\forall \in A} (\gamma_{SM} c_a + \lambda_a^w) x^w_a \\
& \quad A^w x^w = b^w \quad \forall w \in W
\end{align*}
\]

$XSM(\lambda)$ is a multicommodity network flow problem separable by commodity, and the subproblem of each commodity may easily be solved by some shortest path algorithms.

\[
\begin{align*}
YSM(\lambda): & \quad L_y(\lambda) = \min_{\forall \in [0,1], \forall \in A} \sum_{w \in W} \sum_{\forall \in A} (1 - \gamma_{SM}) c_a - \lambda_a^w y^w_a \\
& \quad \sum_{w \in W} y^w_a \leq q_a \quad \forall a \in A
\end{align*}
\]

The YSM($\lambda$) is difficult to solve, so we will need to define an heuristic to solve it.

4. Lagrangean Decomposition algorithm and computational experiments

The TP Lagrangean Decomposition Algorithm TP_LDA is defined by iteratively solving the TP_LDP and the SM. To solve the TP_LD the Subgradient Method (SGM) has been used. The SGM establishes a sequence of dual variables $\lambda^k$, trying to follow an ascent direction from $L(\lambda)$.

The subgradient of $L(\lambda)$ at iteration “$k$” is $sg_a^k = x_a^k - y_a^k$, where $x_a^w(\lambda^k)$ and $y_a^w(\lambda^k)$ are the XSM($\lambda^k$) and YSM($\lambda^k$) solutions at iteration “$k$”. The dual sequence of $\lambda$ is defined by the direction $d^k = sg^k$ of the following way: $\lambda^{k+1} = \lambda^k + \theta^k sg_{a,w}^k$. 

737
Where \( \theta^k \) is the step size in this iteration. An extension of the subgradient method is justified in terms of the KKT optimal conditions [1]. In this case, the step size is defined by

\[
\theta^k = \eta^k \frac{UB^k - L(\lambda^k)}{\|sg^k\|^2}.
\]

(A coefficient \( \eta^k \) is added to step expression)

TP Lagrangian Decomposition Algorithm:

- **Initialization**: Set \( k = 0, \lambda^0 = 0, LB^0 = 0, UB^0 = \infty \).
- **While (k<k_max)**
  - Solve XSM(\( \lambda^k \)) and YSM(\( \lambda^k \)). Let \( x^k \) be the optimal solution of XSM(\( \lambda^k \)) and \( y^k \) the solution of YSM(\( \lambda^k \)). Let \( L_x(\lambda^k) \) and \( L_y(\lambda^k) \) be the values of the respective objective functions and set \( L(\lambda^k) = L_x(\lambda^k) + L_y(\lambda^k) \).
  - Update the lower bound \( LB^k = \max \{ LB^{k-1}, L(\lambda^k) \} \).
  - Find an TP feasible solution \( x_f^k \) solution. Update UB^k:
    \[
    UB^{k+1} = \min \{ UB^k, z_{TP}(xf^k) \}.
    \]
  - If some termination criteria are fulfilled, stop.
  - Update \( \lambda : \lambda^{k+1} = \lambda^k + \eta^k \frac{UB^k - LB^k}{\|sg^k\|^2} \).

Where \( K_{\max} \) is the maximum number of iterations. The algorithms used to solve the SM and to obtain the TP feasible solution will be defined in the paper but in this extended abstract are not space. The same happened with the computational experiments.

In the experiments have been observed that the LD code is more efficient than the B&B, in some items more than 30 times better. They have been implemented with actual data from Madrid-Barajas airport. Other advantages of LD as opposed to using B&B (Cplex-Gams) are that the LD code was written in C++, which guarantees the flexibility and the independence concerning any software constraints.

5. Conclusions

The Taxi Planning model is a new computational tool to represent the conflicts and the airport congestion associated with taxiways in the big airports. The model is defined using Mathematical Programming by a large space-time multicommodity network with node and arc capacity constraints, and other side constraints.

A Langrangean Decomposition algorithm has been defined to solve it. Its efficiency solving the submodels and implementing the subgradient let us to solve the model in a few
seconds. This result may be compared using Branch and Bound (Gams/Cplex), which needed several minutes to solve the large cases.

References


ALLOCATION PLANNING OF HANDLING DEVICES
FOR BARGES UNLOADING

Milorad VIDOVIC¹, Katarina VUKADINOVIĆ²

Abstract. Providing efficient and cost effective service of loaded river barges needs appropriate allocation plan for handling equipment. In order to utilize handling facilities efficiently, and to minimize waiting time of barges, it is necessary to consider both: assignment of handling equipment to unloading locations, and order of servicing different unloading locations. This paper formulates the problem of handling devices allocation, and proposes a heuristics solution.

1. Introduction

Gravel distribution by inland water transportation includes three main phases: loading of gravel by a suction dredger into barges, transport of gravel to the ports or unloading locations, and unloading of gravel by a handling facility that usually consists of pontoon mounted crane and belt conveyor. Because of high costs, a number of handling facilities is usually relatively small, and requires successive relocation of handling equipment between different unloading locations. Accordingly, providing efficient and cost effective service of loaded river barges needs appropriate allocation plan for handling equipment, which means defining sequence of unloading locations that should be served by each handling device.

In order to utilize handling devices efficiently, and to minimize the waiting time, as well as the total service time of barges, it is necessary to consider assignment of handling equipment to unloading locations, and sequencing orders.

It is important to note that this problem is not only applicable to gravel distribution. It can be applied to other transportation modes, where resources are cars, trucks, containers,...etc. However, in this paper, the problem and objective function are adapted to describe gravel distribution processes realized by river barges and corresponding handling equipment.

¹ Faculty of Transport and Traffic Engineering, University of Belgrade, Vojvode Stepe 305, 11000 Belgrade, Serbia, mvidovic@sf.bg.ac.yu
² Faculty of Transport and Traffic Engineering, University of Belgrade, Vojvode Stepe 305, 11000 Belgrade, Serbia, k.vukadinovic@sf.bg.ac.yu
There have been several approaches and studies related to the described problem. One group of related problems covers fleet sizing and allocation problem, where objective is to find answers to the following two questions: how many units should a fleet contain, and how to redistribute units that are not used in a given location among the locations where needed. Units may be cars, trucks, railcars as well as containers and material handling equipment. For both the fleet sizing problem and the empty vehicle reallocation problem, voluminous literature exists ([1], [3], [4], [6], [9]...). The approaches to solve the problems extend from linear and non-linear network programming models via queuing theory models and dynamic programming models, approximation strategies for solving multistage stochastic allocation problems, to simulation, and evolutionary optimization. However, no closed-form solutions are available. The majority of papers deals with algorithms that define an approximate solution for the discrete-time case with known demand and infinite transportation capacity. Another group of related problems are “the berth allocation problem”, i.e. “the berth planning (scheduling) problem” ([7], [8], [10], [11], [13]), and “the crane assignment (scheduling) problem” ([2], [5], [14]). In case of the berth planning problems, the aim is to construct the berth schedule, and the calling schedule of vessels, while crane schedule covers crane assignment to vessels. Both problems are mostly solved by heuristics algorithms, under the objective of delay costs minimization.

Nature of the problem studied here is much closer to the second group of related problems, particularly to “the berth allocation problems”, in case of discrete quay location space. Namely, in the berth allocation problem, berth is server that can serve one or more vessels simultaneously, and vessels are clients to be served. Vessel loading-unloading operations occupy berth, and cause delay of all other vessels waiting for berthing and service. Also, total delay for one vessel comprise loading-unloading time of all preceding vessels. Similarly, in case of unloading gravel, a barge arrived at certain port or unloading place waits for handling facility until all preceding barges, located in other ports or unloading locations are unloaded. However, while in case of the berth allocation problem, vessel transfer time from/to vessel may be neglected when compared with vessel loading-unloading time, handling facility transfer time between unloading locations must be taken into consideration. This fact increases problem complexity, and makes main difference between the berth allocation planning problems, and the gravel unloading problem studied here.

Hence, contribution of this paper is in the problem formulation, which differ from previous formulations of both, berth planning, and crane scheduling problems, and in solution approach proposed.

Remaining part of the paper is organized as follows. Section 2 puts forward the problem formulation, and section 3 proposes solution approach. In the section 4, proposed solution approach is tested on numerical examle. Section 5 gives some concluding remarks.

2. Problem formulation

This paper addresses the problem of handling devices allocation (HDAP) to loading/unloading tasks which are spatially distributed on transportation network.

The problem may be introduced in following way. For a given collection of barges find a set of assignments to minimize the sum of the service times including waiting for service
and handling devices transfer times. The problem of this type may be considered as
dynamic handling devices scheduling problem, where tasks service ready times are known
after the beginning of the planning interval, or as static problem, where all tasks are already
known when the scheduling plan is determined. In this paper, only static problem (SHDAP)
is studied.
In this case, given a collection of N tasks (barges) are ready to be served at known
network nodes \( j, k = 1, ..., N \) \( \forall j, k \in B \), at the beginning of the planning horizon, where
\( B \cup \{0\} \) is set of nodes, representing ports and unloading locations , as well as barges, and
node 0 is handling facility depot. There are I handling devices \( (1 \leq N) \), \( i = 1, ..., I \) \( \forall i \in H \),
where H is set of handling facilities. Each task (barge) has its own time of appearance \( A_k \).
In general case, each handling device (i) has known productivity \( p_{ik} \), dependant on the
barge to be served. Since each barge has known capacity of \( q_k \), task (k) completion time \( C_{ik} \)
depends on the productivity of device used for the service. That is, \( C_{ik} = q_k / p_{ik} \). Transfer time
\( t_{ijk} \) of handling device (i), between nodes (j) and (k) depends on the distance between nodes
\( d_{jk} \), travel speed \( V_i \), and on the handling device preparation time \( D_i \) after handling of barge
(j), as well as on the preparation time \( E_i \) before handling barage (k). That is, \( t_{ijk} = d_{jk} / V_i + D_i + E_i \).
To formulate SHDAP, let \( \{S(\theta_i, n_i)\} \) be a set which represents \( \theta_i \)-th
sequence of \( n_i = S(\theta_i, n_i) \) nodes (barges) served by handling device (i), arranged, and
renumerated by service order, where \( |S(\theta_i, n_i)| \) is the cardinality of the set \( S(\theta_i, n_i) \). Each
sequence starts from node 0, and represents one permutation of the \( n_i \) nodes assigned to the
handling device (i), i.e. \( \theta_i = 1, ..., \Theta_i \), where \( \Theta_i = n_i! \). Therefore, each sequence of handling
device (i) comprises nodes assigned to it, and its order of servicing. Defined sequences
should satisfy following conditions:

\[
0 \leq n_i \leq N, \forall i \in H, \sum_{i=1}^{I} n_i = N \quad (1)
\]

\[
S(\theta_i, n_i) \subseteq B, \forall i \in H \quad (2)
\]

\[
S(\theta_p, n_p) \cap S(\theta_q, n_q) = \emptyset, \forall p, q \in H \quad (3)
\]

Total aggregate time \( T(\theta_i, n_i) \) that all \( n_i \) barges assigned to handling device (i), and
served in order \( \theta_i \), defined by the sequence \( S(\theta_i, n_i) \) spend in the system, both, while waiting
for service, and unloading, when handling device starts from the depot node 0, is given by:

\[
T(\theta_i, n_i) = \sum_{j=0}^{n_i-1} (n_i - j) \left( t_{i,j,i+1} + C_{i,j+1} - A_{j+1} \right) \quad (4)
\]

Obviously, total aggregated time \( T_{TOT} \), that all barges spend in the system, when each
handling device (i) serve \( n_i \) assigned nodes ordered in \( \theta_i \)-th sequence, then is given by:

\[
T_{TOT} = \sum_{i=1}^{I} T(\theta_i, n_i) = \sum_{i=1}^{I} \sum_{j=0}^{n_i-1} (n_i - j) \left( t_{i,j,i+1} + C_{i,j+1} - A_{j+1} \right) \quad (5)
\]

Hence, in order to find minimal aggregated time \( T_{TOT} \) that all barges spend in the
system, it is necessary to find optimal relocation plan of handling devices for unloading of
barges, which is determined by optimal sequencing vector \( \|\Theta_i\| \), for all handling devices
under conditions (1), (2), and (3):
By introducing binary variable $\theta$:  
$$
\begin{cases}
\theta = 1 & \text{when handling device } i \text{ serves } 0\text{th permutation of } n_i \\
0 & \text{otherwise}
\end{cases}
$$

(7)

objective function may be simple derived from (5), and the problem may be formulated as a three-dimensional assignment problem – formulation F1:

$$
\text{Minimize } \sum_{i=1}^{I} \sum_{n=0}^{n_i} \sum_{j=0}^{n_i} x_{i,n,j} \left( \sum_{h=0}^{n_i-1} (n_i - j_h) \left( t_{i,j_h,i,j} + C_{i,j_h,j} \right) - A_{i,j} \right)
$$

(8)

Here, $\theta$ denotes ordinal number of permutation, and indices $(j)$, with sub indices $\theta$, denotes $j$-th in the order node (barge) in $\theta$-th permutation, served by handling device $(i)$ in the sequence of $n_i$ nodes. The constraints are given by expressions (1), (2), (3), and (7).

Another way to formulate SHDAP is based on similarity with the static berth allocation problem SBAP ([7]). In the SBAP objective is to minimize the sum of waiting times for the availability of the berth assigned to each ship, plus the handling time it spends at the berth, while there are $I$ berths, and $N$ ships in the system. Similarly, in the SHDAP, there are $N$ barges (arrived in the $N$ nodes), waiting for the service from one of $I$ handling facilities, which are able to operate in each node. While berths are static and wait for the ships to be assigned, handling devices are assigned to barges waiting in nodes. Berth is released after assigned ship is served, and it is able immediately to serve next ship due to defined service order. However, barge should wait for an arrival of the assigned handling device, even though handling device is released, because of transfer time. From there, main difference between SBAP and SHDAP lies in fact that beside the handling time, arbitrary barge should wait for the transfer time, required for preparing and moving device from a preceding node and a next assigned node.

Considering the above described difference, SHDAP could be formulated as follows – formulation F2:

$$
\text{Minimize } \sum_{i=1}^{I} \sum_{n=0}^{n_i} \sum_{j=0}^{n_i} \sum_{s=0}^{n_i} x_{i,n,j,k,s} \left( (n_i - s + 1) \left( t_{i,j,k} + C_{i,k} \right) - A_{i,k} \right)
$$

(9)

Subject to (1), (2), (3), and

$$
\sum_{i=1}^{I} \sum_{n=0}^{n_i} \sum_{j=0}^{n_i} x_{i,n,j,k,s} = 1, \quad \forall k \in B
$$

(10)

$$
\sum_{n=0}^{n_i} \sum_{j=0}^{n_i} \sum_{s=0}^{n_i} x_{i,n,j,k,s} = n_i, \quad \forall i \in H
$$

(11)

Constraint set (10) ensures that every barge must be serviced by some handling device in any order of service. Constraint set (11) enforces that each handling facility services assigned sequence, up to one barge at any time.

Binary variable $x_{i,n,j,k,s}$ takes following values:

$$
\begin{cases}
1, & \text{when device } (i) \text{ serve barge } (k) \text{ after barge } (j) \text{ as a } s - \text{th in order of } n_i \text{ nodes} \\
0, & \text{otherwise}
\end{cases}
$$

(12)
3. Solution approach

In order to solve SHDAP, the following heuristic algorithm is proposed. The CLASORD (CLustering ASsignment ORDering) heuristic algorithm is based on the idea of classifying nodes into clusters whose number is equal to the number of handling devices, assigning each of handling devices to those clusters, then sequencing nodes (barges) in each cluster in non decreasing order of its service times, and finally improving the obtained solution. Classifying nodes into clusters is based on K-Means Clustering technique ([12], source [15]), where few distance measures are analyzed (distance between nodes, transfer time, unloading time, and compound time measures). However, when it is possible, clusters may be defined intuitively. Sequencing of nodes in each cluster is defined by ordering service times \( \tau_{ijk} \) in nondecreasing order. That is, for any arbitrary node \((k)\), belonging to the cluster \((i)\), service time \( \tau_{ijk} \) is calculated as \( \tau_{ijk} = t_{ijk} + C_{ik} \), where indices \((j)\) denote preceding nodes to the node \((k)\). Assignment of clusters to handling devices is solved as classical two dimensional assignment problem, where costs were defined as sum of weighted service times. Finally, it is experimented with improvements of solution obtained through application of previous three steps, by exchanging nodes between clusters.

In general, CLASORD heuristics application is composed of the following steps:

CLUSTERING PHASE

1) Place I points into the space represented by the nodes (barges) that are being clustered. (creating initial centroids).
2) Assign each node (barge) to the group that has the closest centroid.
3) When all nodes have been assigned, recalculate the positions of I centroids.
4) Repeat Steps 2 and 3 until the centroids no longer move

ORDERING AND ASSIGNMENT PHASE

5) For all nodes assigned to a cluster, calculate service time \( \tau_{ijk} \) when nodes are served by handling devices \((i)\):
   i) Starting from depot \((j=0)\) find closest nodes \((k)\) so that is satisfied
   \[ \tau_{ijk} = \min(\tau_{i01}, \tau_{i02}, ..., \tau_{i0k}) \]
   ii) Set \(j=k\), and find next closest node to \(j\)
   iii) Repeat ii) until all nodes in the cluster are ordered, denoting order of nodes in the cluster \(z\), when served by device \(i\), as \(\theta_{zi}\)
6) Calculate the total waiting time \(T(\theta_{zi})\) of all barges in the cluster by using expressions (4)
7) Repeat Steps 5 and 6 for all of I clusters, and all of I handling devices
8) Assign handling devices to clusters by solving the two dimensional assignment problem
9) Keep record of the solution

SOLUTION IMPROVEMENT PHASE

10) Exchange nodes between clusters, accordingly to predefined rule
11) Repeat Steps 5 to 9
12) If algorithm’s interrupting condition is satisfied then STOP, otherwise return to Step 10
4. Numerical Example

In order to demonstrate the proposed heuristic approach, the following example was tested. There are three handling devices (I=3) and seven nodes (barges) (N=7) to be served. Handling devices move at the same speed, which is 5 km/h. Their unloading productivities $p_{ik}$ are 100 tons/h, 150 tons/h and 200 tons/h respectively.

The assumption is that nodes have equal demand of 1000 tons, which also equals barges’ capacity $q_k$. River and canals network, with distances between network nodes is given in Figure 1 below.

Handling devices are placed at the depot node 0. As a first step of the algorithm the following clusters are formed: C1 which contains nodes 1, 2 and 3; C2 which contains nodes 4 and 5, and C3 which contains nodes 6 and 7.

Service order of nodes defining, in the ordering and assignment phase (step 5), requires the following calculation:

Cluster 1 served by handling device 1

$$
\tau_{101} = t_{101} + C_{11} = d_{01}/v_1 + q_1/p_1 = 90/5 + 1000/100 = 28
$$
$$
\tau_{102} = t_{102} + C_{12} = d_{02}/v_1 + d_1/p_1 = 50/5 + 1000/100 = 20
$$
$$
\tau_{103} = t_{103} + C_{13} = d_{03}/v_1 + q_3/p_1 = 100/5 + 1000/100 = 30
$$

The first node (k) in C1, to be served by handling facility 1 can be determined from:

$$
\min(\tau_{101}, \tau_{102}, \tau_{103}) = \tau_{102} \Rightarrow k = 2
$$

To define the next node in the C1, to be served after the node 2 by handling device 1, it is needed to calculate:

$$
\tau_{121} = t_{121} + C_{11} = d_{21}/v_1 + q_1/p_1 = 40/5 + 1000/100 = 18
$$
$$
\tau_{123} = t_{123} + C_{13} = d_{23}/v_1 + q_3/p_1 = 50/5 + 1000/100 = 20
$$

The next node (k) in C1, to be served after the node 2 is determined from:

$$
\min(\tau_{121}, \tau_{123}) = \tau_{121} \Rightarrow k = 1
$$
From the previous calculation, order of nodes in C1, when served by handling device 1 should be \( \theta_{11} = \{2, 1, 3\} \), which is the same, in this example, for the C1, for handling devices 2, and 3 as well. Hence, \( \theta_{11} = \theta_{12} = \theta_{13} \).

By using expression (4), waiting time of barges \( T(\theta_{11}) \) may be calculated as follows (it is assumed that \( A_k = 0, k = 1, \ldots, 7 \)).

\[
T(\theta_{11}) = 3d_{12} / v_1 + 3C_{12} + 2d_{21} / v_1 + 2C_{11} + d_{13} / v_1 + C_{13} = \\
3 \cdot 50 / 5 + 3 \cdot 1000 / 100 + 2 \cdot 40 / 5 + 2 \cdot 1000 / 100 + 90 / 5 + 1000 / 100 = 124
\]

Similarly, by applying the same approach it is obtained that \( \theta_{21} = \theta_{22} = \theta_{23} = \{4, 5\} \), and \( \theta_{31} = \theta_{32} = \theta_{33} = \{6, 7\} \). Waiting times of barges \( T(\theta_{zi}) \) in all clusters, when clusters are served with different handling devices, determined by the same approach, are shown in the table 1 below.

<table>
<thead>
<tr>
<th>Handling device</th>
<th>Clusters</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>124</td>
<td>46</td>
</tr>
<tr>
<td>2</td>
<td>104</td>
</tr>
<tr>
<td>3</td>
<td>(93)</td>
</tr>
</tbody>
</table>

Table 1. Waiting times of barges \( T(\theta_{zi}) \)

Optimal assignment of handling devices to clusters, when considering waiting time of barges, according to the step 8, is then obtained by solving the two dimensional assignment problem. In this example two optimal assignments with the same total waiting time of 188 hours were obtained (circles, and squares in the table 1). One solution is: device 1 serves C3, device 2 serves C2, and device 3 serves C1. The second solution is: device 1 serves C2, device 2 serves C3, and device 3 serves C1. Service order of nodes is given above by sequences \( \theta_{zi} \).

With the objective to improve solution, C2, and C3 were changed in the following way. Node 4, from C2 has moved to C3, where after recalculations in steps 5 to 8, two new solutions with the same total waiting time of barges (188 hours) were obtained. First solution assignment is: device 1 serves changed C2, device 2 serves changed C3, and device 3 serves C1. The second solution assignment is: device 1 serves changed C2, device 2 serves C1, and device 3 serves changed C3. Obviously, new optimal solutions were determined by nodes exchange step, but didn’t improve the total waiting time of barges, although in this case time was influenced by input values.

Finally, it should be mentioned that the heuristic can be easily implemented to the problems where handling devices are located in different nodes, not in the same depot.

5. Conclusions

It can be concluded that the proposed algorithm is very simple for application, and gives good solutions in a very short time, which suggest a possibility for practical use. Since the CLASORD heuristic is promising, and gives encouraging results, it is necessary to continue with its implementation to additional examples, particularly to larger problems (for example 20 nodes), which could be more appropriate as typical dimension of the real
barges’ unloading problem. Another research direction is in the heuristic performances
determination, and its adoption to real system requirements, particularly in the context of
node priorities. Additional research opportunity is a dynamic problem solving.

References

Transportation Science 25, 19–45, 1991
European Journal of Operational Research 144, 83–107, 2003
random arc capacities, with an application to dynamic fleet management. Operations 
Research 44, 951–963, 1996
Operational Research 122, 272–288, 2000
no.3, 159-175, 1989
[8] A. Imai, E. Nishimura, S. Papadimitriou, Berth allocation with service priority 
Transportation Research Part B 37, 437–457, 2003
with moving resources: Application to the fleet sizing and allocation problem. Int. J. 
Production Economics 81–82, 443–459, 2003
Transportation 26, 45-60, 1992
Observations, Proceedings of 5-th Berkeley Symposium on Mathematical Statistics
and Probability, Berkeley, University of California Press 1, 281-297, 1967
system by genetic algorithms, European Journal of Operational Research 131, 282–
292, 2001
UNIVARIATE AND MULTIVARIATE ARIMA MODELS FOR AIR TRANSPORT DEMAND ESTIMATION: A STUDY ABOUT A REGIONAL AIRPORT

Alberto Andreoni, Maria Nadia Postorino

Abstract: In this work, both univariate and multivariate air transport demand ARIMA models are proposed to estimate the demand levels about Reggio Calabria regional airport (South of Italy), in order to make available a tool able to analyse the impact of recent modifications in the supply (new links, new destinations and lower fares) that are expected to produce an increase in the air transport demand, also due to induced trips generated by the new supply. Differences between the univariate and multivariate models are highlighted. The models calibrations have been obtained by using the Box-Jenkins procedure.

1. Introduction

Usually, the estimation of the air transport demand can be obtained by different models/methods, among which time series models and market surveys (RP – SP methods) are the most used by airline companies. In the literature there are different works on the topic: [8] examines the relation between the demand and the distance on short range; [6], [7], [1], propose time dependent econometric models; [2] and [4] examine the impact of September 11 on air travel demand, with seasonal time series models. However, mainly simple autoregressive time series models have been used, even if with explanatory variables, while there are very few examples of ARIMA models and no one on the calibration of univariate and multivariate ARIMA models in the specific topic of the air transport demand simulation for a regional airport. The purpose of this paper is to calibrate and compare univariate and multivariate trend models to estimate the demand levels about Reggio Calabria regional airport (South of Italy). The application is interesting because due to the fare policy adopted by the main airline company operating at the airport and the variations in the number and schedule of the flights, the passenger demand at the airport changed in the last ten years against expectations (a promising positive trend has been...
followed by a very strong demand reduction). Recent modifications started by the local airport authority in the supply (new links, new destinations and lower fares) are expected to produce an increase in the air transport demand, also due to induced trips generated by the new supply. This study was partially developed within the research project on “Guidelines for the planning of the development of Italian regional airports”, financed by the Italian Ministry of Higher Education and Research.

2. Time series models

A time series is a stochastic process where the time index takes on a finite or countable infinite set of values \([3], [5]\). To describe such a stochastic process, its mean and its variance are used as well as two functions called AutoCorrelation Function (ACF), \(\rho_k\), \(k\) being the lag, and Partial AutoCorrelation Function (PACF), \(\pi_k\), \(k\) being the lag. The ARMA (AutoRegressive Moving Average) models are a class of stochastic processes for which the research of a statistical model has considerable probability of success; they are expressed as:

\[
j_t = \sum_{i=1}^{p} \phi_i X_{t-i} + \sum_{j=1}^{q} \theta_j a_{t-j}
\]

where \(\phi\) and \(\theta\) are the model parameters. To estimate these models, the series must be stationary and ACF and PACF must be time-independent. The non stationarity in variance can be removed if the series is transformed with the logarithmic function. The non stationarity in mean can be removed by using the operator \(\nabla = 1 - B\) applied \(d\) times in order to make the series stationary. In this way, if the B operator such as \(X_{t-1} = BX_t\) is introduced, the ARMA model becomes an ARIMA (AutoRegressive Integrated Moving Average) model:

\[
\nabla^d \phi(B) \cdot X_t = \theta(B) \cdot a_t
\]

The Box-Jenkins approach is the most known method to find, for a given set of data, an ARIMA model that effectively represents the data generating the process. The method requires three stage: identification, estimation and diagnostic checking. Preliminarily a data analysis should be carried out in order to verify the presence of outliers and examine the stationarity of the time series thanks to the use of the ACF and PACF correlograms. The identification stage provides an initial ARIMA model specified on the basis of the estimated ACF and PACF: if the autocorrelations decrease slowly or do not die out, then there is non stationarity and the series should be differenced until stationarity is obtained. After that, an ARIMA model can be identified for the differenced series. If there is no evidence for an MA or an AR then a mixture ARMA model may be adequate. After an initial model has been identified, the AR and MA parameters have to be estimated, generally by using least squares (LS) or maximum likelihood (ML) methods. Finally, the diagnostic check verifies if the residuals of the calibrated model can be considered belonging to a white noise series. To this aim, the significance of the residual autocorrelations is often checked by verifying if the obtained values are within two standard error bounds, \(\pm 2/\sqrt{N}\), where \(N\) is the used sample size. If the residual autocorrelations at the first \(N/4\) lags are close to the critical bounds, the reliability of the model should be verified. To check the overall acceptability of the residual autocorrelations
the portmanteau test statistic can be used. To check the normality of the residuals, the Jarque-Bera test (JB) can be used.

Starting from an univariate ARIMA model, some explanatory variables can be inserted. In this case, the dependent variable $x_t$ depends on lagged values of the independent variables. The length of the lag may sometimes be known a priori, but usually it is unknown and in some cases it is assumed to be infinite. Generally, if one dependent variable and one explanatory variable are considered, then the model has the form:

$$x_t = \alpha + \beta_0 y_{t-1} + \beta_1 y_{t-1} + \ldots + \beta_p y_{t-p} + e_t$$  \hspace{1cm} (2.3)

To reduce the number of parameters, it can be assumed that $\beta_i = 0$ for $i$ greater than some finite number $P$, that, when fixed, is called lag length. The obtained models are then called finite distributed lag models, because the lagged effect of a change in the independent variable is distributed into a finite number of time periods. In vector and matrix notation, eqn. (2.4) can be written as:

$$X = Y\beta + e$$  \hspace{1cm} (2.4)

As it can be seen in the above matrix, further $P$ observations are supposed to be available in addition to the values $y_1, \ldots, y_T$. If $e \sim (0, \sigma^2 I)$ and $y_t$ are fixed, then, based on the sample information $X$, the LS estimator is the best linear unbiased estimator of $\beta$. If the true lag length $P$ is unknown but an upper bound $M$ is known, then the LS estimator of $\beta^* = (\alpha, \beta_0, \beta_1, \ldots, \beta_M)^T$ is inefficient since it ignores the restrictions $\beta_{P+1} = \ldots = \beta_M = 0$. In order to compute $P$, this sequential hypotheses can be set up:

$$H_0^i: P = M - i, \Rightarrow \beta_{M-i+1} = 0 \quad \text{versus} \quad H_a^i: P = M - i + 1, \Rightarrow \beta_{M-i+1} \neq 0$$

The null hypotheses are tested sequentially beginning from $H_0^1$. The testing sequence ends when the first null hypothesis is rejected for the first time. The likelihood ratio statistic for testing the $i$-th null hypothesis can be written as:

$$\lambda_i = \frac{\text{SSE}_{M-i} - \text{SSE}_{M-i+1}}{\hat{\sigma}_{M-i+1}^2}$$  \hspace{1cm} (2.5)

where SSE is the sum of the squared errors for a model with lag length $P$. This statistic has an F-distribution with 1 and $(T-M+i-3)$ degrees of freedom if $H_0^1, H_0^2, \ldots, H_0^i$ are true.

When the lag has been computed, the independent variables can be inserted in the univariate model, in order to derive a so-called multivariate ARIMAX model:

$$\nabla^d \phi(B) \cdot X_t = \psi(B) \cdot a_t + \sum_{i=0}^{p_0} \beta_i (y_{t-i}^{(1)} + \ldots + \sum_{i=0}^{p_1} \beta_i (y_{t-i}^{(2)} + \ldots \ldots$$  \hspace{1cm} (2.6)

where: $y_{t-i}^{(j)}$ is the $j$-th independent variable at the time $(t-i)$ and $\beta_{i+1}^{(j)}$ is the corresponding parameter.

### 3. Application

Both univariate and multivariate models have been estimated by using the same data that refer to planned and enplaned passengers at the Reggio Calabria airport (sources: Official Statistic Institute – ISTAT; Ministry of Infrastructure and Transport; Association of Italian Airports - Assaeroporti). The data referring to 2004 should be considered as outlier.
because the airport was closed during the months of March, April and May 2004. The model has been built by means of the Box-Jenkins approach. Following the general procedure, first of all the ACF and PACF have been estimated. The analysis of the correlograms shows that the ACF decreases linearly and the value of the PACF at lag 1 is close to 1, i.e. there is a non stationarity in mean. Then, the time series has been differenced once. To eliminate the non stationarity in variance the time series has been transformed by using the logarithmic function. Finally, the analysis of the correlograms suggests an AR(1) component and a MA(2) component.

3.1. Univariate model

The presence of the outlier at the year 2004 has suggested to estimate two univariate ARIMA (1,1,2) models: in the first one the outlier has been removed (ARIMA\textsubscript{SO}); in the second one the outlier has been suitably modified (ARIMA\textsubscript{CO}) in order to better follow the natural trend of the series. The parameters estimates of the ARIMA\textsubscript{SO} model are reported in Table 1, and the estimated series w.r.t. the true one is depicted in fig. 1 (note that the figure reports forecast till 2006, used for the testing application described in section 3.3). The residuals series obtained has been used to carry out the diagnostic checking, particularly the Ljung-Box test can be considered satisfied for \( k = 12 \) and the Jarque-Bera test that provides a value of 0,31, i.e. the null hypothesis of residual normality can be accepted.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>AR1 ( \Phi )</td>
<td>0.909</td>
</tr>
<tr>
<td>MA1 ( \theta_1 )</td>
<td>0.858</td>
</tr>
<tr>
<td>MA2 ( \theta_2 )</td>
<td>-0.053</td>
</tr>
</tbody>
</table>

**Table 1. ARIMA\textsubscript{SO}: model parameters**

The ARIMA\textsubscript{CO} requires a suitable estimate of the outlier. This has been performed by means of a monthly AR(1) model (table 2), using the time series monthly data of the latest years. Data referred to 2004 have been forecast by the model and then the ARIMA\textsubscript{CO} model has been estimated (table 3, fig. 1).

![Figure 1. True and ARIMA estimated series](image_url)

Again, the Ljung-Box test can be considered satisfied for \( k = 12 \), and the Jarque-Bera test...
provides a value of 0,11, i.e. acceptance of the null hypothesis.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>AR1</td>
<td>φ 0,665</td>
</tr>
<tr>
<td>Constant</td>
<td>c 10220</td>
</tr>
</tbody>
</table>

Table 2. AR(1) model parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>AR1</td>
<td>φ -0,410</td>
</tr>
<tr>
<td>MA1</td>
<td>θ 1 -0,470</td>
</tr>
<tr>
<td>MA2</td>
<td>θ 2 -0,846</td>
</tr>
<tr>
<td>Constant</td>
<td>c 0,056</td>
</tr>
</tbody>
</table>

Table 3. ARIMA\(\phi\) model parameters

The comparison between the univariate models shows that both models fit well the true series and are statistically satisfactory. It is important to notice that the estimated value 2006 is the same for both models, so they validate mutually (fig. 1). Moreover, the official passenger data for the year 2005 for the airport of Reggio Calabria, not used to estimate the models and then considered a hold-out sample data, is 398089 (source: Assaeroporti, www.assaeroporti.it), which confirms the goodness of the estimated models.

3.2. Multivariate model

Starting from the univariate models, a multivariate model with two explanatory variables (pro capite income, It, and yearly number of movements from/to the Reggio Calabria airport, mt) has been considered. The sequential testing procedure previously described allows the P values for both variables to be identified; particularly, the demand in the year t depends on the movements in the same year t and on the income from year t to year t-6. The results of the model estimation are reported in Table 4 and fig. 2.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>AR1</td>
<td>φ 0,44</td>
<td>ln Iₜ₋₂</td>
<td>α₃ 0,21</td>
</tr>
<tr>
<td>MA1</td>
<td>θ₁ 1,08</td>
<td>ln Iₜ₋₃</td>
<td>α₄ -1,33</td>
</tr>
<tr>
<td>MA2</td>
<td>θ₂ -0,09</td>
<td>ln Iₜ₋₄</td>
<td>α₅ 1,22</td>
</tr>
<tr>
<td>ln mₜ</td>
<td>δ 1,22</td>
<td>ln Iₜ₋₅</td>
<td>α₆ 0,26</td>
</tr>
<tr>
<td>ln Iₜ</td>
<td>α₁ -0,004</td>
<td>ln Iₜ₋₆</td>
<td>α₇ 0,002</td>
</tr>
<tr>
<td>ln Iₜ₋₁</td>
<td>α₂ 3,45</td>
<td>Constant</td>
<td>κ -0,08</td>
</tr>
</tbody>
</table>

Table 4. Parameters of the ARIMAX model
3.3. Some application tests

Starting from the estimated models, particularly the multivariate models, some tests have been carried out in order to provide forecasting of the demand level both for the current year and the next year, and to verify the potential demand increase due to the new offered services with the capacity in terms of offered seats. At the end of 2005 and beginning of 2006, the local airport authority has promoted the introduction of new links with low-cost carriers from/to the airport of Reggio Calabria. Estimation of the demand level for the current year 2006 and for the next year 2007 by using the multivariate model requires the knowledge of the income and movements explanatory variables. The income for the years 2006 and 2007 has been obtained by means of an ARIMA model (ARIMA_{inc}). Specifically, the available data on income have been used to build an income time series univariate model, starting from the hypothesis that the boundary conditions are stable. As before in this paper, the model has been estimated by using the Box-Jenkins methodology. The analysis of the income series shows the presence of non stationarity, that have been removed by differencing twice. The Dickey-Fuller test applied to the differenced series confirms its stationarity. The identification of the ARIMA_{inc} model has been performed by analyzing the ACF and PACF correlograms, that suggest an AR(1) component and a MA(1) component. The results of the estimated ARIMA_{inc} are reported in Table 5.

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>AR1</td>
<td>$\Phi$</td>
<td>-0.979</td>
</tr>
<tr>
<td>MA1</td>
<td>$\theta_1$</td>
<td>-0.800</td>
</tr>
</tbody>
</table>

Table 5. Parameters of ARIMA_{inc}

With reference to the second explanatory variable, the number of movements for the year 2004 and 2005 has been obtained by official data (source: Assaeroporti, www.assaeroporti.it), while the number of movements for the current year 2006 has been computed by assuming the available data for the first forth months, and assuming that the
remaining months have the same number of movements of April. This can be considered realistic, because at the moment there is not any developing plan about new links (more destinations) or flights (increase in the frequency). The same value of movements has been considered for the year 2007, starting from the previous considerations. The ARIMA model applied to forecast the demand level for the current and the next years provides respectively 709468 passengers (year 2006) and 741248 passengers (year 2007). Following the estimated values, there is a growth of 78% w.r.t. 2005, whereas for the year 2007 there is an increase of 4,5% w.r.t. 2006. These results can be considered reliable, because the first months of the year 2006 have registered a considerable increase in the number of planed/enplaned passengers at the Reggio Calabria airport, due to the significant increase of the air transport offer, particularly if compared with the offered service in the previous ten years. After this encouraging answer of the demand to the new supply system for the current year, the successive more moderate increase of about 4% agrees with the forecast demand rate provided by Eurocontrol (www.eurocontrol.int) and IATA (www.iata.org) for the European market. Finally, in order to verify the ratio between the forecast demand level and the capacity in terms of offered seats, an estimate of the seat number for the year 2006 has been done; such value has been obtained by assuming the same hypotheses as before, i.e. the number of flights on April 2006 remain the same until December 2006. Then, 1164522 seats have been estimated; given that the estimated demand level for the same year is 709468, the average load factor is 0,61. Then, the actual supply can be considered sufficient to satisfy the forecast demand. Furthermore, the number of movements for the current year is beneath the runway capacity and the landside capacity, thus suggesting a potential for growing both in terms of supply and demand.

4. Conclusions

The comparison between the univariate and the ARIMAX models (fig. 3) shows that both the models provide satisfactory results, even if the univariate models fit better than the ARIMAX model when there are some peaks. However, it is not possible to assert in absolute that the estimated univariate models are better than the multivariate models and vice versa. As obtained in this study, the better forecasting power of univariate models is offset by its limits of validity, which depends on the stability of the boundary conditions. On the other hand, even if the multivariate model solves this problem with explanatory variables, however their time series are often difficult to find. That is the reason why no more independent variables (as for example the fare) have been introduced, given that the data are not always available for the examined period. In any case, the estimated model shows a reasonable explanation power and it can be used to test policies about the development of the Reggio Calabria airport. Particularly, some tests have been carried out following the current development plan of the airport authorities, in order to verify the limits to the airport growth. On the contrary, the univariate models can only forecast the demand level all the underlying conditions being the same and then they cannot be used to simulate the effects of different policies. Further developments concern the use of more explanatory variables, such as the number of served destinations and the fare (this latter should be estimated by using suitable approaches), as well as the implementation of a specific procedure to test and evaluate the effects of different developing policies,
particularly in terms of both landside capacity (due to the forecasted demand increase) and airside capacity (due to infrastructure characteristics and ATC systems that set a limit to the maximum number of movements at the airports).

![Graph](image.jpg)

**Figure 3.** Trends of the original and estimated series.

### References


CROSSING TIME MICRO-ANALYSIS:
DIFFERENT CROSSING SCENARIOS FOR DIFFERENT ZONES
OF APPROACH TO THE CROSSROADS

Edoardo DE LIDDO¹

Abstract. In the present paper a method to apply micro-analysis criteria to the
calculation of manoeuvre times and spaces is described. The method is valid for
all the kinds of manoeuvres at junction. In the following the case of crossing
manoeuvre will be faced.

1. Introduction

Traditional manoeuvre analysis is leaded trying to mediate between results accuracy and
calculation simplicity and rapidity. Complex simulation and calculation are needed to find
times and spaces values near to the real ones; but at the same time, simplification
hypothesis, that makes calculation easier and quicker, yields to less realistic simulations.
Since in many practical cases results accuracy is considered less important compared to
other aspects (i.e. time and cost saving), most widespread techniques accept rough
assumptions for manoeuvre simulation and analysis; nevertheless there are other cases (i.e.
safety checking at free junctions, traffic micro-simulation algorithms, calculation of time
delay due to junctions, etc.) in which precise calculation procedures and results the most
close to real values are necessary. In order to reach this aim, we have to reduce
simplifications and turn to manoeuvre micro-analysis procedures, that take into account all
the phases that compose the manoeuvre that are usually neglected.

This paper describes a method to apply micro-analysis criteria to the calculation of
manoeuvre times and spaces. The method is valid for all the kinds of manoeuvres at
junction; in the following the case of crossing manoeuvre is considered.

¹ DVT-Technical University of Bari, Italy. e-mail: e.deliddo@poliba.it
2. The proposed method

The crossing manoeuvre micro-analysis consists of mathematical models that are the easy ones of cinematic; indeed calculation complexity is not due to mathematical formulations of models but to the multiple possibility of succession of the manoeuvre phases; the occurring succession of manoeuvre phases depend on several variables. Therefore the problem is to define and to analyse all the possible scenarios that, at normal condition (excluding car accident hypothesis), can occur at junction. Each scenario changes at changing of the number and type of phases that compose the manoeuvre and it depends on several variables.

All these variables (geometry and cinematic characteristics of vehicles, junction and meeting roads geometry, conditions of visibility, etc.) have an influence on the manoeuvre execution modalities; consequently changing the variable values, the number and type of manoeuvre phases change and than the possible crossing scenarios change.

A preliminary analysis of the vehicle phases of approach to the junction area shows as the real crossing executions depend on the position of the point in which the crossing manoeuvre begins (decision point). At changing the decision point position, the crossing scenario changes. In details, seven zones of the section of approach to the junction (we’ll call its “zones of approach”, see fig. 1,2) are singled out.
The main characteristic of a zone is that all the manoeuvres for which the crossing begins in the zone, have the same crossing phases succession (that is the same crossing scenario). Points A, B, C, D, E, F, G, and H in the figure 1 and 2 delimit the seven zones at stake.

Therefore, in the following, possible crossing scenarios will be analysed and classified by “zones of approach”.

Mathematical expressions of the coordinates of these points shows that points C, D and E, coordinates depend, between the others variables, on junction speed \( V_i \), that is the speed at which the main part of vehicles cross the junction; on the contrary point F and G coordinates depend on slow advancing speed \( V_{al} \), that is the speed at which vehicles that arrive to the junction area limit lacking the conditions of minimum visibility move cautiously forward through the junction area.

The basic assumptions for the manoeuvre analysis are the followings:

1. When the driver of the vehicle approaching the junction arrives at distance “\( S_{st} \)” (stopping space) from the junction area, if he has not the minimum visibility conditions he begins a slowing down manoeuvre with a constant deceleration “\( d \)”, that allows him to stop at the junction area limit.
2. If during the approaching phase with constant deceleration “\( d \)” the driver reaches a \( V_{al} \) speed before the decision point, he continues approaching with constant speed \( V_{al} \).
3. When the driver arrives at the point in which visibility is enough for him to decide (decision point), he verifies that the crossing manoeuvre is possible and he begin the manoeuvre (that is to say to neglect that situations in which the vehicle need to stop and to wait that good crossing conditions arise).
4. Within the junction area the speed \( V_i \) is never exceeded.
5. Acceleration phases are all executed with constant acceleration “\( a \)”.
6. The manoeuvre finishes when the vehicle reach again the speed \( V_1 \) that is the normal speed kept far from junctions.

In Fig. 3 the crossing scenarios corresponding to the seven zones of approach are represented, showing characteristics space/speed. In the figure is put in evidence that the state of characteristics change with: i. the \( x \) position of the decision point, ii. the vehicle speed \( V_x \) at decision point, iii. the vehicle speed \( V_{P_x} \) at the end of the action/reaction space after the decision point position.

You can observe that zones 1 and 2 characteristics are here described to show an exhaustive analysis example, but its have little practice relevance because in main urban and suburban environments condition of visibility in the areas of approach at junctions are never the good one to decide the possibility to cross.
Figure 3: Crossing scenarios corresponding to the seven zones of approach
LEGENDA

PRECEDENTI DI CALCOLI UTILIZZATI DELL’AREA AMBIENTALE DELLE AERODELTA MIGLIA REIPERRE (PMR)

<p>| | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Ascissa del punto iniziale della manovra</td>
<td>Ascissa del punto di separazione fra le zone 1 e 2</td>
<td>Ascissa del punto di separazione fra le zone 2 e 3</td>
<td>Ascissa del punto di separazione fra le zone 3 e 4</td>
<td>Ascissa del punto di separazione fra le zone 4 e 5</td>
<td>Ascissa del punto di separazione fra le zone 5 e 6</td>
</tr>
<tr>
<td>Vx</td>
<td>Velocità del veicolo nel generico punto di ascissa “x”</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>tral</td>
<td>Tempo impiegato dal veicolo per raggiungere il punto di ascissa “x”</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VPx</td>
<td>Velocità del veicolo dopo il tempo psico-tecnico “tpt”</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stpt</td>
<td>Spazio percorso dal veicolo nel tempo psico-tecnico “tpt”</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sres</td>
<td>Spazio rimanente perché il veicolo attraversi e liberi l’area d’incrocio</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sacc</td>
<td>Spazio necessario perché il veicolo raggiunga la velocità d’incrocio con accelerazione</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vu</td>
<td>Velocità del veicolo nel punto di uscita dall’area d’incrocio</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>tacc</td>
<td>Tempo impiegato dal veicolo per percorrere il tronco di “Sres” in accelerazione</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>tcost1</td>
<td>Tempo impiegato dal veicolo per percorrere a velocità costante lo spazio precedente l’area d’incrocio</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>tcost2</td>
<td>Tempo impiegato dal veicolo per percorrere a velocità costante lo spazio entro l’area d’incrocio</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>t1</td>
<td>Tempo di attraversamento a partire dalla posizione “x”</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>tm</td>
<td>Tempo totale di manovra</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sf</td>
<td>Spazio finale percorso in accelerazione nel tempo “tf”</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sm</td>
<td>Spazio totale di manovra</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pm</td>
<td>Perditempo di manovra</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Tutte le Zone

Figure 4: Algorithm for crossing maneuver parameter evaluation

According to this aspect characteristics of the zones 5, 6 and 7, in which with a bigger probability decision points fall.

In the following of the paper an algorithm is suggested (see Fig. 4) by which it has been possible to automate the calculation of manoeuvre time and space. This automation has allowed to repeat calculations for different operative conditions, that is to say for different...
values of parameters at stake. Particularly interesting is the results of variation of the real crossing time “t₁” (that is the time the vehicle spends to cover the distance between the decision point and the exit point of the junction area) with the position of the decision point; considering the influence that this time has on the critical interval and on the gap acceptance for the crossing manoeuvre.

Using algorithm in Fig. 4 has been possible verify that “t₁” is not always decreasing when “x” decrease. In Figg. 5, 6, 7, 8 e 9 shapes of this variation are represented changing several parameters as “V₁”, “Vᵢ”, “Val”, “a” e “d”.

Those parameters influence “t₁(x)” in a different and not uniform way:
- “V₁” influence faraway zones from crossing area; that is when “x” is greater then 20-25 m;
- “Vᵢ” influence complete branch of approach to the junction when his value is smaller then 25 Km/h; whereas, when his value is greater then this, his influence is appreciable only for distances greater then 10,00 m from crossing area;
• "Val" influence only on nearest or internal zones to crossing area; no one influence we see far away more then 10,00 meters from crossing area;

• "a" shows his influence all in phases of manoeuvre in which it take his effect; this fact determine a strong influence of this parameter inside crossing area and in nearest zones as far as a distance of 50,00 meters from crossing area;

• "d" finally shows his influence all in phases of manoeuvre in which it take his effect; this fact extend his influence all in branch of approach.

By those diagrams it’s possible to observe that "t1" has an upwards concavity inside Zone 5. Particularly (v. Fig. 8), this concavity is as much large as "a" is smaller. On the contrary, it is as much great as "d" is greater (v. Fig. 9).
If we consider that in the overwhelming majority of urban intersections decision point is very near crossing area (or even inside them), we understand that parameters which have much more influence on $t_1(x)$ and, consequently, on exercise and capacity of uncontrolled junctions, are

- in major way, “a” (v. Fig. 8) and “$V_i$”, for smaller values then 25 Km/h;
- in lower way, “$V_{dl}$”.

In any case, the shape of the function $t_1(x)$ reveals that the crossing time has a minimum value, that follows, almost in all the cases, in a zone of the section of approach included among about 10 and 20 meters from the junction area; we can argue that when visibility condition are such to bring the decision point this zone, crossing times are minimums and, consequently, minimums are the idle times of the crossing manoeuvre.

3. **Conclusion**

Consideration and analysis carried out in the present paper show how the micro-analytical approach to manoeuvres analysis offers two main advantages:

- characterize suitable physical and geometric conditions for a better use of free at-level junctions
- open new and several perspective to understand drivers behaviours when they meet particular conditions that occur during the manoeuvres executions.

In this second perspective further studies will be devoted to other recurring manoeuvre analysis, both close to and distant from at-level junctions, considering also parameters that have an influence on visibility conditions, especially taking into account the weight that these conditions have on drivers behaviours.

With regards to this last aspect, actual studies are going to be advanced on: i. role of decision point as a point of crucial interest for manoeuvre scenarios analysis, ii. calculation of decision point position with regard to both the time "$t_1$" and the visibility conditions on the branch of approach to the junction.

A first comparison between theoretical results and empirical results of crossing times and gap acceptance critical intervals (obtained in other authors studies) shows similar results. Nevertheless, it is important for further research in this field to carry out experimental checks of this approach to manoeuvres analysis, in particular with regard to other Italian and European urban realities.

This is an other of objectives of research in progress.

**References**

*T. M. Matson, W. S. Smith, F. W. Hurd: TECNICA DEL TRAFFICO STRADALE (1961).*

*Her Majesty’s Stationery Office - London: RESEARCH ON ROAD TRAFFIC (1965).*

*P. Ferrari, F. Giannini: GEOMETRIA E PROGETTO DI STRADE (1977).*

SIMULATION AND VALIDATION OF PARKING KINEMATICS CYCLES

Borja BELTRAN, Stefano CARRESE¹, Emanuele NEGRENTI²

Abstract: In the frame of the FP5 HEARTS project a set of models has been developed in order to simulate the kinematics cycles related to parking processes. This work is focused to deepen the methodologies used for studying parking management procedures and improve the assessment of impacts from transport systems. In the proposed methodology a specific effort has been dedicated for modeling parking which generates a higher amount of hot emissions and the re-entering traffic phase which is relevant for cold start and evaporative emissions. Parking phase is represented by dedicated ‘searching speed’ and ‘searching time’ models. A multi-disciplinary approach from fuzzy logic to deterministic models has been proposed to achieve the objectives. Eventually model results have been validated with data obtained from a field survey where data related to parking has been collected.

1. Introduction

VPQ HEARTS project (Health Effects and Risks of Transport Systems) address the need for more integrated methods for health risk assessment which consider the full range of exposures and health effects and can be applied in the policy or planning process. The new version of TEE model (Transport Energy Environment) has been developed in the frame of the FP5 ISHTAR and HEARTS Projects with several features for better analysing transport related direct impacts, taking into account vehicles kinematics, cold start emissions distribution, parking processes, noise emissions and accident occurrence. Parking and re-entering traffic are a source of congestion and pollution: half of the cars driving downtown during peak hours in heavily congested areas cruise for parking [1]. Parking, on street and off street, can create pollution problems regarding additional emissions because of vehicle movements at low speed and low gear, number of decelerations and accelerations, long searching time, low temperature of engine during cold and warm start operations, presence of ramps on parking lots and so on [9]. In facts, the

¹ Università degli Studi di Roma TRE, beltran@uniroma3.it, carrese@uniroma3.it
² ENEA, ENE-TEC negrenti@casaccia.enea.it
relevance of parking phase in the impact assessment is out of discussion, but there are not enough studies or experimental data in order to predict the main parking characteristics related to emission calculation such as searching time, searching speed or time of warm up. The behaviour of the parking and inserting vehicle flows (going to and leaving from parking areas) can be described by means of simplified speed cycles, based on the kinematics of the vehicle during the parking and the re-insertion. Parameters defining the shape of such cycles are searching speed, searching time and time of warm up, on which our research is focused. Determination of emissions due to parking is possible after integration in a model for emission calculation as TEE, which takes into account the additional emissions at lower speed and under cold start conditions. TEE is an innovative model of traffic emissions aimed at providing detailed calculation of the several factors influencing vehicles fuel consumption and pollutant emissions [13]. The TEE model link representation is characterized by three ‘flow modes’ corresponding to transit, parking and inserting vehicles. Inserting and parking modes, defined by characterized kinematics cycles, are the focus of this research.

2. Methodologies and simulation results

2.1. Simulation of parking

Parking search phase is represented by dedicated ‘searching speed’ and ‘searching time’ models. Such models are further split in two sub-models depending on the category of parking: on-street and off-street. In the on-street parking, searching speed is modeled by a fuzzy model as a function of the linear traffic density and the occupation rate of the parking facility. Searching time for on street parking uses a probabilistic approach to calculate the searching time based on the occupation rate of the link connecting parking areas and searching speed.

In the off street parking, a detailed simulation of choice of parking place is proposed in order to reproduce the vehicle movements from entrance to parking place which allows to assess the searching speed and searching time in order to reproduce the kinematics cycles of vehicles in a parking facility.

The output of models allows TEE to build the simplified speed cycle (speed versus time), calculate the additional emissions and support decisions about parking management. TEE2005 software with detailed parking models estimates fuel consumption and pollutants emissions with high spatial and temporal resolution, and it can be used for the design and for the assessment of parking management systems including a variety of specific telematic technologies such as VMS for parking guidance or routing.

2.1.1. Searching speed model for parking on street

Searching speed model for on street parking allows to evaluate the speed adopted by the user in the parking and it is related to meaningful variables of link and parking conditions provided as input by a traffic model (occupation rate and linear traffic density). The model
is based on fuzzy logic. In fuzzy logic, linguistic values and expressions are used to describe physical variables [17]. Each linguistic term is described by a fuzzy set. It is thus unequivocally defined mathematically by the two statement basic set \( G \) and membership function \( \mu \). The membership functions states the belonging of every element of the universe of discourse \( G \) in the set in the form of a numerical value between zero and one. If the membership for a specific value is one, then the linguistic statement corresponding to the linguistic term applied in all terms. If, in contrast, it is zero, then there is absolutely no agreement.

A fuzzy set is defined as the set of ordered pairs \((x, \mu(x))\) where \( x \) is an element of the universe of real input and \( \mu(x) \) is the membership function, that attributes to each input \( x \) a real number \( U \), describing the degree which the variable \( x \) belongs to the set. At this moment the rules describing the driver’s behaviour in order to associate input and output are defined by a matrix of cases. This choice is a limit in the simulation that can be overcome using collected data from our field survey.

An example of searching speed for several traffic and parking conditions in a test network is shown in the figure 1.

![Figure 1: Relationship between searching speed as function of flow speed for several occupation rate (OCRate)](image)

2.1.2. Searching time for parking on street

Searching time model estimates the time employed by a driver from the moment he decides to park up to he finds the first available space lot to park. A great number of parking models are present in literature and many approaches have been used to understand and replicate parking choice behaviour. These models are used to investigate parking policies, while the approach of this research is addressed to the study of the searching time in a known link, therefore it can not be represented with network models [6], [8], [2], [3].
Searching time is a function that increases with the number of parked cars and consequently of occupation rate. The function will stretch asymptotically towards infinite when the number of occupied parking places is close to the capacity of the road. Obviously an infinite searching time is not compatible with a real model and then accordingly to the observed data a threshold of the maximum occupied parking rate has been proposed in order to determine the maximum searching time in a link.

Given an uniform distribution of the parking along a link, the user, once he reaches this arc, will proceed at searching speed and examine sequentially all the places up to he finds a free slot that satisfies all the conditions. The model of the searching time is then based on the analysis of probability number of attempts that a driver must carry out before finding a useful place to park.

\[
T_{\text{searching}} = \mu \frac{L_p}{V_{\text{searching}}} = \frac{L_p}{V_{\text{searching}}} \cdot \frac{N + 1}{\text{occ}_2 - \text{occ}_1} \cdot \log \left( \frac{N + 1 - \text{occ}_2}{N + 1 - \text{occ}_1} \right)
\]  

(1)

where \( \mu \) is the mean value of attempts, \( N \) is the total parking places, \( L_p \) is the length of considered parking place \( \text{Occ}_1 \) is the number of occupied places in the beginning of the study interval and \( \text{Occ}_2 \) at the end of interval. The previous expression calculates the mean searching time in the slice of time (usually one hour) because original expression has been integrated among occupation in the beginning and the end of time interval. So, consequently it can also consider the trend of the parking in the hour [3].

Figure 2: Searching time and occupation rate for several densities
2.1.3. Models for parking off street

The methodology is extended in order to consider parking off street and a detailed simulation of choice of parking place is proposed to reproduce the vehicle movements from entrance to the parking lot. In off street models a great number of information about parking geometry and possible driver’s destinations influencing the choice of parking must be known in order to simulate car movements with a major detail.

Off street searching speed depends on occupancy rate and geometry characteristics describing parking lanes. Young [16] has shown that driver in parking establishment have a desired speed that is influenced by the characteristics of road, particularly its length. For this purpose data supplied by Young has been fundamental in the modelling of vehicle flow speed in parking facilities. Relationship based on length of road and occupancy rate has been found and used to calculate the searching speed. The link are further classified in two typologies influencing speed, one way link and two way link.

Searching time in a parking off street has been studied analyzing the parking lot choice. The choice of parking place consists in two level decisions: the choice of isle inside the parking and the choice of the space inside the isle. The choice of isle is modeled with a Logit model proposed and calibrated by Van der Waerden [15] and it supplies the distribution of parking demand among the isles.

\[ U_i = \beta_c * N_i + \beta_d * d_i + \beta_{TM} * TM_i + \beta_E * E_i + \beta_{P&R} * P & R_i \]  \hspace{1cm} (2)

\[ P_{ik}(i/K) = \frac{\exp(U_i/\theta)}{\sum_{i \in K} \exp(U_i/\theta)} \]  \hspace{1cm} (3)

Where:
- \( U_i \) = utility choosing aisle I
- \( N_i \) = capacity of aisle i
- \( d_i \) = distance from aisle i to desired destination (shopping,..)
- \( TM_i \) = distance from aisle i to the ticket machine
- \( E_i \) = distance from aisle i to pedestrian exit
- \( P&R_i \) = distance from aisle i to Park & Ride system (rail, metro,..)

The first decision level requires also the knowledge of the link-path incidence matrix. Once the user has chosen the isle maximizing its utility, the time employed inside the isle is calculated with the same probabilistic model used in the on-street case. So, the searching time is the sum of time required to reach the isle and the time employed searching a free space inside of it.

2.2. Simulation of inserting phase

From statistical data relating to the Drive modem and Hysem driving patterns [11], 41.9% of the trips are started with a fully warmed-up engine (water temperature exceeding 70°C) for passenger cars, while about 19% only of the trips are actually started with a cold engine,
i.e. with an engine temperature equaling the ambient temperature. Similar results (20%) were recorded during a measurement campaign conducted in Austria at a small scale. About two fifths of the trips are started with an intermediated engine temperature lying between ambient temperature and fully-warmed up engine temperature (70°C) [11].

The effective conditions of start up, defining cold, warm or hot start, are fundamental for an accurate assessment of emissions. In cold start and transitory warm up the engine has not reached its operating temperature and enrichment of intake mixture occurs. If we assume that short journeys, mostly consisting of home to work, home to shopping or home to school, have large fuel consumptions and emission values due to cold start, it is necessary to propose a methodology in order to determine the real operating conditions of the engine and the temporal length of the cold start phase.

The catalyst converter is also a critical factor in the overall production of hydrocarbon emissions from a vehicle during cold start period. The catalytic material is generally non active at low temperatures and it is therefore necessary for the catalyst to achieve the light-off temperature in order to convert efficiently pollutant emissions [12].

As regarding the re-insertion phase the problem is to simulate the real operational conditions of the engine (engine temperature) and catalyst system (catalyst temperature and pollutant conversion efficiency) during the warm up phase. For this purpose a suite of models have been proposed to describe the dynamic behaviour of the system during the transient period of warm up in different operating modes. The requested output of different models are:

1. Temperature of engine and the engine exhaust second by second (temperature profile) derived from speed cycle of the vehicle.
2. Catalyst temperature profile and conversion efficiency of pollutants
3. Vehicle cooling profile for the assessment of the real temperature of start up.
4. Determination of number of vehicles that are in cold conditions or cold start fraction in the link.

The first model simulates engine heat transfers into the combustion chamber and into the catalyst converter second by second, in respect of energy balance. The input of the model is the inserting speed cycle and derived fuel consumption and effective power used by the vehicle to accelerate. The meaningful output is the vehicle cooling profile where time of warm up can be easily deduced from. Only a small rate of thermal energy set free during combustion is converted in useful mechanical energy for motion. The second step is to calculate the whole combustion heat release given by fuel combustion and assess the rate of conversion of this energy to useful energy on the wheels. The whole of heat losses can be viewed as an energy storage device composed of coolant, engine oil, engine block and engine head. The energy level can be expressed as the temperature of the system. There is a source of energy in an internal combustion engine and several sources of energy loss. Heat losses model consists in two models: one is the engine coolant system and the other is the radiator coolant temperature model [10], [5], [7].

The catalytic converter warm up numerical models is based on heat transfer analysis through the exhaust system. The model accounts for the engine out mass flow rate, temperature of exhaust gas and concentrations of pollutants. The heat losses in the catalyst system are account according to convective heat transfer and convective heat transfer from the wall to ambient air. The catalytic conversion efficiency is estimated according to a set of fit curves in function of Equivalence Ratio Ø_exh and temperature of catalyst T_cat and
described as a S shape Wiebe function with a minimal amount of data collection necessary [5].

In order to determine the real temperature of start up a model has been proposed that predicts the average temperature of coolant after the engine is switch off during a specific period of time, usually the parking duration. The temperature of engine after a specific time is calculated as function of temperature of engine at the moment of arrival, ambient temperature and the parameter k independent of ambient temperature and only dependent on characteristics of the vehicle. Heat losses to ambient air are mainly due to natural convection and in this case we must use the law of Dulong and Petit. $T_{\text{eng}}$ is the start temperature of engine and $T_{\text{amb}}$ the ambient temperature [4].

$$
(T_{\text{eng}} - T_{\text{AMB}}) = \frac{256}{\left(\frac{4}{(T_o - T_{\text{AMB}})^{1/4}} + kt\right)^4}
$$

(4)

Afterwards the next step is to calculate the cold start fraction, i.e. the effective rate of inserting vehicles that cruises in cold start conditions. This ratio, CFEMICLD, will be zero when all of vehicles are in hot conditions, otherwise (all vehicles in cold conditions) CFEMICLD value will be one. This ratio is calculated using the vehicle cooling profile, calculated by means of expression (4), in order to know the start-up temperature of coolant $T_{\text{eng}}$.

$$
\text{CFEMICLD} = f(T_{\text{hot}} - T_o) = f(\Delta T) = \frac{T_{\text{hot}} - T_{\text{eng}}}{T_{\text{hot}} - T_o}
$$

(5)

Figure 3: CFEMICLD for ambient temperature of 25°C
3. Validation of models with survey data

The last step of the work concerns the validation of the models with the data collected during the field survey. The experimental campaign has been conducted in conjunction with the IM-CNR in Naples and it has been conceived to simulate parking in a set of streets monitored by traffic cameras at the intersections. Monitoring cameras in the street provide traffic data as flows and speeds that are necessary to correlate the results given by the models.

The car used for the measures is a Fiat Punto that it has been equipped with an on-board system that allows the data acquisition of engine, vehicle and geographic position by means of a GPS receiver. The system supplies monitoring position, instant speed, gear, engine and catalyst temperature second by second by means of a data acquisition program developed with Labview. The engine variables have been used to validate the models regarding the time of warm up and the vehicle and kinematics variables have been used for the validation of searching models, i.e. searching speed and time.

The data were collected during the last week of January 2006 from Monday to Friday in the peak hour morning (9-11 AM) and in the afternoon (14-16 PM).

The first validation regards the vehicle cooling profile for validating the time of warm up. Two kinds of validation have been achieved: time of warm up in cold start conditions in the morning when temperature of engine was equal to temperature of ambient, and time of warm up in warm conditions in the afternoon after two or three hours of pause. In this way it has been also possible to validate the model concerning the temperature of start up after vehicle is parked for a period of time (i.e. parking duration).

The first results from the field survey show an accurate calibration of simulated vehicle heating profile (figure 4). The simulated and observed measures are linearly correlated with a $R^2$ over 0.90 for any temperature of start up during the phase of warm up. The vehicle cooling profile model offers a good approximation of temperature of start up with an error non exceeding 20% in all considered cases.

![Vehicle cooling profile with temperature of start up equal to 13 °C](image)

*Figure 4: Vehicle cooling profile with temperature of start up equal to 13 °C*
The validation of searching models requires the use of available traffic data in the monitored streets. The first results of the validation show that models fit quite properly the phenomena.

Figure 5: Correlation of measured and modeled searching speed

The searching speed model is correlated with a $R^2$ of 0.405 (figure 5) considering all the sample. The searching speed depends on linear density and the occupation rate. After a detailed analyze of the sample, the measures with a major error or worse explained by the model are the speed related to small densities in the sample (around 60 veh/km). In fact, if only the densities greater than 80 veh/km are considered (85% of the sample) the $R^2$ of the correlation is 0.698, reaching a $R^2$ equal to 0.73 densities greater than 90 veh/km.
This result brings out an important contribution to the methodology. The model fits accurately the searching speeds related to high densities. However the small densities are not well explained by the model. The only reasonable explanation is the calibration of the current fuzzy model that penalizes the low densities. The experimental data will be used to recalibrate the model in order to define the new membership functions.

The modeled searching time is correlated with experimental data with a $R^2$ of 0.52 and results show the difficulty of simulating the phenomena with a discrete probabilistic approach. The model tends to overestimate the searching time, and only in a small number of cases the time is hugely underestimated. The correlation of searching time is shown in figure 7.

4. Conclusions

The methodology is a first step in order to characterise the parking procedures and its environmental and transportation effects. The validation of models with survey data offers an opportunity to evaluate the robustness of the methodology and brings out new prospective to deep and improve the models. The first validation has provided several correction factors that improve the accuracy of the simulation and they have contributed to calibrate the models and demonstrate that methodology is appropriate to study the problem. The results encourage the next step focused on improvement the detail of the approach using new experimental data.
Next step will regard the calibration of parking fuzzy model with a Neural Network approach in order to avoid using a fixed matrix of rules, and validation of parking off streets with a new field survey.

5. References


COLLABORATIVE DEMAND AND SUPPLY NETWORKS: MODELING, DESIGNING, MANAGEMENT AND PERFORMANCES EVALUATION

Maurizio BIELLI and Mariagrazia MECOLI1

Abstract. The globalization of business and production around the world is going to generate a revolution in the supply chains and logistics systems design and management. The real innovation is evolving along two main lines: production of final goods is going to be performed by groups of firms integrated into “network of enterprises”, while logistics and distribution by large-scale world-wide service enterprises. Therefore, the topic of collaborative demand and supply networks and the relative methods and tools for modeling, designing and evaluation is becoming very relevant for efficiency and competitiveness of enterprises clusters.

1. Introduction

The globalization of business and production around the world is going to generate a revolution in the supply chains and logistics systems design and management. The real innovation is evolving along two main lines: on one hand, logistics and production are becoming tasks complementary but performed by different and separated companies; on the other, while production of final goods is going to be performed by groups of firms integrated into “network of enterprises” [22] - [23], both internal logistics and distribution is going to be performed by large-scale world-wide service enterprises [6]. Industrial logistic systems, either internal or dedicated to large-range distribution, are now objects of a significant technological innovation owing to their crucial role. That is their increasing utilization in supplying producers and markets in a wider and wider space and in due time, even in front of relevant variations in the demands. Present evolution of supply chains and

1 Istituto di Analisi dei Sistemi ed Informatica “Antonio Ruberti”, Consiglio Nazionale delle Ricerche, Viale Manzoni 30, 00185 Roma, Italy – e-mail : bielli@iasi.cnr.it
logistic networks is pushing managers to investigate the opportunity of increasing their system effectiveness and competitiveness with the utilization of models and design tools so far developed by the scientific community.

In this paper particular reference is made to collaborative demand and supply networks where main problems concern the coordination of distributed and autonomous agents and the application of suitable methodologies and technologies for the innovation in enterprises networks.

2. CO-DESNET

The European Co-ordination Action CO-DESNET has been designed to promote the diffusion of the European scientific knowledge on the problem of designing and managing large-scale multi-functional multi-agents Collaborative Demand & Supply Networks. Its focus is concentrated on large-scale networks of production, logistics and service enterprises operating within a common industrial sector. The main activities carried out in this project consist of setting up a common information system for design and management issues, tools, procedures, performances evaluation, and best practices in supply chain and logistics fields. In particular, a Virtual Library was organised in order to collect scientific papers and a set of design and management procedures for networking services devoted to SMEs. Then, a Virtual Laboratory provides a benchmark of case studies on existing clusters of enterprises in Europe, as well as sector studies, collected and standardized, by presenting an aggregate analysis of their financial, economic, and operational structures, the organisational issues, the interactions with the socio-economic environment and the expected development. One of the key contributions of this action is the identification of a number of performance indicators specific to this context.

More details can be found in the project web portal [9], acting like a “dedicated information system”, where many contributions have been reported according to a standard format. Moreover, information and news are included about funding programs supporting the technological innovation and the re-engineering of enterprises networks, and some ‘gates’ connecting to other European projects or institutes.

3. Modeling

Traditionally the design and management of efficient supply chains (SC) and logistics networks required the optimization of different functions and subsequent coordination of several separate solutions, then many mathematical optimization techniques have been proposed, and successfully used.

Such methods are based on the use of a mathematical model of the Supply Chain, where flows and nodes are present, a number of constraints in the form of mathematical inequalities represent the functioning of the system and one or more objective functions must be minimized respecting the constraints. Remarkable results were obtained in this
area, and an accurate overview of many different approaches to this problem is presented in [18].

In most optimization models the objective is to optimize and control the flow of materials and information among all the components of the Supply Chain system, referring to the three stages [19]: procurement, production and distribution. Each of the three stages has different characteristics, dimensions, and complexity in real problems; it is thus necessary to use different modeling solutions, and related algorithmic machinery to identify solutions to these models. Unfortunately, the nature of most real Supply Chain problems is complex and the mathematical models that must be adopted are either complex in structure or very large in size, or both. A factor that has a great influence on the complexity of the Supply Chain problem is the time horizon adopted for planning: from strategic planning (typically, network design) to tactical and operational planning we experience a dramatic increase in problem sizes that may often require sophisticated algorithmic techniques customized ad-hoc.

Trends in modelling supply chain and logistics networks have been widely outlined [4, 15], and decision support models for global supply chain design have been widely classified and discussed [14].

4. Designing

Supply chains and logistics networks represent the complex of all activities devoted to modify time and space attributes of any resource is managed within a production and distribution system. Then, their design has to consider a multi-actor situation and a multi-attribute attitude of the decision-making process. Traditional design approaches are based on the procedural paradigm that performs a deterministic sequence of design steps, in order to define system parameter values. Moreover, the design involves a progressive selection of attributes and a tradeoff analysis of costs and performances. Therefore, they are not able to tackle systems with functional and conceptual complexities such as the case of logistic systems, because, in order to improve enterprises’ business it is necessary to set up a structured and integrated design approach.

A framework of design models and methods has been proposed in [2,5,7], and in particular, a method based on the workbench and blackboard approaches is outlined, suitable to providing tools for supporting the decisions of logistic managers. Then, an application to the design of a large-scale connective network of enterprises and logistic services is also investigated.

The design methodology so far developed focused on logistic manager point of view: nevertheless, quality, cost service, reliability, safety and other important parameters must be taken into account. In fact time and cost performances for handling materials and finished goods distribution are strongly affected by performances of logistic service providers. For this reason we believe that a breakthrough advantage may come from an extension of key concepts of the logistic workbench to all the elements of the logistic chain. The enhancement of the entire logistic system involves effective and efficient decision processes and the quality of available information.
These requirements should be pursued by integrating the elements (forwarders, carriers, wholesalers, distributors, etc.) by means of suitable support to electronic data exchange. Moreover, decision support modules have to be specifically tailored for each step of the logistics chain.

Other interesting research topics to be addressed concern the use of new computation techniques (heuristics and AI) to meet the flexibility and quality requirements in logistic design [13].

5. Distributed decision-making

As much as the dimension of a supply chain and a logistic network increases, as much as the interest toward a distributed management organization also increases. A distributed system is represented by a decentralised collection of autonomous organizations that cooperate with each other to achieve a common global goal. Therefore, the new problem to tackle is the trade-off between decision decentralization and global efficiency in large-scale networks. In this framework, models of decentralized management organization and local autonomous agents coordination have been investigated and developed [3, 20, 21].

The challenge is to optimally support distributed, less structured and highly dynamic processes such as distributed problem solving and decision-making processes, and coordination of expert knowledge and its integration into distributed cooperation processes. This requires the development of an effective decision support environment able to improve collaborative value chains. For example, as regards the collaborative planning in inter-organisational supply chains, agents technology is also well suited for coordinating decentralised plans. In the framework of this approach, a contribution to remark presents a Cooperative Interaction method via Coupling Agents, maintaining the information of subsets of coupling constraints and resolving conflicting interests among Autonomous Agents participating on the corresponding set of constraints [12].

Moreover, a negotiation-like coordination scheme has been proposed, which can be used to synchronizing master plans performed decentrally by supply chain partners, with the use of mathematical programming models [11].

Enabling developments in Information and Communication Technologies (ICT) have made possible to extend the themes of collaboration and coordination to cross-functional business processes and across organizational boundaries to inter-organizational processes and relationships.

Moreover, web-based and mobile solutions are becoming a corporate communication infrastructure for setting up integrated information services and mobile collaborative processes in SC and logistics. In fact, they are able to creating a dynamic collaborative work environment by merging databases, information management systems, decision support systems, internet web sites and portals to provide services and new work methods for the mobile actors of transport logistics [8].
6. Conclusions

Current scenario in supply chain and logistic networks is characterised by networking organizations, industrial districts and SMEs networks which are formal or informal connected to follow the common aim of development improvement [10]. The diffusion of ICT and Internet applications provides SMEs with new tools and new commercial opportunities for the international competitiveness and e-business, and a gain in efficiency through a reduction in transaction costs. However, SMEs are performing incremental paths in increasing the web technology features and innovations, e.g. with the help of ASP (Application Service Providers) solutions [1]. On the other hand, in large companies the main problem is the hierarchical coordination of the intra-organizational supply chain, while the real challenge is the integration of inter-organizational supply chains and the cooperation in the partnership. Therefore, the task of the Supply Chain Management has to focus on how integrating organizational units along a SC and coordinating materials, information and financial flows in order to fulfil customer demands and improving competitiveness of the SC as a whole, with the use of Advanced Planning Systems [18].

Moreover, in a SC, as regards the mechanisms to manage the inter-firm relationships (such as market price, transfer prices, incentives), the real challenge is to create balance among all chain partners, with the use of collaboration platforms. Moreover, in complex environment, the coordination problems are beyond optimization and synchronization, and are typically reduced to harmonization of inter-firm networks [17].

Then, decision support systems for the design and management of collaborative supply and demand network organizations must be developed as an innovative platform able to integrate Information Technology-driven models and data mining procedures.

References


A DYNAMIC NETWORK LOADING MODEL FOR THE EVALUATION AND CONTROL OF TRAFFIC POLLUTION

Maurizio BIELLI*, Mauro DELL’ORCO**, Ergun GEDIZLIOGLU***, Murat AKAD***1

Abstract. In this paper the state of the art of traffic emission models are reviewed and a study of a mesoscopic model for a traffic network is presented, taking into account the speed variability over time. Then, implementation of proper emission models with the explicit use of the vehicles acceleration/deceleration is discussed, with aim to study an application to a real traffic environment.

1. Introduction

Several emission models have been developed and presented in the specific technical literature. Most of them (see for example Horowitz [1982] or Joumard et al. [1994]) calculate gas emission through the values of speed and acceleration in a standard urban driving cycle. For classes of instantaneous speed and for classes of speed by acceleration, the MODEM model by Joumard et al. provides instantaneous emission of four pollutants (CO, HC, NO, and CO₂) concerning vehicles with engine capacity less than 1400 cm³ to more than 2000 cm³.

Some other models consider the elementary component of vehicles movement (Crauser et al. [1989], Maurin and Crauser [1989]) in order to characterize pollutants emission (André et al. [1995], André [1996], Trigui et al [1996]). In these models, the vehicle speed is practically the only cinematic variable to be considered, but sometimes acceleration is needed to describe changes over time.

The prediction of noise, produced by road traffic, is also an important problem (Shaw and Olson [1972], Yamaguchi and Kato [1989]). As for acoustic pollution, research is generally addressed to find indices characterising the event in a global way (Favre [1984], Rapin et al. [1994]). In practice (Beruard [1986], Mitani and Ohta [1989]), the equivalent noise level

1 *Istituto di Analisi dei Sistemi ed Informatica “Antonio Ruberti”, Consiglio Nazionale delle Ricerche, Viale Manzoni 30, 00185 Rome, Italy – e-mail bielli@iasi.cnr.it
** Dipartimento di Vie e Trasporti, Politecnico di Bari, Bari, Italy
*** İ.T.Ü. – İnşaat Fakültesi, Ulaştırma Anabilim Dalı, 80 626 İstanbul Turkey
$L_{eq}$ is widely used since it allows a simple quantification of noises which may often vary in a highly non-stationary manner.

Although vehicles acceleration is then an important element for the evaluation of traffic pollution, usually existing models for Dynamic Network Loading (DNL) consider constant or stepwise variable speed (Lighthill and Whitham [1955], Merchant and Nemhauser [1978], Leonard et al. [1989], Daganzo [1994], Jayakrishnan et al. [1995], Astarita [1996], Han [2003]). Then, a new DNL model allowing the explicit calculation of acceleration should be useful for a more precise evaluation of traffic pollution.

2. The modeling approach

In the framework of a bilateral research project between Italy and Turkey, funded by Italian Research Council and Tubitak, our work aims to provide with an accurate evaluation of traffic pollution, through a model of Dynamic Network Loading that takes into account explicitly also the acceleration of vehicles. In this case, since the speed of vehicles can be calculated as a function of time and acceleration, or as a function of density, a fixed-point problem arises.

From a topological point of view, a network is made of link and nodes; so, starting from an existing validated model for Dynamic Link Loading (Dell’Orco, forthcoming), that take into account acceleration of vehicles, in the proposed research work a Node Management Model will be carried out, in order to upgrade the Link Loading model to a Network Loading model. Once obtained a validated DNL model, appropriate emission models both for gas and noise will be used, to explicitly incorporate in the models the values of acceleration.

In the case of gas emissions, the model by André and Pronello could be favourably used, because it considers instantaneous values of vehicles emissions as a function of speed classes and speed acceleration classes. For acoustic pollution, usually an index like the equivalent continuous sound level $A$ or $L_{eq}$ is calculated as the ratio of decimal logarithm of the average pressure energy and the reference pressure of $2 \times 10^{-5}$ Pa. Then, a relation between the instantaneous speed of vehicles and pressure energy has to be found, in order to incorporate the acceleration values in that index.

An accurate evaluation of traffic pollution, as a function of speed and acceleration of vehicles, allows designing more proper interventions to control and possibly to reduce the pollution itself, reaching the threshold values recommended by the World Health Organization (for example, 65 dB(A) in the case of noise pollution). The model to be implemented will be able to calculate more realistic values of emission of pollutants, with respect to other models that use the average of instantaneous values of emissions.

The proposed model will use a mesoscopic model of traffic, to calculate the Network Loading. A Node Management Model will be implemented, to pass from a Link Loading model to a Network Loading model. Nodes are the points where two or more links converge and/or diverge: a link ends at a junction, where possibly more than one link also end and/or begin. So, a Link Loading model needs a new control, whether the vehicles arrive at the end of the link or not. In the first case, a portion of arriving vehicles enters the successive link. Of course, each link arriving to the node has to be checked, considering the
possible available paths; therefore, there will be the need of a model managing the conflicts among the flows arriving from the different links to the node. The obtained model will provide with vehicles cinematic parameters, that are used as input in the pollutants emission models.

3. Expected benefits

The project carried out has its main application in control and reduction of traffic pollution. The introduction in the model of vehicles cinematic parameters, like speed and acceleration / deceleration, will lead to more realistic results in the simulation of pollution phenomena. Outcomes of the research can be easily used in the evaluation of gas and noise emissions, simulating scenarios with different parameters of traffic demand and supply, like traffic flow and links performances. Consequently, effective interventions for control and reduction of pollution are possible, with advantages in finding the best scenario to respect local or international recommendations about the thresholds values for pollutants emission, and in designing higher quality of transportation services also in terms of respect of the environment.

Therefore, the bilateral project is carried out according to the following steps:

a) Detailed study of the state of art of existing up-to-date emission models,
b) Study of a mesoscopic model for a traffic network, taking into account the speed variability over time,
c) Extension of a Link Loading model to a Network Loading model, taking into account the acceleration/deceleration of vehicles,
d) Using the DNL model developed in the previous step, proper emission models will be implemented with the explicit use of the vehicles acceleration/deceleration, and with more realistic values of pollutants emissions calculated through traffic pollution models incorporating the implemented DNL model,
e) Finally, the resulting emission models incorporating the variability over time of vehicles speed will be applied at a real case study, in order to validate the developed models.

References


MODELLING URBAN TRAFFIC NETWORKS BY A CONSUMER/PRODUCER ARCHITECTURE

Marco CALABRESE †*, Vincenzo DI LECCE.*, Michele OTTOMANELLI *,
Domenico SASSANELLI**

Abstract. In this paper a proposal for modelling urban traffic network is presented. The proposed traffic model is specified by investigating the possible isomorphism between traffic system and operating system. In this short paper we show that such an analogy is consistent and it could be helpful when dealing with dynamic network loading. In fact, with respect to a given urban traffic network (region of interest), we assume each road (link) to act as resource consumer or as buffer. The relationships with the environment outside the region of interest is modelled by a set of external producers (centroids). The work is just at the beginning and the research on practical performances of such an approach are on going.

1. Introduction

Generally, Urban Traffic Simulation (UTS) is faced by modelling the urban transportation system by means of a system of models constituted by the following sub-models [1]:
- Supply Model;
- Demand Model;
- Traffic Assignment Model.
Supply model aims to model elements, rules and relevant performances of transportation supply system.
The following entities of the transportation supply system are generally taken into account: roads, intersections and control rules [2].

† Corresponding author: calabrese.marco79@libero.it
* Dept. of Environmental Engineering and Sustainable Development – Technical University of Bari, Italy {dilecce, m.ottomanelli}@poliba.it
** Dept. of Highways and Transportation – Technical University of Bari, Italy
E-mail: d.sassanelli@poliba.it
Roads and intersections are usually modelled respectively as arcs and nodes in a digraph according to a predefined topology. Actually, graph theory is used to model not only transport infrastructures or real point of the territory but are used also to model transportation service and other relationship useful to represents transportation systems in all is aspects.

For the sake of modularity, it has been assumed that a road is a street between two consecutive intersections so that a path between a given origin-destination couple is represented by a sequence of arcs and nodes.

In our proposal (see figure 1) each road is modelled as a traffic consumer operating periodically when the traffic light is green, while it acts only as a buffer when the light is red. This assumption is general and also holds for the case when the road has no traffic light. The set of the roads considered for simulation is the region of interest. Its boundaries represents the sources of external traffic flows coming into the area under scope (region of interest). It’s fairly evident that this traffic sources play the role of external producers. On the other hand, when a vehicle passes from one street to another, the consumer related to the street where the vehicle comes from becomes an internal producer for the street where the vehicle moves to.

![Figure 1: Consumer/producer model in the case of two intersecting roads](image)

The assumptions about the function describing the traffic flow can be several. As many authors assume [3], the Poisson distribution with a time-dependent parameter lambda can be the simplest and suitable one. The time dependency, however, may be neglected if the principal aim is to model congested traffic networks. In this case, the system can be stressed simply assuming a constant high lambda.

External producers are supposed to generate traffic flows made up of different-sized vehicles picked up from a set of archetypes. With a rate given by lambda parameter, these vehicles comes into the region of interest following a pre-specified route. It’s up to the
buffer entity implementing an event-driven mechanism to maintain instantaneous buffer occupation coherence when a vehicle enters the buffer.

Buffer occupation formally represents the extent of the ineligible portions of the road due to vehicle presence at any time instant. During red period it is simply the vehicle queue forming before the intersection. The instantaneous buffer state may be difficult to model due to unpredictability of the single driver’s behaviour that can affect the local traffic flow. In first approximation however, the vehicles can be thought as one-dimension space-consuming entities moving in a predefined direction with a given average speed. Due to this assumption, the buffer is compelled to have a first-in first-out (FIFO) policy, thus neglecting other issues related to individual behaviour and street geometry. Abrupt changes imputable to undisciplined drivers can be considered as noise reducing average traffic speed.

Local congestion affects adjacent roads, self-propagating to other roads (traffic consumers) connected to the buffer. This implies that, as the flows continue being buffered, the traffic congestion spreads across the consumer/producer units. When intersection congestion occurs, every producer in front of the intersection is blocked, that is, passes from an active state to a blocked state, thus acting only as a mere buffer. An immediate parallel with reality may be this: when a car occupies an intersection without being able to move ahead, it causes a blocking or at least a bottleneck condition for any other driver who wants to cross the intersection itself. The extreme consequence of such situation is that a closed path interested by the congestion can degrade into a deadlocking condition as can be seen in figure 2.

Once a deadlock occurs, there is no chance to exit from within the congested situation, only from the contour conditions. Its fairly evident that the best strategy is

![Figure 2. Deadlock example. A vehicle stopped in a critical section determining a no-exit configuration.](image-url)
preventing these critical situations more than recovering from them. In the field of Operating Systems the deadlocking configuration has been widely studied [4][5] and it was solved with the concept of *critical section*. In a multi-threaded environment, a critical section object can be owned by only one thread at a time, which makes it useful for protecting a shared resource from simultaneous access.

In urban traffic network, the critical section is simply the intersection among crossing roads. The traffic light (assumed as control unit) might delay access to the intersection if the exiting flow is going to fill up the road (buffer) beyond the intersection. This can be achieved by increasing in advance the time of the red lights nearby the intersection. As soon as a region approaches a deadlock configuration, the affected roads may tune their traffic light parameters by consequence. This approach however regards Urban Traffic Control (UTC) and goes beyond the purpose of this paper.

2. **The isomorphism: traffic networks and operating systems**

Summarizing all the above considerations, this isomorphism can be found: the roads are modelled as traffic consumers with buffers, the intersection status (free/occupied) drives the blocking condition (active/blocked) in related consumers, the producers are both external (traffic flow parameters defined by input file or by graphic user interface) and internal (caused by vehicle passing from one road to another). Finally almost a direct correspondence can be found between intersection congestion in a closed path and deadlock configuration, although this can be considered as a theoretical extreme situation and rarely happens in reality.

For each consumer/producer unit a persistent state must be kept up during simulation. This performs a memory-consuming activity, while the control logic concerning intersections and buffer fetching requires a heavy computational effort proportional to the number of elements of the traffic net. In order to employ both computing and storing capabilities with little effort in terms of programming ease, Matlab has been considered as the appropriate framework to deploy such kind of simulation. In particular, because of its inner nature of a modelling environment at abstract level, Matlab Simulink [6] offers effective tools to model event-driven configurations. That’s the reason why our further works on implementation of a UTS system will be employed using Matlab technology. In particular, Matlab offers several means to render surfaces on maps, which can be useful if we want to visualize traffic peaks on a given urban topology.

The proposed architecture for UTS is distributed, very scalable and easy to implement. The simple model herewith described can be easily exported to grid architecture for massive simulation on large traffic networks. A real sized urban transportation network can be split in an appropriate number of regions of interest each one deployable on a grid node.
3. Conclusion and further research

In this contribution we proposed a new insight for model real-sized traffic network. In particular, we have pointed out to review the traffic supply system from the view of Operating Systems. We have shown that such an analogy is theoretically consistent so that network congestion problem can be solved by using the methods of O.S. field.

This research is on going and the first results achieved using a test network are very interesting. The simple model herewith described can be easily exported to grid architecture for massive simulation on large traffic networks for dynamic network loading problem.

References


PRESERVING UNCERTAINTY IN MODELING CHOICE BEHAVIORS: FROM POSSIBILITY TO PROBABILITY

Mauro DELL’ORCO¹, Shinya KIKUCHI²

Abstract. This paper focuses on how to pass from choice Possibility to choice Probability. Traditionally, the modeling framework has been rooted in Probability Theory, in which both the analyst’s uncertainty about the explanatory capability of the model and decision-maker’s uncertainty, due to the lack of complete information about choices, are expressed in terms of probability. However, while the analyst’s uncertainty is related to the statistical nature of events, the decision maker’s uncertainty has rather a perceptive nature. Therefore, a modeling framework using the Possibility Theory seems more proper to account for this kind of uncertainty. The results of this modeling framework are obviously expressed in terms of Possibility, and need to be transformed into Probability values, in order to be usable by analysts. In this work, we use the Principle of Uncertainty Invariance to perform this transformation, preserving the Uncertainty embedded in human choice process. The paper discusses the thought process, mathematics of possibility theory and probability transformation, and examples.

1. Introduction

In response to a given a set of stimuli humans, individually or collectively, make certain choices. To predict these choices, models of human stimulus-response pattern are used. Traditionally, the stochastic modeling framework has been used to estimate human choice

¹ Dept. of Highways and Transportation, Technical University of Bari, Italy. E-mail: dellorco@poliba.it
² Dept. of Civil and Environmental Engineering, Virginia Polytechnic Institute and State University. Northern Virginia Center - Falls Church, VA - USA, e-mail: kikuchi@vt.edu
based on such models. Various formulations of stochastic models, with different degrees of sophistication, have been devised to handle the uncertainty embedded in human choice process.

In this paper, we use an alternative approach to model human preference and choice by using Possibility Theory. This approach focuses on two aspects of uncertainty that is in the mind of the decision-maker: one related to his lack of complete understanding about the alternatives, and another one related to his lack of clarity about goals or objectives while selecting alternatives. In the proposed approach, decision-maker’s utility of an alternative is therefore treated as the fuzzy number. Hence, evaluation of the alternatives is based on comparing utilities expressed in fuzzy numbers. As a result, one’s inclination to make a particular choice is expressed in possibility and necessity measures. The paper then shows that possibility measure can be transformed to the probability measure through an appropriate mathematical treatment (uncertainty invariance), so that the choice can still be predicted in terms of probability, which is useful for practical application. In this paper, human choice in transportation options has been considered, but the underlying logical basis should be valid for any modeling of human choice.

2. Motivation and Background

2.1. Motivation

In the traditional stochastic choice models, the utility of an alternative is generally expressed by a linear combination of attributes in two parts: one, a set of values of the attributes of the alternative, called the fixed term; another, a random term, which is expressed by a random variable with a certain probability density function. Let \( V_A \) be the fixed term of utility of alternative \( A \) and \( \varepsilon \) its random term, the value \( U_A \) of utility of alternative \( A \) can be expressed as follows:

\[
U_A = V_A + \varepsilon \quad (1.)
\]

Calling \( x_{Ai} \) the value of \( i \)-th attribute of the alternative \( A \), the fixed term can be express as a linear combination of the attributes of that alternative with each attribute weighted by a coefficient \( a_i \):

\[
V_A = a_0 + a_1 x_{A1} + a_2 x_{A2} + a_3 x_{A3} + \ldots + a_n x_{An} \quad (2.)
\]

and, then:

\[
U_A = a_0 + a_1 x_{A1} + a_2 x_{A2} + a_3 x_{A3} + \ldots + a_n x_{An} + \varepsilon \quad (3.)
\]

In this formulation, all kind of uncertainty is gathered in the random term \( \varepsilon \): possible errors in measurements, variations of users’ taste, analyst’s lack of knowledge about the completeness of the model, user’s lack of complete understanding about the alternatives, and lack of clarity about goals or objectives.

From Eq. (3), with the hypothesis that the random terms are independent and identically distributed (IID hypothesis) according to the Weibull-Gumbell distribution, the probability that the \( k \)-th decision maker chooses the alternative \( i \) is given by:
\[ p_\alpha = \frac{e^{-\alpha V_\alpha}}{\sum_{j} e^{-\alpha V_j}} \]  

where \( n \) is the total number of alternatives, and \( \alpha \) is a calibration parameter. The Eq. (4) is called Logit model. The IID suggests that the analyst’s degree of uncertainty about the representation of the choice situation is the same for all the alternatives. Basically, this means that the alternatives are compared by their fixed terms, and the variance of the utility (or decision-maker’s uncertainty about the alternatives) does not affect the final choice probability. As a consequence, alternatives having the same utilities with different uncertainties will have the same probability.

In summary, in stochastic models only the fixed terms of the attributes of the alternatives affect the probability of choice. Consequently, the degree of the analyst’s uncertainty does not appear to matter in the end. A different mathematical framework is then needed to model different dimensions of uncertainty.

2.2. Background

In choice processes, the knowledge of alternatives is rarely complete and precise. Therefore, uncertainty affects any decision. Traditionally, randomness has been used to deal with uncertainty, and therefore, it is the potential for manifestation of different choices. Starting in early 1970’s, from seminal works by Domencich and McFadden (1975), several scientists carried out relevant theoretical works about random utility models and, in particular, about the multinomial logit model. Theoretical fundamentals of random utility models have been analyzed by Stopher and Meyburg (1976), Williams (1977), Ben Akiva and Lerman (1985). Nested logit models have been investigated by Ortuzar (1983) and Sobel (1980), and the probit model has been deeply analyzed by Daganzo (1979).

Uncertainty embedded in different situations of choice has been studied by these stochastic approaches: De Palma et al. (1983) developed a model for stochastic equilibrium for departure time choice. Afterwards, Ben Akiva et al. (1986) extended this model including route choice and the option of not making the trip. More recently, Cascetta (1989) has analyzed day-to-day dynamics and Cascetta and Cantarella (1991) within-day dynamics in a transportation network. Influence of Information - or its dual uncertainty - on users’ behavior has also been studied by means of simulation frameworks proposed by Kaysi (1991) for ATIS services and by Hu and Mahamassani (1995).

During the same period, a set of new paradigms of uncertainty was being developed. This development started with Fuzzy Set Theory in the late 1960’s and Evidence Theory in the 1970’s. Different measures of uncertainty emerged in the 1980’s, and in the 1990’s, treatment of Uncertainty has been systematized by Klir (1998). Among the new theories, Possibility
Theory is said to be amenable to the framework for representation of human perceptive uncertainty. This point has been suggested by prominent systems scientists such as Shackle (1969; 1979) and Cohen (1970). They argue that the traditional approaches for choice modeling using Probability Theory do not completely represent the true level of uncertainty in people’s behavior. These new theories, instead, have provided a better insight into the understanding of uncertainty and definitely redefined the place of probability theory when dealing with uncertainty.

3. The Possibilistic Approach

In the following, we assume that readers have some knowledge about Fuzzy Set Theory and Possibility Theory; anyway, fundamentals of Fuzzy Set Theory are found in Klir and Yuan (1995), Zadeh (1978), and Klir (1993). For the Dempster-Shafer Theory, which subsumes Probability and Possibility Theories, a number of references is available, among them are Shafer (1976) and Yager et al.(1994).

In the present approach, we introduce fuzzy numbers to represent the decision-maker’s uncertainty about the attributes of individual alternatives. Utility of each alternative, thus, becomes a fuzzy number, as shown in the following. Moreover, we introduce a comparison among the fuzzy utilities when they are very close one another.

Consider n alternatives, $A_1$, $A_2$, ..., $A_n$. The generic alternative $A_i$ is characterized by a vector of m attributes $\tilde{X}_1, \ldots, \tilde{X}_m$. Thus, utility of alternative $A_i$ is expressed by: $\tilde{U}_i = a_1i \cdot \tilde{X}_1 + a_2i \cdot \tilde{X}_2 + \ldots + a_mi \cdot \tilde{X}_m$, which is similar to Eq. (1). However, now the values assigned to $\tilde{X}_1, \ldots, \tilde{X}_m$ can be either an exact number or a range (or an approximate number). In our approach, both of these values are treated as a fuzzy number (or a fuzzy set).

Utility is now represented by a sum of fuzzy numbers; this means that its value is a range characterized by the membership function.

4. Comparison of fuzzy utilities

The best alternative from the decision-maker’s point of view is now obtained comparing the fuzzy values of utility of each alternative. This comparison is not always simple, especially when the ranges of the utilities of two alternatives overlap. However, it is possible to set up a procedure allowing the calculation of the degree to which one utility value (a fuzzy value) is greater than the other. Suppose that fuzzy utilities of two alternatives $A$ and $B$ are denoted $\tilde{U}_A$ and $\tilde{U}_B$; the problem is how to compare these two fuzzy numbers. Let $W$ be the set
representing “less than $\tilde{U}_B$”. Then, the degree to which (or Possibility) that $\tilde{U}_A$ is a subset of $W$ indicates the degree that $\tilde{U}_A$ is less than $\tilde{U}_B$, and can be calculated by:

$$\text{Poss} (\tilde{U}_A < \tilde{U}_B) = \max \min (h_A(x), \pi_B(x)) \text{ for } x \in X \quad (5.)$$

where $X$ is the universe, $h_A(x)$ is the membership function of fuzzy set $\tilde{U}_A$, and $\pi_B(x)$ is the possibility distribution derived from fuzzy set “less than $\tilde{U}_B$”. In other words, $\pi_B(x) = h_{<B}(x)$.

Using the principles shown in Eq. (5), Figure 1 illustrates the way to calculate $\text{Poss}(\tilde{U}_A \leq \tilde{U}_B)$ and $\text{Poss}(\tilde{U}_A \geq \tilde{U}_B)$ for given $A$ and $B$.

**Figure 1. – Comparison of fuzzy utilities**

5. **Possibility to Probability Conversion: Uncertainty Invariance**

Although the Possibility represents better the decision-maker’s uncertainty about the attributes of individual alternatives, its values cannot be used directly by the analysts; a conversion to Probability values is needed. The idea is to move from one theory to another on the basis of a justifiable principle. Klir (1990) proposed the principle of uncertainty invariance, subsequently systematized by Klir and Wang (1992). This principle specifies that the uncertainty in a given situation should be the same, whatever is the mathematical framework used to describe that situation.

In the traditional probabilistic choice modeling situations, the amount of uncertainty associated with the probabilities can be assessed in terms of entropy of the probability distribution, the so-called the Shannon entropy, given by:

$$H = \sum (p_j \log_2 p_j) \quad (6.)$$

where $H$ is the measure of entropy and $p_j$ is the probability of choosing the alternative $j$.

The same choice situation can be represented in the context of the Possibility Theory. The principle of uncertainty invariance states that the uncertainty measure associated with the
possibility distribution should be equivalent to the entropy measure of the corresponding probability distribution.

Given an ordered Possibility distribution \{\pi_1, \pi_2, \ldots, \pi_n\} for which is always the case that \pi_i \geq \pi_{i+1}, the possibilistic counterpart of the Shannon entropy, called U-Uncertainty, is given by the following function (Klir, 1988):

\[
U = \sum_{i=1}^{n} (\pi_i - \pi_{i+1}) \log_2 i
\]  

(7.)

where \pi_{n+1} is always 0 by convention.

According to the Uncertainty Invariance Principle, both the H and U in Eq. (6) and (7), respectively, contain the same amount of uncertainty, that is:

\[
\sum_j (p_j \log_2 p_j) = \sum_i (\pi_i - \pi_{i+1}) \log_2 i
\]  

(8.)

Additionally, Possibility and Probability distributions are subject to the normalization requirement, requiring that:

\[ \sum p_j = 1 \] (probabilistic normalization)

\[ \max (\pi_i) = 1 \] (possibilistic normalization)

Information-preserving transformation have been discussed by Klir (23) for different scales. It has been found that ratio and difference scales possess only one free coefficient, which can be determined by either probabilistic or possibilistic normalization. Therefore, the uncertainty equivalence expressed by Eq. (8) is ineffective. Instead, log-interval scale transformations of the form:

\[
\pi_i = \beta \cdot (p_j)^\gamma
\]  

(9.)

with \alpha and \beta positive constants, have two free coefficients that can be determined respecting both normalization requirements and the equivalence in Eq. (8). In fact, from Eq. (9) we obtain

\[
p_j = (\pi_i / \beta)^{1/\gamma}
\]  

and, applying the probabilistic normalization requirement:

\[
\beta = (\sum_j (\pi_i)^{1/\gamma})^{\gamma}
\]  

(10.)

Hence:

\[
p_j = \pi_i^{1/\gamma} / (\sum_j (\pi_i)^{1/\gamma})
\]  

(11.)

Replacing in Eq. (8) the expression of \pi_i obtained by Eq. (11), we have:

\[
\sum_i \pi_j^{1/\gamma} \cdot \log_2 \pi_j^{1/\gamma} - \sum_j \pi_j^{1/\gamma} \cdot \log_2 \pi_j^{1/\gamma} - \sum_i [\pi_i - \pi_{i+1}] \cdot \log_2 i = 0
\]  

(12.)

Then, solving numerically the equation (12), the constant \alpha, can be calculated, and the probability value is found through Eq. (11).
6. Example: Use of the Proposed Framework

A survey has been carried out in the London Underground, among travelers going from Leicester Square to somewhere between Euston Road and King’s Cross. Travelers can choose either the path A (Leicester Square – Tottenham – Euston), or the path B (Leicester Square – Holborn – King’s Cross), as shown in fig. 2. The actual travel time is approximately the same for the two paths – 6 min and 7 min, respectively. From the survey came out that the perceived travel times were two fuzzy numbers with an approximately trapezoidal shaped membership function (fig. 3). Numerically, the Possibility distribution produced by the two fuzzy numbers can be express in the following tables:

<table>
<thead>
<tr>
<th>time (min)</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td>Possibility</td>
<td>0.00</td>
<td>0.83</td>
<td>0.99</td>
<td>1.00</td>
<td>1.00</td>
<td>0.89</td>
<td>0.86</td>
<td>0.79</td>
<td>0.74</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Table 1. – Possibility of perceived travel time to Euston Road

<table>
<thead>
<tr>
<th>time</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
</tr>
</thead>
<tbody>
<tr>
<td>Poss.</td>
<td>0</td>
<td>0.82</td>
<td>0.85</td>
<td>1</td>
<td>0.97</td>
<td>0.95</td>
<td>0.88</td>
<td>0.83</td>
<td>0.82</td>
<td>0.77</td>
<td>0.70</td>
<td>0.70</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 2. – Possibility of perceived travel time to King’s Cross

Let us consider the choice between path A and path B. Provided that the destinations are equally distant from both Euston Road and King’s Cross, we can neglect the walking time at destination. Therefore, the utilities \( \tilde{U}_A \) and \( \tilde{U}_B \) are only the perceived travel times to Euston Road and to King’s Cross, respectively. From fig. 3 it easy to see that \( \text{Poss}(\tilde{U}_A > \tilde{U}_B) = \pi_2 = 0.94 \), and \( \text{Poss}(\tilde{U}_A < \tilde{U}_B) = \pi_1 = 1 \).
What we see above is the possibilistic views of preference of each path; in other words, for example, the possibility that the sentence “the path B is the preferred one” is true is 0.94. Therefore, based on the previous computation, the alternative A has the higher preference, and would be the route preferred by a traveler who has no other biases in choosing the route.

Replacing the numerical values in the Eq. (12), we obtain the Eq. (13):

\[
\frac{1^{1/\gamma}}{1^{1/\gamma} + 0.94^{1/\gamma}} \log_2 \frac{1^{1/\gamma}}{1^{1/\gamma} + 0.94^{1/\gamma}} + \frac{0.94^{1/\gamma}}{1^{1/\gamma} + 0.94^{1/\gamma}} \log_2 \frac{0.94^{1/\gamma}}{1^{1/\gamma} + 0.94^{1/\gamma}} - 0.94 \log_2 2 = 0 \quad (13)
\]

Afterwards, solving numerically this equation, we can calculate \( \gamma = 9.5 \), and \( 1/\gamma = 0.105 \). Consequently, from the Eq. (11) the probabilities of choosing path A and B are, respectively:

\[ p_A = 1/1.99 = 0.501; \]
\[ p_B = 0.99/1.99 = 0.499. \]

\[ \text{Figure 3. – Comparison of perceived travel times} \]

What is shown here is a simple example that involves two alternatives of choice (paths), and the travel time as the only attribute or decision criterion. Each value of the travel times, however, has been considered as an interval incorporating the notion of uncertainty. Possibilities that each path will be chosen by the decision-maker are computed. Finally, the corresponding probability of choice is derived. We believe that this presentation provides more information about the preservation of uncertainty of the estimate.
7. Conclusion

When the decision-makers’ information about the attributes of the alternatives is approximate or incomplete, the classical stochastic choice models are unable to produce a range of possible values for choice probability. For example, when applying a survey for construction of a choice model, the decision-makers are asked about choice of either the past behavior or an hypothetical situation. Although the information on the attributes of alternatives is inevitably not clear or complete in these cases, the strict requirements of the probability theory dictate the mathematics of deriving the choice probability. Hence, the vagueness in the information that the decision-maker originally harbored is not reflected in the model building process. It appears that the classical probabilistic approach mixes fuzziness in perception by the decision-maker and randomness of the manifested outcome in one mathematical framework.

Instead, introducing fuzzy set for representation of uncertainty values of the attributes allows us to separate the uncertainty experienced by the decision-maker and the uncertainty that the modeler feel as to the incompleteness of the model. Using fuzzy numbers for utility as the modeling framework for choice, one can predict choice that consistently preserves uncertainty.

The approach to model choice used in this paper differs from the traditional ones in three aspects. First, the decision-maker’s uncertainty about the information on the attributes is accounted for by the use of the approximate number (fuzzy number). Second, the values of utility of different alternatives (expressed in fuzzy numbers) are compared in terms of the possibility measures. Third, the possibility based comparison of the utilities are converted to the probability that utility of an alternative is greater than the other, using the technique of Uncertainty Invariance.

The proposed framework will be useful in dealing with choice situation when the information about the value of attributes is incomplete, e.g., in approximate number, interval, or linguistic expression. The approach can respond to the degree of accuracy of the data, and yields choice probabilities faithfully to the quality of information. It is the authors’ belief that different mathematical frameworks are possible depending on the nature of uncertainty embedded in the problem situation. The authors do not undermine the classical stochastic choice model approach if the nature of uncertainty in the problem is purely probabilistic. This research presents what possibility theory and uncertainty invariance can offer in dealing with situation in which the quality of information about the attributes and the decision-maker’s perception of the information is in question.

References


Williams, H. C. W. L. On the formation of travel demand models and economic evaluation measures of user benefit Environment and Planning 9, 1977.


Ortuzar, J. O. Nested logit models for mixed-mode travel in urban corridors Transportation Research, 17A. 1983


Daganzo, C. F. Multinomial probit: the theory and its application to demand forecasting, Academic Press, New York, 1979,

De Palma, A., Ben Akiva, M., Lefevre, C., and Litinas, N. Stochastic equilibrium model of peak period traffic congestion Transportation Science 17, 1983


Cascetta, E. A stochastic process approach to the analysis of temporal dynamics in transportation networks Transportation Research, 23B(1), 1989


Hu, T. and Mahamassani, H. Evolution of network flows under real time information: a day-to-day dynamic simulation assignment framework, 74th Annual Meeting of Transportation Research Board, 1995


PETRI NET MODEL OF TRANSHIPMENT ACTIVITIES EXECUTED BY HUMAN OPERATORS IN A CONTAINER TERMINAL

G. Maione¹, M. Ottomanelli¹

Abstract: This paper proposes a Petri net representation of the activities performed by the most important human operators during processes of downloading/loading containers in an intermodal terminal. These processes are the core of export, import, and transshipment cycles executed in the terminal. The aim of this research is to take into account the human component together with the material handling resources, i.e. cranes, trailers, stacking areas, etc., such that a complete model could be defined to describe how all the system resources are used and coordinated to fulfil the task of serving vessel or feeder ships. The developed model can be translated to a simulation environment, where it is useful to validate the approach, to test the efficiency of the processes, and to propose alternative solutions to the different problems that could arise to optimize the activities of the terminal.

1. Introduction and literature review

Achieving the best performance of a transportation systems is one of the goals for transport managers and planners. An important element of the general intermodal transportation system are the container terminals that can be considered as complex systems instead of merely node of the intermodal network. In fact, transshipment terminal are complex systems whose performances affect, also in strong way, the whole transport system. In container terminals take place processes that should allow the integration of maritime, rail and road service, technology and human factor (in different measure). Since fully automated terminal are not a common practice, in such a processes an important role is covered by human operators that operate to integrate and support the terminal service chain as an important ring of it.

¹ Dept. of Environmental Engineering and Sustainable Development – Technical University of Bari. Viale del Turismo, 8 – 74100 Taranto (Italy). {gmaione, m.ottomanelli}@poliba.it
The aim of defining the best organization of a container terminal has been collecting great attention in literature, by analysing different solutions to different problems concerned. The problem of containers transfer from ship to stack and vice-versa is discussed by many authors considering both full-automated and not automated transfer systems [6].

The allocation of ships to berths consist in assigning incoming ships to berthing positions in order to minimize the waiting time of the vessel. Problem solutions have been discussed in [7],[8]. The ships loading/unloading operations cover many problems, for example the quay cranes assignment [12], the unloading plan [9], [10], the stowage plan [11].

Simulating the whole system suffers complexity since many problems are linked by shared resources whose availability is time dependant. Nevertheless, efficiency of terminal management can be achieved by means of system simulation. Namely, it is possible to design and evaluate control strategies, different system layouts, new and powerful technologies and processes. To this aim, a resources allocation tool is proposed in [13] where the model is based on operations research methods and allows managers to develop resources allocation plans. A distributed discrete-event container terminal simulator based on object-oriented modeling technique was also recently proposed in [14]. A comprehensive review of terminal containers simulation models is given in [6].

Petri Nets (PN) have been used in [15],[16],[17] to model and analyze container terminal systems with respect to different aspects of terminal operations. In [15] a stochastic PN models containers transfer in a railroad transshipment yard. In [16][17], PN describe the synchronization problems among resources, and the analysis of faulty situations during operations at terminal. Actually, few attention has been addressing to the function of the human resources within the process chain in which many operation need of the operation by different agents whose role may affect the achievement of the expected performances. In the following, as part of a research project aimed to a more sophisticated container terminal simulation model, we face the problem of modeling the human agent. This paper proposes a Petri Net representation of the activities performed by the main human agent during processes of downloading/loading containers in an intermodal terminal. These processes are the core of export, import, and transshipment cycles executed in the terminal. The aim of this research is to take into account the human component together with the material handling resources, i.e. cranes, trailers, stacking areas, etc., such that a complete model could be defined to describe how all the system resources are used and coordinated to fulfil the task of serving vessel or feeder ships. The developed model can be translated to a simulation environment, where it is useful to validate the approach, to test the efficiency of the processes, and to propose alternative solutions to the different problems that could arise to optimize the activities of the terminal transshipment process.

1.1. The main processes handling containers

At this time, the real transshipment process begins involving terminal resources and operators. Usually, containers are landed from the ship to trailers, transferred to yard blocks, stacked in array positions. After a certain amount of time in which TEUs are kept in these positions or are moved in other more appropriate locations (consolidation process), containers are picked-up by trailers and delivered to the quay area where they are embarked on ships leaving for another destination.
Activities in the yard area and the distribution of containers in the stacking blocks are managed and supervised by a yard planner, who decides in which block-row-tier position to allocate each container.

Sometimes, containers are landed, stacked and then delivered to trains departing from the terminal and connecting to the railway system, or to trucks leaving for roadways. Also, cargo may arrive on trains or trucks entering the terminal. Activities in the railway connection area and in the gate area where the trucks enter/exit, are organized by a rail and a truck planner, respectively.

In synthesis, the main processes in the terminal area are:

- Unloading containers from vessel or feeder ships (landing)
- Loading containers on ships (embarking)
- Internal transport of containers from quay to yard area, and vice versa
- Loading containers on trains/trucks
- Unloading containers from trains/trucks

The organization of these processes may be modelled by PNs, as illustrated in [1], in which the focus was on containers movement, represented by the token flow through the net.

1.2. Human resources and material handling equipment

Several specialized operators work to fulfil different tasks for moving containers in the terminal environment, according to pre-specified rules and methods. Basically, we may distinguish two class of operators:

- Office operators working in planning activities
- Operators working in quay or yard area.

There exist communication links between these two classes of operators typically using PDAs. The current containers location is stored in central databases, and moving and stacking strategies are selected and managed by a central data processing supervisory system. We will skip the ICT details and we will concentrate on the activities executed by the second class of operators.

In particular, we will focus on the activities executed in unloading/loading processes. For each active quay crane, a team of 25 people is involved: 1 foreman, 2 checkers, 2 quay crane operators, 4 raisers, 5 trailer drivers, 8 yard crane operators, 1 yard operator, 1 reach stacker driver, 1 side loader driver.

The foreman (Fo for brevity) is head of the team. He receives instructions and messages from the ship planner, to start embarking or landing containers. He sends the container identification code to checkers and raisers. He compiles reports to the planning team, for example when a damage occurs to an unloaded container.

Checkers monitor the process: the first one, Ch1, works inside the kiosk of the quay crane, retrieves (or stores) and verifies container information from database, and communicates to the crane operator the bay-position of the container to be unloaded/loaded; the second checker, Ch2, is located below the crane, verifies external conditions and type of containers, and makes containers to be embarked ready to be fixed or sets them free from fixtures.
The first quay crane operator, $Qo1$, is actually operating the crane to pick-up the container from a trailer (ship cell) and put it in a ship cell (trailer), during the embarking (landing) cycle; the other, $Qo2$, works on board to help his colleague during manoeuvres.

Raisers ($Ra$ for brevity) lock/unlock containers on/from the ship by special fixtures and constraints. They receive commands from $Fo$ to fix (or free) containers to (or from) their locked positions in the ship hold or cover.

Trailer drivers ($Td$ for brevity) deliver containers from/to the yard side to/from the quay crane. Each driver always waits a start-signal from $Ch2$, after receiving or delivering the container and before leaving for the yard.

Yard crane operators pick-up containers from trailers or trucks and deposit them in a yard block, or pick-up them from the block and load them on transport vehicles. The yard operator manages a parking area for trailers: he opens it and registers the trailers going in or going out, then he closes the parking. He also verifies the efficiency of transport vehicles, quay and yard cranes. Finally, he schedules turn-overs, on and off-duty periods, and vacations. Reach stacker and side loader drivers operate particular cranes to stack or retrieve containers in or from the blocks. Obviously, they work together with trailer and truck drivers.

In this paper, for sake of space, we only model the activities of $Fo$ and $Ch2$, while those executed by quay crane operators, raisers and trailer drivers in the quay side are just generally described. The formalization of other operations can be easily derived and will be specified in future works.

2. Petri net modelling of human operations

The Petri Net modelling the considered human activities can be defined by integrating different modules, each associated with a specialized operator. Each module can be easily defined by the standard construction rules of Petri nets. Here, we developed the module for $Fo$ and $Ch2$. For sake of space, we synthesized the modules representing the activities executed by quay crane operators, raisers, and trailer drivers by using incidence matrices. Future works may develop all the modules and merge them with those associated to the operation of unmanned resources in the same transshipment processes (see [1], [2]). The complete model could be converted to a discrete-event simulation environment, both for validation purpose and for testing the terminal efficiency. Now, we re-elaborate some useful formal notations about Petri nets.

2.1. Theoretical background

A PN is a bipartite graph, basically formalized by a five-tuple $PN = (P, T, A, W, M_0)$, in which nodes are determined by places in set $P$ and transitions in set $T$, and directed edges or arcs constitute set $A$ and link places and transitions together [3]. Each place, graphically depicted as a circle, represents a condition. Transitions, showed as bars, represent events changing the state. All arcs are drawn as arrows.

Input places to a transition $t$ are associated to the pre-conditions for that event to occur, or the resources needed for a certain task. Output places from $t$ are associated to the post-
conditions consequent to the event, or the resources released after the task. The pre-
conditions enable $t$, the post-conditions are verified after the occurrence of $t$.

Usually, places model the execution of activities or the availability of resources, where
each resource $r_i$ has a finite capacity $C(r_i)$ (number of units). A token (black dot) marking a
place indicates the truth of the condition associated to that place, or an available resource
unit. Therefore, the system state is defined by the distribution of tokens in the net, and the
state changes are triggered by transitions, corresponding to tokens flow through the arcs.
Each arc is labelled by a weight function $W: A \rightarrow \{1, 2, \ldots\}$, specifying how many tokens flow
through the arc. Unity weights are omitted.

The state is defined by the marking vector $M: P \rightarrow \{0, 1, 2, \ldots\}$, where $M(p_i)$ gives the
current number of tokens in $p_i \in P$. No token resides in an unmarked place and $M(p_i)$ tokens
in a marked place $p_i$. $M_0: P \rightarrow \{0, 1, 2, \ldots\}$ is the initial marking. The markings change
according to two basic rules:

- **enabling rule**: a transition $t$ is enabled if each of its input places $p$ is marked with at
  least $w(p, t)$ tokens, where $w(p, t)$ is the weight of arc from $p$ to $t$, but an enabled
  transition may or may not fire (depending on whether or not the event actually takes
  place);

- **firing rule**: a firing of an enabled transition $t$ consumes $w(p, t)$ tokens from each input
  place $p$ and produces $w(t, p)$ tokens for each output place $p$, where $w(t, p)$ is the weight
  of arc from $t$ to $p$. We assume that only one transition fires at a time.

In addition, to allow the representation of a decision, choice or conflict, a place may be
connected to more than one output transitions, which, in this way, represent nondeterministic events, each associated with a probability or possibility of occurrence (e.g. transitions $t_{31}$ and $t_{32}$ in figure 1).

We consider the operation cyclically executed by human operators in the same way as
production sequences of workstations in manufacturing systems. Then, human activities are
specified by a sequence of interleaved transitions and places.

Each timed transition corresponds to a step that needs time to be executed in the
operator cycle. Each immediate transition occurs instantaneously, and it is associated to
logical state changes with zero firing time. Each place is related to a logical condition of the
operator, or to the availability of a resource. This common interpretation allows us to build-
up a Generalized Stochastic Petri Net [4], which is useful to evaluate the performance of
the modelled processes. It is defined as $\langle GSPN, FIR \rangle$, where $GSPN$ comes from the net $PN$
when $T$ is partitioned in a set $T_I$ of immediate transitions and a set $T_E$ of timed exponential
transitions. Immediate and exponential transitions are depicted as white and black bars,
respectively. Besides, $FIR$ is a firing function for exponential transitions [5]. The firing time
of $t \in T_E$ is a continuous random variable with exponential distribution. Then, for each $t \in T_E$
in every reachable marking $M$, the real number $FIR(M, t)$ establishes the rate of firing of $t$.
The mean firing time of $t \in T_E$ gives the average duration of the associated process.

### 2.2. PN model of the foreman activities

The flow of tokens in the PN of figure 1 describes the foreman activities as follows.
Figure 1 depicts the foreman initial state, before starting his operation in an
unloading/loading task.
Fo receives information from the ship planner about the containers to be unloaded/loaded (t1). Information includes the ship arrival time, where the ship is moored, the unloading/loading plan, the working schedule, the quay crane to manage. No other foreman will have this crane assigned for other contemporaneous unloading/loading operations. After all necessary data are received (p1), if the quay crane QC and both checkers Ch1 and Ch2 are available, Fo commands and supervises setting the quay crane in a proper position (t2), with the help of checkers (arc from Ch to t2 has a weight of 2). After the crane is ready, he starts (p2) to check if it is working fine (t31 and t32 represent two conflictive immediate events): if not, he starts (p3) a restoring procedure and waits for a recovery period (t4), until the crane is repaired; if the crane works, Fo begins (p4) to bring tools close to the crane (t5), as tools are necessary for freeing/fixing unloaded/loaded containers. When he has completed this task (p5), he lets the sanitary check begin (conflict between t61 and t62): if some health problems have occurred, then he starts (p6) a resolution procedure and waits for a health recovery period (t7), until every problem is solved; if nothing happened, then Fo is ready to begin the operations on board (p7). To start them, he needs the availability of 4 raisers (arc from Ra to t8 has a weight of 4) and two quay crane operators (arc from Qo to t8 has a weight of 2), to whom he sends a start-signal (t8) to let them get on board.

Therefore, Fo starts (p8) to verify if the containers distribution according to the ship officer matches the plan made by the ship planner (conflict between t91 and t92): if this is not the case, Fo begins (p9) to manage a reorganization of loading/unloading tasks to the assigned quay crane, for which he takes some time (t10); otherwise, Fo may start (p10) the
normal supervision of loading/unloading processes, which undergoes a certain amount of time \((t_{11})\). The operations are triggered when \(Fo\) signals the containers to be loaded/unloaded to checkers, quay crane operators and raisers. When these processes are completed \((p_{11})\), \(Fo\) sends a signal \((t_{12})\) to the four raisers (arc from \(t_{12}\) to \(Ra\) has a weight of 4) and to \(Qo\) (arc from \(t_{12}\) to \(Qo\)), such that they may get down from the ship. When all on-board operations are finished \((p_{12})\), \(Fo\) commands \((t_{13})\) the checkers and \(Qo\) to set the crane in rest position, after which they will be available again (arc from \(t_{13}\) to \(Ch\) of weight 2, arc from \(t_{13}\) to \(Qo\)), and the crane will be ready for new operations (arc from \(t_{13}\) to \(QC\)). The foreman activity terminates in \(p_{13}\).

Note that in figure 1 some places, namely \(QC, Ch, Qo,\) and transition \(t_8\) are duplicated, for sake of clarity, but they do represent the same condition and should be considered one time only. For the same reason, the net is split in two parts. Table I synthetically describes transitions and places.

<table>
<thead>
<tr>
<th>Transition</th>
<th>Event</th>
<th>Place</th>
<th>Condition</th>
</tr>
</thead>
<tbody>
<tr>
<td>(t_1)</td>
<td>Receiving info from SP</td>
<td>(p_1)</td>
<td>Info available</td>
</tr>
<tr>
<td>(t_2)</td>
<td>Setting QC ready</td>
<td>(p_2)</td>
<td>Ready for functionality check</td>
</tr>
<tr>
<td>(t_{31})</td>
<td>Observing QC normally working</td>
<td>(p_3)</td>
<td>Start waiting QC repair</td>
</tr>
<tr>
<td>(t_{32})</td>
<td>Observing QC fault</td>
<td>(p_4)</td>
<td>Start recovering tools</td>
</tr>
<tr>
<td>(t_4)</td>
<td>Waiting QC repair</td>
<td>(p_5)</td>
<td>Ready for health check</td>
</tr>
<tr>
<td>(t_5)</td>
<td>Getting tools to quay</td>
<td>(p_6)</td>
<td>Start waiting health problem solution</td>
</tr>
<tr>
<td>(t_{61})</td>
<td>Observing no health problem</td>
<td>(p_7)</td>
<td>Start on-board activity</td>
</tr>
<tr>
<td>(t_{62})</td>
<td>Observing health problem</td>
<td>(p_8)</td>
<td>Start checking bay plan</td>
</tr>
<tr>
<td>(t_7)</td>
<td>Waiting solution of health problems</td>
<td>(p_9)</td>
<td>Start redistributing tasks to QCs</td>
</tr>
<tr>
<td>(t_8)</td>
<td>Command operators to get on board</td>
<td>(p_{10})</td>
<td>Start supervising unloading/loading process</td>
</tr>
<tr>
<td>(t_{91})</td>
<td>Observing match of bay plan</td>
<td>(p_{11})</td>
<td>End supervising unloading/loading process</td>
</tr>
<tr>
<td>(t_{92})</td>
<td>Observing need of bay redistribution</td>
<td>(p_{12})</td>
<td>End on-board activity</td>
</tr>
<tr>
<td>(t_{10})</td>
<td>Redistributing tasks to QCs</td>
<td>(p_{13})</td>
<td>End of activity</td>
</tr>
<tr>
<td>(t_{11})</td>
<td>Supervising unloading/loading process</td>
<td>(QC)</td>
<td>QC available</td>
</tr>
<tr>
<td>(t_{12})</td>
<td>Command operators to get down</td>
<td>(Ch)</td>
<td>Checkers available</td>
</tr>
<tr>
<td>(t_{13})</td>
<td>Repositioning of QC</td>
<td>(Qo)</td>
<td>QC operators available</td>
</tr>
</tbody>
</table>

Table I. Places and transitions in the foreman GSPN module

2.3. PN model of the second checker activities

The PN module for the activities normally executed by \(Ch2\), below the quay crane, is specified in figure 2. Even if places and transitions are consecutively numbered as in the \(Fo\) PN module, they model different conditions and events.

\(Ch2\) receives assignments for loading/unloading containers from \(Fo\) \((t_1)\). When ready \((p_1)\), he checks the trailer position below the quay crane (conflict between \(t_{21}\) and \(t_{22}\)); if he notes the trailer out of place, he blocks the quay crane operator and starts \((p_2)\) to give directions to the trailer driver for adjusting trailer position \((t_3)\); if the trailer is correctly in place, \(Ch2\) begins the next step \((p_3)\).
When unloading containers, \textit{Ch2} verifies container external conditions (conflict between \texttt{t41} and \texttt{t42}). If damages are detected, \textit{Ch2} starts (\texttt{p4}) to signal this occurrence to \textit{Fo}, \textit{Ch1} and to the quay crane operator. After this event (\texttt{t5}), he is ready for another operation (\texttt{p6}). The crane operator moves the damaged container to a safe quay area, \textit{Ch1} updates container information in the database, and \textit{Fo} compiles a report. If no damage is detected, \textit{Ch2} is ready (\texttt{p5}) to check the type of container (conflict between \texttt{t61} and \texttt{t62}): if it is full, he starts (\texttt{p6}) reading and communicating the seal on it (\texttt{t7}); otherwise, if the container is empty, \textit{Ch2} starts (\texttt{p7}) to take locking tools off the container. After taking the lockers off and signalling the quay crane operator (\texttt{t8}), \textit{Ch2} begins (\texttt{p8}) to give the start-signal to the trailer driver (\texttt{t9}). After this communication, he starts (\texttt{p9}) verifying the end of the unloading process (conflict between \texttt{t10} and \texttt{t11}): if the end is reached, all containers have been loaded and \textit{Ch2} terminates his activity (\texttt{p11}); if the process is not finished, some containers are left and \textit{Ch2} waits (\texttt{p10}) for a new unloading task assigned by \textit{Fo} (\texttt{t1}).

Figure 2. GSPN module of the second checker activities
When loading containers, $Ch2$ verifies if the ship is equipped with automatic or semi-automatic lockers (conflict between $t_{43}$ and $t_{44}$): if not, he has to take and manually set lockers on containers and communicate with the crane operator ($t_{12}$); otherwise, if the containers are automatically locked, he is soon ready ($p_8$) to give the start-signal to the trailer driver and, thereafter, to check if the loading process is completed ($t_{11}$) or not ($t_{10}$) because of other loading tasks from $Fo$.

3. **Activities executed by quay crane operators, trailer drivers, and raisers**

   Considering the first quay crane operator, $Qo1$, we may synthetically describe his activities as follows. He receives from the first checker, $Ch1$, assignment to unload (load) a specific container from the ship bay (the trailer) to the waiting trailer (the ship bay). He moves the QC arm down to the bay (trailer). Then, he opens the arm spreader on the container and proceeds to hook it. At this time, after getting an ok-signal from his substitute on board $Qo2$ (the second checker $Ch2$), $Qo1$ lifts the container from the bay (trailer) and puts it on the trailer (in the bay). Then, he releases the container upon receipt of an ok-signal from $Ch2$ (from $Qo2$). Now, if the unloading (loading) process is terminated, $Qo1$ parks the crane and raises its arm. Otherwise, he continues by repositioning the crane on the ship bay (trailer).

   A trailer driver receives tasks from the yard planner. He drives his vehicle to the quay if going to take containers unloaded from the ship, the yard if going to pick-up containers from blocks. After the container is put on the trailer, the driver verifies the weight on a PDA: if something goes wrong, he notifies the yard planner; otherwise, he waits for the ok-signal to leave. Then, he delivers the unloaded (loaded) container to the yard crane (quay crane). After the crane pick-up, he waits for an ok-signal to leave for its next task in the quay or yard or for the parking area.

   A raiser gets assignments from $Fo$ to free containers to be unloaded or to fix containers being loaded. He gets on board and works in the bay indicated by $Fo$. He may deal with containers located in the ship cover or hold. After task completion, he communicates it to $Fo$ and gets down from the ship.

   The PN modules associated to the quay crane operators, trailer drivers, and raisers may be constructed by following the same formal procedure used to specify the formerly developed nets. But here, we omit them and refer the reader to future papers.

4. **Discussion about future developments**

   By integrating the model of the behaviour of each human operator and the model of the sequence of operations to be executed by each resource (crane, trailer, etc.), we will be able to complete an accurate model of the complex processes in the terminal. This formalization will be translated to a discrete event simulation environment, which can be useful to monitor and test the system, observe its performances, study more efficient methods of handling, stacking and transporting containers in the terminal area.
The aim is to achieve the highest efficiency of all operations handling containers, while reducing costs. Namely, this brings economical benefits both to the TCT company managing the terminal and to all navigation companies that use the terminal services and pay for them on a time basis.

Currently, the terminal stacking space is under-utilized on the basis of the daily demand, while material handling equipment is used at full capacity during transshipment cycles. In particular, trailers are sufficient to serve one vessel at a time only. So, to these authors, it seems evident that an efficient and effective management of unmanned and human resources would require different control policies of all scheduled operations, in case future customer demands will ask for more space and faster operations. Especially if one cannot invest much budget for acquiring new equipment to expand terminal capabilities.

References


Abstract. In this paper we deal with intermodal paths in urban transportation networks; in particular, we focus on the evaluation and comparison between the generalized cost of origin – destination paths requiring at least one change of modality and that associated with monomodal paths. With the aim of increasing the usage of intermodality, we analyse different pricing tools in such a way to minimize the negative components of the transition costs, that is the cost due to the change mode nodes. We give preliminary results related to the city of Genoa, Italy, that are obtained by analysing different scenarios via a spreadsheet tool.

1. Introduction and problem definition

In the last decades the mobility within urban areas is changed, and, in particular, the quota of users choosing urban paths that required more than one transportation modality is increasing [3]. Anyway, the private modality is the most preferred one and, in many cases, is still proportionally greater than the available infrastructures, thus determining congestion in urban areas.

Motivated by the above considerations, we decided to analyze the generalized cost of origin – destination (o-d) paths according to different modalities in order to evaluate the alternative paths and promote the intermodality. More precisely, given the generalized cost function for computing the shortest paths, related to both the mass transit transportation network and the private one, our main idea is to find the gain for which a user is willing to move from the private network to the public one; that is we would like to determine what could be for a driver the added value for reaching the destination on the public

---

1 This work has been developed within the MIUR PRIN 2005012452-003 project.
* DIEM - Università di Genova, Via Vivaldi, 5, 16126 Genoa, Italy
ambrosin@economia.unige.it; sciomach@economia.unige.it
CIELI – Centro Italiano di Eccellenza per la Logistica Integrata - Università di Genova
transportation network instead of continuing the desired path on the private network, i.e. by car. As a matter of fact, the selection of the path, either intermodal or monomodal, is performed directly by the user on the basis of the foreseen traveling cost, that includes some subjective elements that are not easy to be evaluated, such as social economic factor, propulsion at walking and discomfort. The shortest o-d path consists hence of a path and a modal combination from o to d such that users have the minimum generalized cost associated with it.

For the computation of the shortest o-d paths we use the algorithm proposed in [2]; other algorithms based on multimodal networks could also be used, such that the bicriterium shortest path algorithm presented in [7], the adaptive methods proposed in [5] and the algorithm for network with dynamic arc travel times [10]. Of course, in the computation of the shortest multimodal paths only multimodal viable paths are considered [4].

In the next Section we present our referring multimodal network model and its associated costs. In Section 3 the proposed approach for making as much as possible more convenient the multimodal paths than the private ones is given, together with some preliminary results related to the centre area of the city of Genoa, Italy.

2. The multimodal network model and the associated cost function

As in [6] we model an urban transportation network by a digraph $G = (V,E)$ that can be actually considered as the union of three subgraphs, representing, respectively, the considered transportation modalities, that is private, public and pedestrian, together with their interconnections. More formally, $G = (V,E)$ is such that $G = G_D \cup G_P \cup G_W$, where $G_D = (V_D,E_D)$ models the private network and $G_P = (V_P,E_P)$ is the public network; $V_D$ and $V_P$ are, respectively, the nodes reachable by car and by the mass transit transport. $G_W = (P,T)$ represents the pedestrian network, where $P$ is the set of available parking places (both on the street and on buildings), and $T$ is the set of transition arcs that represent the possibility of commuting between the private and the public modalities.

In order to better understand the transition cost of the paths in a multimodal network, let us consider the simple network represented in Figure 1, where nodes 1, 2, 3, and 4 $\in P$, that is are existing car parks, bold arcs belong to $G_D$, normal lines depict the arcs of $G_P$ and dotted lines represent the transition arcs of $T$ towards parking cars.

Let us assume that possible alternative paths connecting the origin node o to the destination node d are the following:

- a path through the private network, in which a driver can reach either node A or node E at the barrier where has to pay the corresponding fee for entering into the road pricing zone (depicted in Figure 1 as the oval zone), and finally park the car at any node between 2 or 3;

- an intermodal path, in which the driver after having reached the barrier can leave the car at a park, that is either node 1 or 4, and then continue the trip by taking a mean belonging to the public transportation network.

Let us now evaluate the cost of the above two paths. In the first case the driver has to travel sequentially for instance along arcs $(o,e)$ $(e,c)$ $(c,p)$ and $(p,d)$. The corresponding
costs are related to the travelling time $t'^p_v$, the road fee $\chi$, the fuel cost $c_c$ and the transition cost, that in this case is given by the parking fee $c_p^p$, the parking time $t'^p_p$ and the walking time $t'^p_c$ towards the destination node. In the second case, that is when choosing the intermodal path, the user has to travel sequentially along arcs (o,a) (a,i) (i,p) and (p,d). The corresponding cost are related to the usage of the car until the intermodal node, that is the travelling time $t'^I_v$, the fuel cost and the cost of the transition arcs, that is the parking fee $c_p^I$, the parking time $t'^I_p$, the walking time towards the bus stop and the waiting time $t'_w$ at the bus stop, the bus tickets $\beta$, the travelling time on the public mean and the walking time $t'^I_c$ towards the destination node.

\[ C'_{ij} = \omega_1(t'^I_v + t'^I_p + t'^I_c) + \omega_2(c'_p + \beta) \]  
\[ C^p_{ij} = \omega_1(t'^p_v + t'^p_p + t'^p_c) + \omega_2(c_c + c_p^p + \chi) \]

Of course, the paths would be identically preferable for the users if they have the same cost. Unfortunately, the cost given in (1) is almost always greater than that in (2).

**Figure 1. An example of path components in multimodal transportation network**

Note that for deriving the traveling times $t'^p_v$ and $t'^I_v$ any of the usual cost functions that can be found in the transportation literature can be used (see for instance [8] and [9]), provided that they are increasing function with respect to the flow on the corresponding arcs.

Let us assume that the generalized cost of the considered paths is hence obtained by introducing weights $\omega_1$ and $\omega_2$, that correspond to the perception of the users of the time and the monetary cost.

It can be easily observed that the generalized cost of any intermodal path is then given by equation (1), while the generalized cost of any private path is then given by equation (2).
A crucial role in the evaluation of the optimal trade-off between benefits and costs in the choice of the path is paid by the modal change nodes. In this paper we start from the analysis of the most “relevant” nodes among the intermodal ones that has been proposed in [1] as follows:

**Definition 1.** We say that a node \( i \in V \) is a primary intermodal node if a) \( i \in V_D \cap V_P \), b) \( i \) is connected in \( G_P \) to at least 60% of the nodes belonging to \( V \), and c) \( i \) is reachable by all nodes in \( V \) on \( G_D \) along a path whose cost is at least 20% less than the average shortest path cost.

**Definition 2.** We say that a node \( i \in V \) is a secondary intermodal node if a) \( i \in V_D \cap V_P \), b) either condition b) or c) of Definition 1 holds c) either condition b) or c) of Definition 1 is satisfied if it is relaxed of about 20%.

In fact, we think that for promoting the intermodality it is crucial to improve the attractivity of the mode changes nodes that already perform very well from a structural point of view, by also locating there services like car parks, info panels and so on.

Our aim is find instruments for increasing the usage of intermodality. In the next Section we analyse some pricing tools that could have a different impact on the cost function and hence make more favourable intermodal paths.

### 3. Making more favourable the intermodal paths

Let us consider the city of Genoa, Italy, and some main traffic paths from east to west side and viceversa that are reported in Figure 2. Nodes 1, 2 and 10 are the main congested ones and represent the zones in which we split the central area of Genoa and focus our analysis, since they are considered primary intermodal nodes. Note that in this area the public network consists on both bus and train connections.

The average transit flow and the modal split in a rush hour time (i.e. from 7.0 a.m. to 9.0 a.m. and from 5.30 p.m to 7.30 p.m.) throughout these nodes is reported in Table 1.

<table>
<thead>
<tr>
<th>Origin</th>
<th>Destination</th>
</tr>
</thead>
<tbody>
<tr>
<td>Node 1</td>
<td>Node 2</td>
</tr>
<tr>
<td>Car</td>
<td>14150</td>
</tr>
<tr>
<td>Bus</td>
<td>600</td>
</tr>
<tr>
<td>Train</td>
<td>2200</td>
</tr>
</tbody>
</table>

**Table 1. Flow of users at the main nodes**

For computing the shortest paths among the above nodes we have considered the average distance from any reachable intermodal nodes, either primary or secondary ones, to parking places in terms of driving time from \( G_D \) and the successive connection to \( G_P \), that includes the waiting time for a public mean. We derived consequently the corresponding generalized cost that is computed by assuming the time conversion factor of 15 Euro / hour. Note that for each node we have the same monetary cost both for the parking tariff, that is 1,5 Euro / hour and the bus ticket, that is 1 Euro for 90 minutes of traveling time.
Figure 2. Connections among the three main nodes

We used the Dijkstra algorithm for determining the shortest paths in the public and private network for each of the o-d selected paths and successively determined the shortest multimodal path by using the algorithm presented in [2].

By analysing the cost components of the above paths according to equations (1) and (2) we have the values reported in Table 2.

By looking at Table 2 the reader can see that none intermodal or public path is more convenient than the private path.

Note that independently on the value reported in Table 2, given the above assumptions about \( c^I_p \), \( c^P_p \) and \( \beta \) we have, under the equality condition (i.e. (1) = (2)), that

\[
\omega_1 = (\chi - 1) \text{ or } \omega_1 = (\chi - 2)
\]

(3)

depending on the travel time along \( G_p \).

Therefore, we face the problem of fixing the pricing tools as road fee, parking fee and bus ticket, via a simulation performed by using the spreadsheet Excel in order to have a given number of intermodal paths preferred to the private ones.

In particular, we investigate different scenarios that reflect different government policies. In order to promote intermodal paths, the decider can act in two opposite way: he/she can decide to reduce the cost of intermodal paths by reducing both the parking fee and the bus ticket; on the other hand, he/she can increase the cost of the private paths by introducing road fees and increasing the cost of the car parks in the centre areas.
The results obtained by using these different policies depend on many factors, among others the perception of the services offered by the public transport, the cost of the travel time and the cost of the transaction arcs \( T \) at the modal change nodes.

For these reasons we try to describe the effect of the different policies by also considering different weight in equations (1) and (2). The results are sensitized in six different scenarios reported in Table 3.

For instance the case of presence of interchange car parks that are characterized by \( c_{i}^{p} = 0 \) and \( r_{i}^{p} = 0 \), together with car parks in the centre areas with a fee of 2 Euros represent a policy in favor of intermodal paths: in such cases the number of the preferred intermodal paths is 10 over 16 and 9 over 16 with and without the road fee (see columns 5 and 6). Moreover, if we differentiate also the social classes such that the time conversion factor is 15 Euro/h as before for the middle high class people and 7.5 Euro/h for the middle low class, we obtain that the number of the preferred intermodal paths is 9 and 16 (see last row of columns 5 and 6).

Note that an high parking cost in the centre areas increases the number of users that choose intermodal paths, but on the other hand the waiting time for an available place in the car park results lower; we can hence observe that for this indirect effect on the total cost of the private path, the number of preferred intermodal paths decreases for the high level social class (see columns 8).

<table>
<thead>
<tr>
<th></th>
<th>( T_{v} P )</th>
<th>( T_{v} I )</th>
<th>( T_{r} P )</th>
<th>( T_{c} P )</th>
<th>( T_{r} I )</th>
<th>( T_{c} I )</th>
<th>( (1) )</th>
<th>( (2) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-2</td>
<td>7.06</td>
<td>3.11</td>
<td>2.887</td>
<td>7.689</td>
<td>1.4</td>
<td>4.667</td>
<td>6.958</td>
<td>4.426</td>
</tr>
<tr>
<td>1-4</td>
<td>7.35</td>
<td>3.11</td>
<td>3.45</td>
<td>8</td>
<td>1.4</td>
<td>4.667</td>
<td>7.165</td>
<td>4.846</td>
</tr>
<tr>
<td>1-9</td>
<td>9.62</td>
<td>3.11</td>
<td>1.15</td>
<td>1.15</td>
<td>1.4</td>
<td>4.667</td>
<td>6.095</td>
<td>5.306</td>
</tr>
<tr>
<td>1-8</td>
<td>10.45</td>
<td>6.27</td>
<td>3.4</td>
<td>6.833</td>
<td>1.2</td>
<td>3.333</td>
<td>7.594</td>
<td>5.457</td>
</tr>
<tr>
<td>1-5</td>
<td>7.72</td>
<td>6.27</td>
<td>3.12</td>
<td>5.8</td>
<td>1.2</td>
<td>3.333</td>
<td>7.327</td>
<td>5.867</td>
</tr>
<tr>
<td>1-7</td>
<td>8.74</td>
<td>3.11</td>
<td>3.45</td>
<td>8</td>
<td>1.4</td>
<td>4.667</td>
<td>7.871</td>
<td>4.826</td>
</tr>
<tr>
<td>1-3</td>
<td>7.86</td>
<td>3.11</td>
<td>3.1</td>
<td>5.667</td>
<td>1.4</td>
<td>4.667</td>
<td>6.828</td>
<td>4.406</td>
</tr>
<tr>
<td>6-2</td>
<td>3.91</td>
<td>0</td>
<td>2.887</td>
<td>7.689</td>
<td>1.39</td>
<td>4.6</td>
<td>6.381</td>
<td>3.825</td>
</tr>
<tr>
<td>6-4</td>
<td>3.58</td>
<td>0</td>
<td>3.45</td>
<td>8</td>
<td>1.39</td>
<td>4.6</td>
<td>6.455</td>
<td>4.382</td>
</tr>
<tr>
<td>6-5</td>
<td>3.63</td>
<td>0.97</td>
<td>3.4</td>
<td>7.667</td>
<td>3.4</td>
<td>6.833</td>
<td>6.965</td>
<td>5.976</td>
</tr>
<tr>
<td>6-7</td>
<td>3.58</td>
<td>0.97</td>
<td>3.45</td>
<td>8</td>
<td>3.4</td>
<td>6.833</td>
<td>6.789</td>
<td>4.557</td>
</tr>
<tr>
<td>6-3</td>
<td>3.59</td>
<td>0.97</td>
<td>2.4</td>
<td>6.167</td>
<td>3.4</td>
<td>6.833</td>
<td>5.971</td>
<td>4.857</td>
</tr>
<tr>
<td>10-4</td>
<td>2.81</td>
<td>0</td>
<td>3.45</td>
<td>8</td>
<td>3</td>
<td>5</td>
<td>6.286</td>
<td>5.255</td>
</tr>
<tr>
<td>10-5</td>
<td>3.54</td>
<td>0.88</td>
<td>3.4</td>
<td>7.667</td>
<td>3.4</td>
<td>6.833</td>
<td>6.745</td>
<td>5.961</td>
</tr>
<tr>
<td>10-7</td>
<td>2.81</td>
<td>0</td>
<td>3.45</td>
<td>8</td>
<td>3</td>
<td>5</td>
<td>6.62</td>
<td>4.652</td>
</tr>
<tr>
<td>10-3</td>
<td>3.63</td>
<td>0.88</td>
<td>2.15</td>
<td>4.5</td>
<td>3.4</td>
<td>6.833</td>
<td>5.638</td>
<td>4.796</td>
</tr>
</tbody>
</table>

**Table 2. Value of equations (1) and (2) for the main paths**

The results obtained by using these different policies depend on many factors, among others the perception of the services offered by the public transport, the cost of the travel time and the cost of the transaction arcs \( T \) at the modal change nodes.

For these reasons we try to describe the effect of the different policies by also considering different weight in equations (1) and (2). The results are sensitized in six different scenarios reported in Table 3.

For instance the case of presence of interchange car parks that are characterized by \( c_{i}^{p} = 0 \) and \( r_{i}^{p} = 0 \), together with car parks in the centre areas with a fee of 2 Euros represent a policy in favor of intermodal paths: in such cases the number of the preferred intermodal paths is 10 over 16 and 9 over 16 with and without the road fee (see columns 5 and 6). Moreover, if we differentiate also the social classes such that the time conversion factor is 15 Euro/h as before for the middle high class people and 7.5 Euro/h for the middle low class, we obtain that the number of the preferred intermodal paths is 9 and 16 (see last row of columns 5 and 6).

Note that an high parking cost in the centre areas increases the number of users that choose intermodal paths, but on the other hand the waiting time for an available place in the car park results lower; we can hence observe that for this indirect effect on the total cost of the private path, the number of preferred intermodal paths decreases for the high level social class (see columns 8).
Finally, in an unrealistic scenario in which it is possible to travel on bus without paying a ticket and it is necessary to pay only for car park in the center areas, the number of intermodal path preferred is 9 (see column 3).

As further analysis, in Table 4 we show the number of preferred intermodal paths when varying $\varpi_1$ and $\varpi_2$ in equation (3), such that $\varpi_1 + \varpi_2 = 1$. By choosing $\varpi_1$ closest to one we consider very important the cost factor of the generalized cost of each path. Otherwise by choosing $\varpi_1$ closest to zero the time factor is considered to be more relevant. Note that by considering only the time factor the number of the preferred intermodal paths is 8.

In Table 4 we report the experiments in which we fixed $\varpi_1$ equal to 0.2, 0.5 and 0.8.

<table>
<thead>
<tr>
<th>Scenario</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta$</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$\text{Cp}^1$</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>$\text{Cp}^2$</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td># intermodal paths preferred</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15 EURO</td>
<td>0</td>
<td>3</td>
<td>9</td>
<td>3</td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>4</td>
<td>9</td>
</tr>
<tr>
<td>7.5 EURO</td>
<td>0</td>
<td>1</td>
<td>9</td>
<td>1</td>
<td>9</td>
<td>16</td>
<td>9</td>
<td>5</td>
<td>15</td>
</tr>
</tbody>
</table>

Table 3: results of different scenarios

Table 4: preferred intermodal paths for two social classes and different weights $\varpi_1$

Generally the number of preferred intermodal paths is higher for people belonging to the low social class, while we have noted the same proportional variation for the users of the two considered social classes when we modify the weight $\varpi_1$ and $\varpi_2$.

Since the greatest impact on the preference of the users expressed in terms of number of preferred intermodal paths is due to a reduction of the cost of transaction arcs at the modal change nodes, it is worst mentioning that the choice of the modal change node among the primary ones for improving the services for users is a strategic factor. Only after having adopted new strategies on primary nodes (such as location of info panels, new services and so on) the decider can also act on pricing tools.
References


Abstract. Urban mobility is characterized by a relevant interaction between travel demand and transportation supply. Investments in mass transportation move mobility demand from private to public transportation, but new demand arises as a consequence of the social growth. As a result the global transportation supply does not satisfy all the mobility demand, producing traffic congestion. Users Route Guidance is needed in the stationary state of the network and in emergency situations. The first case concerns an average working day with the expected flow pattern. The second case concerns a network with an event requiring the forced flow re-routing or the evacuation of the zone, in case of very a dangerous situation. Therefore it is worthwhile to introduce new relevant optimisation problems in the urban traffic system design, such as infomobility service facility location (sensors and cameras, VMS) [4] [8], and o/d path computation optimising a particular performance index of the network. The job program of the research unit is based on the development of two innovative themes. The first is made up of models and methods for finding the optimal location of plants and services for infomobility. The second is made up of models and methods for finding the optimal paths to choice to minimize traffic congestion on the network or the evacuation time.

1. Introduction

Urban mobility is characterized by a relevant interaction between travel demand and transportation supply. Investments in mass transportation move mobility demand from private to public transportation, but new demand arises as a consequence of the social growth. As a result the global transportation supply does not satisfy all the mobility demand, producing traffic congestion. In this context, it is important to guide users of
private transportation on paths devoted to increasing park and rides, to reduce conflict
nodes and traffic pollution.

In the last decade, route guidance systems of traffic network users have been introduced to
obtain a best flow pattern and reduce traffic congestion. The term Route Guidance has been
introduced to define a system that gives drivers useful information useful for choosing the
best o-d path. A Route Guidance system can be static and dynamic. The former uses
historical data about flows and travel times on the links, while the dynamic system uses
information about the real state of the network.

To this end, telematic technologies are needed to develop monitoring systems (sensors,
cameras), capable of detecting the flows on some road links and information systems
(variable message systems) in real time, to inform the users about the network state in terms
of congestion levels, opening/closure of roads and convenient routes.

The development of these technologies is changing the management of urban traffic, from a
classical static approach to an innovative dynamic one, and new approaches to the urban
traffic management have been developed, based on ATIS (Advanced Traveller Information
System) and/or on ATMS (Advanced Traffic Management System).

Users Route Guidance is needed in the stationary state of the network and in emergency
situations. The first case refers to an average working day with the expected flow pattern.
The second case refers to a network with an event that requires forced flow re-routing or
the evacuation of the zone, in the case of a very dangerous situation.

Therefore it is worthwhile to introduce new relevant optimisation problems in the urban
traffic system design, such as infomobility service facility location (sensors and cameras,
VMS)[4] [8], and o/d path computation optimising a particular performance index of the
network.

The job program of the research unit is based on the development of two innovative
themes. The first is made up of models and methods for finding the optimal location of
plants and services for infomobility. The second is made up of models and methods for
finding the optimal paths to choose to minimize traffic congestion on the network or the
evacuation time.

For the first theme, there are contributions to literature devoted to the formulation of
models for the location of discretionary services, i.e. of services that do not have
necessarily to be used by the customers of the network during their journey. Variable
message systems (discretionary services of an involuntary type) belong to this category of
services. In general we can talk about problems of location of systems and services that
intercept flow, which can be defined as path covering problems [2]. There are not many
contributions in literature to this topic. Network location problems have up to now
concerned systems and services, which generate and attract flow.

The models proposed in literature for flow intercepting facility location are based on the
hypothesis that the network flow pattern is assigned, in terms of link flows, and in terms of
path flows, and moreover that it is not modified by the plant location. The models are
therefore articulated on the basis of the information available on flows:

- flows on arcs,
- flows on the distances.

We found models of binary integer programming where the variables are associated to the
plants to locate (in vertices or on links) and to the paths intercepted from the same plants.
The objective function aims to:

- maximize the flow intercepted from a prefixed number of plants;
minimize the number of plants necessary in order to intercept the whole flow or an assigned quota of the total flow.

The plant location can be made in vertices or nodes. The choice to localize on vertices is mainly used in the case of localization of sensors and cameras, while the choice to locate on links is mainly used in the case of the variable message panels.

The research theme is therefore particularly interesting and innovative, from the application point of view, because it can be inserted in the context of urban traffic management and from the methodological point of view, because it addresses a little-explored field of network location.

The second research topic concerns the modelling of particular “extraordinary” situations, in which users are constrained to follow a predetermined path and, consequently, the assignment of users to paths on a traffic network can be preliminary performed. This phenomenon occurs in evacuation problems that represent one of the aspects of emergency processes, which can be simply defined as the removal of residents of a given area considered dangerous to safety zones as quickly as possible and with utmost reliability.

The complexity of an evacuation process depends on various factors including, for instance, the extension of the involved area, the average distance and the average time evacuees need to move toward the safety area (known as egress time), the time needed to recognize danger and to decide which course of action to take and the total time needed to complete an evacuation process.

The overall organization of an evacuation process is defined in an evacuation plan which includes objectives, procedures, activities, resources (economic, infrastructure and human resources), times and methods to respect. In this context planning and operational organization of the flow distribution on the network so that users can easily and quickly reach the safety zones are crucial. In order to perform these tasks mathematical programming models able to describe either planning phases or management activities are available.

In the planning phase models provide solutions in terms of network design variables to be introduced with the objective of improving network performances in terms of evacuation time and congestion risk.

In the operational phase, on the basis of a given network as defined during the planning phase, optimisation models determine the set of paths and the starting time that users leaving each origin node should follow in order to minimize overall measures of performance.

Mathematical models can be effectively used in practical applications if they are able to represent an evacuation process for a given study area in a realistic way. In addition they should be effectively included in flexible decision systems providing solutions, which can support the decision maker in the definition of the evacuation plan to be adopted.

Existing evacuation approaches can solve problems of limited dimensions and are therefore currently are not suitable for tackling realistic applications.

The objective of the proposed research project is primarily focused on the improvement of existing models, solution techniques, which could represent the framework for the design of a decision support system. In particular the aim is to implement a software tool whose characteristics, (i.e. flexibility, portability, integrability) fosters practical applications.
2. Research Program

The work program of the research unit is based on the development of two innovative themes:
- Infomobility plant location
- Path finding for congestion level reduction and/or for zone evacuation.

2.1. Infomobility Plant Location

The available models are quite general and so it is necessary to formulate specific constraints for the applied problems described. In particular for localization of the sensors and cameras for flow monitoring we must insert the model in the context of the estimation process of the origin/destination travel demand matrix. It is clear that the systems for the monitoring flows have to be located to maximize the likelihood of the estimated O/D matrix, compared to the real O/D matrix [15] [9]. Particular attention is therefore necessary for the problem of multiple counting, in order to avoid the same flow being counted several times in the course of its movement from the origin to the destination.

Instead, for the location of the variable message panel we must adapt the formulation of the model to the specific role that this type of system plays in the context of operation of a network with monitoring and information of the customers. In this case, as an example, it could be useful for a single flow to intercept a flow two or more times, i.e. in other words, that it is possible to inform customers two or more times of a path constituting an O/D flow. In any case the binary integer programming models formulated require elevated computation times for the optimal solution, because of the NP character of the problem to solve. In the course of the first experiments carried out with a commercial solver it was possible to solve small and medium networks in low computation times. On one hand it is therefore necessary to estimate the possibility of a polyhedral approach that concurs to improve the formulation of the problem and, on the other, to construct and use also inexact and heuristic algorithms for the solution of the proposed problems [14]. The heuristic algorithms available in literature for the solution of problems of path covering are based in general on the coverage matrix of the graph, i.e. in other words, the incidence matrix link/path. They produce feasible solutions in very low computation times that in some cases can be far from the optimal solution of the problem. It is therefore necessary to study the behavior of available heuristics and to predispose modifications to reduce the distance from the optimal solution [15].

The research team has developed models and methods devoted to location of fixed VMS plants for user information in the previous project. Two models were formulated: the first, named M1, was based on the maximization of the intercepted flow, with an assigned number of plants. The second, M2, was based on the minimization of the number of plants needed to intercept a fixed percent of the total flow. The research program 2003 contains the experimentation of these two models on a real case of the traffic network in Naples.
The research program 2005 contains the development of models and methods devoted to location of mobile units, according to travel demand and flow pattern evolution.

The whole day is then subdivided into periods and each origin/destination demand is assigned to paths from origin to destination. This information can be determined starting from the flows on the links, detected on-line, and by an o/d matrix estimation. Through the calculation of the flow distribution on the network it is possible to determine the flows on the paths between each o/d pair. Thanks to this information, optimal location of the available mobile units can be performed using the previously formulated and tested M1 and/or M2 models. The acquisition (on-line or off-line) of the same information for the next period allows us to define new locations for the available systems. For this reason it is necessary to formulate a suitable model for the repositioning of the mobile units on the base of the link costs, with the objective of minimizing the total cost of repositioning units, under further constraints on the maximum distance that each unit can cover in order to reach the new assigned location. It is evident that the two objectives (maximization of the intercepted flows and minimization of the repositioning costs) can be in conflict. For this reason iterative and/or dynamic procedures could be proposed with mono-objective or bi-objective functions to optimize.

2.2. Path Finding for Evacuation Problems

The second research topic regards traffic assignment in emergency situations. The topic was developed within the PRIN project on Infomobility carried out by the research team of the Dept. of Management Engineering of the “Federico II” University of Naples. The main result of this project was the proposal of a prototype of a decision support system for the planning and the management of evacuation processes based on the adaptation and the modification of mathematical models for network flow optimization.

The decision support system consists of two fundamental elements: a procedure for the definition of paths on a network which users should follow during an evacuation process; the calculation of the dynamic scheduling of the flows along the individuated paths; a user friendly interface for data entry and the representation of the provided results.

The procedure is represented by a classical local search heuristic, which uses two fundamental models (Quickest Path Problem, Minimum Turnstile Cost [1]). The proposed procedure solves the evacuation problem from many origins to many destinations on a network whose links are characterized by a travel time and a capacity on the link. In particular, given a time horizon subdivided into a set of intervals of given lengths, the procedure provides the dynamic evolution of the flows on the network during the entire evacuation time.

The computational times of the procedure are strongly influenced by the number of intervals and, as a result, by the length of each interval.

The user interface allows one to perform data entry (network, parameters associated to the network, population data) and provides a visual output and a final report containing all the information needed to practically implement the evacuation process. The prototype interface also includes tools for the import of maps, which make input and output operations easier to realize.
Starting from the availability of this prototype, the research project aims to design and implement a decision support system able to solve the problem of the planning and/or the management of real practical applications.

In particular the research project will focus on the following fundamental steps:

1. Definition and implementation of a user interface compatible with a geographical information system in order to exploit the functions of these systems to better perform input and output procedures.

2. Improvement of the proposed heuristic with particular attention to the quality of the solutions provided and the computational efforts required. With this aim the current heuristic based on a local search procedure could be modified by the definition of more sophisticated neighborhoods and/or the development of metaheuristics. In order to perform this step computational experiences to identify the more suitable characteristics of the procedure are required.

3. Identification of flexibility requirements to include in the system; this objective will be achieved through a preliminary analysis of characteristics that a system should present in order to better perform planning and/or management operations during an evacuation process.

4. Application of the proposed system to solve a practical case study. This phase will be realized thanks to the cooperation and collaboration of local authorities and institutions.

2.3. Conclusions

The cultural background of the research team is the result of various research projects that the team had developed together with the municipality of Naples (ATENA and Centaur, sponsored by the MURST). The research team has also been involved, together with other research groups, in the preliminary analysis of the development of a coordination plan for the evacuation of the so-called “red zone” in proximity of the Vesuvius volcano. The theme of the research unit is fully placed in the context of an integrated system of traffic management. Therefore it strongly interacts with the work of all the other participants in the co-financing proposal, and in particular with the units in Rome and Genoa. To this end, it will be useful to plan an exchange of data and integration of procedures for the construction of a procedure for the integrated management of the urban traffic system. Models and methods have to be tested and experienced on real network in the cities of Naples and Genoa.
References


INFOMOBILITY AND LOGISTICS ON URBAN AND REGIONAL NETWORKS

Alberto COLORNI1, Roberto WOLFLER CALVO2, Renato DE LEONE3, Paolo DELL’OLMO4, Nicoletta RICCIARDI4, Daniela AMBROSINO4, Anna SCIOMACHEN5, Giuseppe BRUNO6, Gennaro IMPROTA6, Antonio SFORZA7

Abstract. The paper deals with a national research project supported by MIUR (Italian Ministry of Instruction, University and Research). The local research units are at the universities of Camerino, Genoa, Naples, Milan and Rome and are coordinated by the University of Rome. The project regards infomobility and the design of an integrated urban transportation system, with specific reference to urban networks adopting telematic technologies and innovative transport services. The research topics of the five units include:

- Optimisation models that take advantage of telematic technologies to make more efficient and integrated public transport services, intermodal exchange nodes and parking areas for private-public connections.

1 Dipartimento INDACO (Industrial Design, Arts & Communications), “Politecnico di Milano” Technical University, via Durando 38/A - 20158 Milano - Italy, alberto.colorni@polimi.it
2 Departement Genie des Systemes Industriels, Universite de Technologie de Troyes, 12, rue Marie Curie - BP2060 - 10010 Troyes Cedex - France, roberto.wolfler@utt.fr
3 Dipartimento di Matematica e Informatica, University of Camerino, via Madonna delle Carceri 9 - 62032 Camerino (MC) - Italy, renato.deleone@unicam.it
4 Dipartimento di Statistica, Probabilità e Statistiche Applicate, University of Rome, P.le Aldo Moro 5 - 00185 Roma - Italy, {paolo.dellolmo,nicoletta.ricciardi}@uniroma1.it
5 Dipartimento di Economia e Metodi Quantitativi, University of Genoa, via Vivaldi 5 - 16126 Genova - Italy, {sciomach,ambrosini}@economia.unige.it
6 Dipartimento di Ingegneria Economico-Gestionale, University of Naples “Federico II”, via Claudio 21 - 80125 Napoli - Italy, {giubruno,improta}@unina.it
7 Dipartimento di Informatica e Sistemistica, University of Naples “Federico II”, via Claudio 21 - 80125 Napoli - Italy, sforza@unina.it
1. Introduction

The effects of congestion on urban networks are well known, as are the consequent social and environmental costs. More recently, travellers and drivers are becoming more familiar with the new capabilities for reducing such phenomena or for improving public transport, which are based on telematic services that both detect traffic flows in real time and transmit and visualise information for users (e.g. Variable Message Signs, information panel at bus stops, etc.). Moreover, travellers are starting to adopt innovative transport services like car pooling, car sharing and Dial-a-Ride, integrated, in several cases with mobile phone services (e.g. text messaging) or the Internet, to improve the coordination between demand and supply. Consequently, the attitude of local authorities is also changing. They are starting to look more favourably at the potential offered by traffic control systems and infomobility systems, not only for everyday management but also for large events (emergency plans, strikes, big celebrations, etc.) which, as a matter of fact, are relatively frequent and indeed very important for security and public order. At the same time, there is a need for new decision models specialized enough to capture the specific characteristics of the different problems, yet easy enough to be integrated in a common framework, to allow coordinated management of intervention in the same area where different flows share the same resources.

The project aims to extend a previous project (MIUR n° 2003095533 2003-2004): "Infomobility and the design of an integrated urban transportation system". This project, which started in November 2003 and lasted 24 months, studied the basic component of a transport system for planning and managing a traffic network, with specific reference to urban networks adopting telematic technologies and innovative transport services.

The extension mainly regards problems not contained or only outlined in the previous project, but that have been clearly identified during research activity. Among which are:

- New optimization models that take advantage of telematic technologies to create more efficient and integrated public transport services, intermodal exchange nodes and parking areas for private-public connections.
- New models and algorithms for locating facilities for flow interception (i.e. Detectors) or to visualize information for users (i.e. VMS).
- New models for the management of evacuation processes and large urban events.
- New algorithms for the optimal management of innovative transportation services.
- Evaluation of the effects of freight transportation on the urban mobility and design of new models of city logistics.
- Developing model integration on a single platform and automating the decision-evaluation cycle.

The methodological base and the global approach to the problem involve a unique viewpoint in dealing with the above-mentioned topics by means of adequate mathematical formulations and the corresponding solution algorithms. The validation and experimentation phase will also be carried out with the support of local authorities.

In the following proposal the problems we would like to address are described in detail, together with the possibilities of formulating them, the solution approach we want to follow and the way we want to perform experimental tests, both in the laboratory and in the field.

2. Objectives of the research program

The growth in the demand for transport that we have been seen in recent years, especially in urban areas, depends on several economic, demographic and social factors. The costs and time required to realize new infrastructures and existing spatial and environmental constraints have not permitted an adequate growth in the supply of transportation. This is especially true in Italy, where major cities are subject to several constraints of a different nature. As a consequence, our transport networks are characterized by a high level of congestion, with well known social and environmental costs.

The increase in urban congestion and the related rise of direct and indirect costs for single users and the community makes traffic management one of the most urgent and important problems to be solved in urban areas.

Only mobility issues devoted to the movement of people have been considered, while freight movement has been quite neglected in this context. However, movement of goods represents a significant problem for urban areas.

The phenomenon is not new, but in the past it has been considered of secondary importance in the planning and analysis of transport systems in urban areas. Today, attention toward this topic is rapidly growing because of increasing awareness on the part of institutions and citizens of the impact of urban mobility on congestion, noise, pollution and so on.

The recent diffusion of telematic technologies, both for real time monitoring of network conditions and for giving information to users, are providing an opportunity to optimally use the capacity of existing infrastructures, to manage flows according to variation in demand and to coordinate services for mobility and freight transportation.

It is quite clear that there is a need for decisional models capable of making the best use of these new sources of information. Indeed, while from a technological point of view the required equipment is becoming more reliable and affordable, the whole process of acquisition, filtering, analysis and organization of data to obtain useful information for traffic management and
planning could be widely improved. This is becoming clearer as soon as the different optimization models are introduced into the system and determine precise information needs. Hence, we are in the phase where "intelligent" components induce guidelines for integration and analysis to be performed on the raw data, defining the requirements of the whole process and the feasible strategies of high level management of flows. While, on one hand, the availability of information makes it possible (at least in theory) to identify multiple strategies of intervention, on the other the complexity of the problem calls for a new generation of decision models and algorithms with the following characteristics: i) use of (real time) information; ii) models integrated in a common architecture; iii) efficient algorithms with an acceptable running time despite the computational complexity of the problem.

In this context, the project participants are cooperating and unifying their experiences. We believe that this effort will contribute to identifying new methodological results in an integrated approach and to verifying them. Our objectives are:

i) To develop new methodologies required for solving the above problems.
ii) To develop and experimentally evaluate new models and algorithms for the optimal planning and for the management of urban mobility services and freight distribution in the urban area;
iii) To educate specialised technicians for implementing advanced tools and adapt them to the specific needs of the users.

We want to deal with four topics in this project:

i) models for routing of flows both in standard and emergency conditions
ii) models and algorithms of network location of facilities and services both for flow interception and traveller information
iii) optimization of transport services’ performance (routing and scheduling) and of their integration: multimodal, public and private transportation and innovative transportation services
iv) freight transportation and urban mobility: city logistic models for sustainable mobility

In the proposal we describe the problems to be faced in detail, the possibility of formalising them in an adequate manner, the solution approach we want to follow, and experimental activity both in the laboratory and in practice.

3. Main research fields

The scientific background of the project consists of international and national research papers and books (the authors of many of them are participants of the present project) and of results of many national and international research projects to which many members of this project have belonged. In particular, DREAMS (Demand Responsive Extended Area Mobility Services; City
of Milan), ATENA (Environment, Traffic and Telematics in Naples) financed by MURST and the European Community, European research projects (AIUTO, TransPrice, Reflex) and national projects (Progetto Finalizzato Trasporti 2) led the operative units of this project not only to develop extensive skills and know-how in the corresponding fields but also to establish collaborative links with many municipalities. The main research fields of the present proposal are described below.

3.1. **Traffic flow management, traffic plans for emergency, large events and evacuation processes**

Phenomena that disrupt the equilibrium of network conditions (interruption of one or more arteries, high transport demand, emergency situations, etc) occur very often in Italian cities. These require the design of more robust and reliable networks and methodologies to re-route flows, which can be implemented quickly. Unfortunately, existing models in literature are based on simplifying assumptions that make their solution quite unrealistic in the above scenarios. As far as re-routing is concerned, recent papers have proposed local re-routing strategies that fail to find paths and ensure the progression of flows to their final destination. This appears to be a major drawback especially for evacuation plans (as in the case of the risk of Vesuvius for the city of Naples) and large events (as in the case of Rome).

Part of the research topics concern modelling particular "extraordinary" situations, in which users are constrained to follow a predetermined path and, consequently, where the assignment of users to paths on a traffic network can be performed beforehand. This phenomenon occurs in evacuation problems that represent one of the aspects of emergency processes, which can be simply defined as the removal of residents of a given area considered dangerous to safety zones as quickly as possible and with utmost reliability.

The complexity of an evacuation process depends on various factors including, for instance, the size of the area involved, the average distances and the average time need to move toward the safety area (known as egress time) the time needed to recognize a dangerous event and to decide which course of action to take and the total time needed to complete an evacuation process.

The overall organization of an evacuation process is defined in an evacuation plan which includes objectives, procedures, activities, resources (economic, infrastructure and human resources) and times and methods to respect. In this context, planning and operational organization of the flow distribution on the network so that users can easily and quickly reach the safety zones are crucial. In order to perform these tasks, mathematical programming models able to describe either planning phases or management activities are available [19].

In the planning phase, models provide solutions in terms of network design variables to be introduced with the objective of improving network performances in terms of evacuation time and congestion risk. In the operational phase, on the basis of a given network as defined during the planning phase, optimisation models determine the set of paths and the starting time that
users leaving each origin node should follow in order to minimize overall measures of performances [48][50].

Mathematical models can be effectively used in practical applications if they are able to realistically represent the evacuation process for a given study area. In addition, they should be effectively included in flexible decision systems providing solutions, which can support the decision maker in the definition of the evacuation plan to be adopted. Existing evacuation approaches can solve problems of limited dimensions and so are not currently suitable for tackling realistic applications.

The objective of the proposed research project is primarily focused on the improvement of existing models and solution techniques, which could represent the framework for the design of a decision support system. In particular the aim is to implement a software tool whose characteristics, (i.e. flexibility, portability, integrability) allows practical applications.

3.2. Infomobility, traffic flows and information services for users.

Most of our cities are equipped with systems to inform users (Variable Message Signs, intelligent information panels at bus stops, etc.). These technologies (ATIS - Advanced Travellers Information Systems) are used to inform users on the congestion status of the network and should permit:

1. users to be reached while they are planning their routes (ENVG- En-Route Vehicle Guidance) so as to avoid congested links at the moment of route selection; the route selection can be static (based on historical data) or dynamic (using real time information);

2. users to be informed during their trip [VMS-Variable Message Signal] so as to help changes of routes in order to avoid congestion (queue length in particular intersections, arteria status classification, alternative route suggestions, occupancy of parking areas, etc).

It is also necessary to introduce two important optimization problems to the design of the traffic system. The first is made up of models and methods for optimally placing infomobility systems and services. The second is made up of models and methods for finding the optimal paths users should choose to minimize traffic congestion or evacuation time.

There are contributions to this theme in literature devoted to the formulation of models for location of discretionary services, i.e. of services that do not necessarily have to be used by the customers of the network during their journey. Variable message systems belong to this category of services. The models proposed in literature for flow intercepting facility location are based on the hypothesis that the network flow pattern is assigned, in terms of link flows, and in terms of path flows, and, moreover, that it is not modified by the plant location. The models are therefore defined according to the information available on arc flows and path
flows. These are typically binary integer programming models where the variables are associated to the plants to be located (in vertices or on links) and to the paths intercepted from the same plants. In these cases, the objective function is aimed at maximising the flow intercepted by a fixed set of plants or at minimising the number of plants needed in order to intercept the whole flow or a part of it.

3.3. Multimodal transportation and innovative transportation services

Intermodal transportation is becoming more and more important for the mobility of people. For this reason, in recent years several works both on methodological and practical aspects have appeared in literature.

In particular, several studies on urban intermodal networks have been made. Algorithms for the determination of shortest paths in urban multimodal networks and multi-objective algorithms based on the preferences of users expressed by utility functions have been presented in [44] and in [1]. In [4] multiobjective algorithms on multimodal networks and are considered.

The problems of effective integration concern both multimodal transportation and innovative transportation services like Dial-a-Ride, Car Pooling and Car Sharing. In all cases, the coordination requirements are not adequately addressed in literature, where contributions are more focused on a single service [46]. As an example, consider the Dial-a-Ride case where the user specifies the starting point, the arrival point and the desired time. Clearly, the practical management of this service poses problems of scheduling of cars, sequencing of requests and route selection with the requirement of demand satisfaction and cost limitation. In this area, due to the large dimension of the problems, several metaheuristic algorithms have recently been proposed (e.g. Greedy Randomized Adaptive Search Procedure) [47]. These algorithms could also be applied to the Dial-a-Ride problem.

An interesting formulation has been proposed by J.F. Cordeau in "A Branch-and-Cut Algorithm for the Dial-a-Ride Problem". This formulation includes both decision variables for the arc choice and variables representing the time in which a specific event (get-on, get-off of a passenger) take place and, finally, variables that represent the number of passengers simultaneously present in the vehicle. Valid inequalities are also proposed and used in branch-and-cut algorithms.

It is also worth recalling the distinction between Off-line and On-line DAR. In the first case, all requests are known in advance and the optimization phase determines the optimal vehicle schedule. In the on line case, new requests can be accepted after the scheduling of vehicles has already been performed.
3.4. Impact of freight transport on urban mobility and models of city logistics

Freight transport represents an important aspect of the problems related to traffic in urban areas. The phenomenon is not new, yet it has been regarded as of secondary importance to the planning and analysis of metropolitan transport systems. Interest in this subject is currently growing rapidly due to the increased awareness of both institutions and citizens on the impact of goods distribution on congestion, noise and pollution. As a consequence, the necessity of studying, measuring, design management strategies and controlling these phenomena is widely recognized.

The flow of goods related to urban logistics deeply affects Italian cities where space is a very scarce resource, especially in the historical central areas. In this environment it is crucial to adapt all logistic activities to the different categories of goods. The differences concern times and delivery frequencies, types of vehicle used and so on. As regards the realization of urban logistic platforms, the work of [2] e [3] have studied a location-routing problem for food distribution. At present, no methodology focused on urban goods distribution is available. Recent studies [13] [20] [21] have been devoted to freight distribution in urban areas, mainly for the location, routing and scheduling of related activities. Yet, most of these efforts mainly represent the business point of view and more effort should be devoted to the institutional point of view.

3.5. Integration of optimization models, simulation models, and databases

Nowadays, there is a need to adopt unifying approaches and systems for integrating heterogeneous information and allowing users to take better advantage of transport systems while, at the same time, giving support to planning and management operations (network condition detection and analysis, design and simulation of re-route strategies, and the like).

Existing literature deals with specific problems of planning, management and control of a transportation system [45] [28] [29] [42] [37] [40]. However, an effective analysis and evaluation of modern transportation systems would require a global approach, not so thoroughly investigated in literature, especially from the mathematical model point of view.

The whole process of data acquisition, structuring, organization, filtering and analysis to obtain useful information to plan and manage traffic flows can be widely improved. We are in the phase of research in which the "intelligent" components of the system induce the analysis and integration requirements to be performed on the raw data to support the entire process.
References


AN INTEGRATED INFORMATION AND MANAGEMENT SYSTEM FOR TRADITIONAL AND DEMAND-RESPONSIVE MOBILITY SERVICES IN THE CITY OF MILAN

D. Ciccarelli, A. Colorni, A. Luè, V. Morrocchi, D. Rossi, R. Wolfler Calvo
University of Milan

An intensive use of Information and Communication Technologies (ICT) can improve substantially the answer to transport demand of men and freights in urban reality. The introduction of these technologies (Web site, SMS, UMTS) allows, for the first time, a rational and integrated use of information. User and managers have an easier access to the resources, moreover it is easier to adapt systems conceived for a specific urban situation to others. At the same time, last years have seen an increasing improvement of models and algorithms. In particular, those addressed as a consequence of the introduction of the Innovative Transport System (ITS) are now robust enough to be applied to real-life situation.

In this context the research group of Milano has improved both their methodological (several algorithms have conceived and realized as prototypes) and practical (development of pilot platforms and experiences of integrated systems) skills in the field of ITS. As for the methodological know-how, the research group has a deep knowledge of the Operation Research models and algorithms, which will allow them to plan and manage in the best way the several ITS [1]. We remind, as an example, the routing and scheduling problems focused on the determination of constrained paths in a Dial-a-Ride system [2, 3] and Car Pooling service [4, 5]. Together with these experiences it is important to recall the software packages already implemented for a number of European (AIUTO, TransPrice, Reflex) and national (PFT2) research projects, with the partnership of local administrations (Regione Lombardia, Comune di Como, Comune di Milano).

As for the development of pilot experiences or integrated systems let us mention for example the following activities: a feasibility study for a Metrotramway in Como, a Dial a Ride door to door service project for Milan and the following project DREAMS (Demand Responsive Extended Area Mobility Services) [6]. The system DREAMS is an advanced technical platform used to organize and manage the collective transports (traditional or innovative) in an urban area. The platform is a useful tool in supporting both the business mobility manager and the urban mobility manager. A single user uses it as a system to information retrieval or to adapt the transport solution to its constraints. It is a typical intervention complementary to the traditional ones for mobility in a urban area. The project, actually in the testing phase, has several characteristics, among them it is important to underline the knowledge management aspects and the integration among different transport databases.
The efforts of the research group of Milano for the next couple of years (2005-2007) will focus on the following four aspects:

The algorithms will be improved both theoretically and practically. In particular, it is important to understand which model, among those proposed in the literature, is more useful for modeling and solving Car Pooling and Car Sharing. The Car Pooling can be modeled as a matching, as a Vehicle Routing Problem with Time Windows (VRPTW) or as a maximal Clique. Each of them is more useful to model certain aspects of the real-life problem. Therefore an analysis of the literature will be done, looking for the best model for the Car Pooling system under consideration. The Car Sharing is often modeled as a fleet location problem. Thus this subject will be developed cooperating with the others research group of the project having an experience on location problems. The research group of Genova has developed a methodology for classifying the nodes of a graph according to two parameters: quality and strategy. Therefore an interesting research topic will be to understand how this classification could be useful and used to solve the Car Sharing location problem. A hypothesis to verify is that nodes with high strategy and low quality can be considered a useful subset of candidate nodes.

During the past PRIN project (2003-2005) several models have been integrated in a framework, showing that it could be a useful tool to answer the demand of sustainable mobility.

2a) since the framework has been applied to a no-real city called Utopia, the purpose of applying it to real-life cities (Milano, Como, Genova, Roma, Napoli) is an important and natural evolution of the research. The idea consists in showing how the system can be easily adapted to several different situations. This part of research will be done in collaboration with all the research groups of the project.

2b) a different but important and complementary aspect consists in creating a Web site useful to both the users and the manager. The manager will use it for updating data about traffic conditions in real-time, for data storage, data retrieval and simulations used for decision making. Moreover, a module called route planner will be developed. The purpose is to conceive an algorithm able to generate several transport solutions. This set of solutions should contain a mix of solutions: those based only on traditional transport systems, those based only on the ITS and those multimodal (combining both of them). Thus the citizens can find information and answers to their transport demand. This research will be held in collaboration with the research group of Genova which has an expertise on multimodal shortest path topic.

The flow of freights in an urban area reduces the network capacity affecting the traffic conditions. Therefore using some of the ITS, such as Car pooling and Dial-a-Ride, for freights, together with the idea of installing a huge depot (warehouse) outside the cities seems an interesting hypothesis to test for controlling the flow of goods. The ITS for goods can be used both off-line (all the demand is known before the system starts) and on-line (the demand can arrives after the service starts). The objective of this part of the research is to show how the framework can be enlarged including the transport of goods allowing the mobility manager of having a complete view of both the two flows. In this way the conflicts among flows and over positions can be avoided. Moreover, the idea of installing the framework on the Web makes any firm able to reach it from everywhere and any time it is necessary.
Transversally at the above-stated research activities (conception and development of algorithms, construction of a website and insertion of freight in the platform) the unit of Milano has the purpose of creating a short period evaluation model able to estimate changes in circulation’s condition. The primary target is the definition of a tool able to predict with sufficient accuracy short/medium time circulation’s condition on the main road network of a determinate area. The innovative idea of the project is the development of a tool able to estimate the circulation’s conditions for a definite day, and not only for a common average weekday, as happens for an output of a traditional traffic model. Forecasts refer to a specific condition characterized by precise occurrences: forecast concerns a close time (till the following three days). The use of the described tool can be strongly useful not only for those who have to planning their own daily or occasional trips, but also for those people who have to know the day-to-day real conditions of road network (transport planner and urban/provincial mobility managers).

References


A WEB-BASED SOFTWARE TOOL
FOR THE MANAGEMENT OF COMMUTERS
CAR POOLING IN THE MILAN AREA*

Alberto COLORNI1,2, Diego CICCARELLI2, Alessandro LUÈ2, Valentina MORROCCHI2

Abstract. Car pooling is a transportation system based on a shared use of private cars. Its main objective is to reduce the number of cars in use by grouping people. This paper presents a web-based software tool for the daily car pooling problem, applied in the area of Milan, Italy. The software has been developed within an ongoing research project, DREAMS (Demand Responsive Extended Area Mobility Services). We developed a system that focuses on the reliability of the overall system: (1) the users are provided with the expected schedule for their trip and they are informed immediately in case of delay or changes; (2) the use of the system is restricted to the employees of companies that have subscribed to the service.

1. Introduction

The growth of the travel demand, combined with a high use of private cars, causes an elevated level of congestion, resulting in environmental nuisance and accidents, and penalizing both users and the economy [8]. Increasing both the efficiency and the quality of public transport systems, and the development of alternative transportation systems may help in decreasing traffic congestion and air pollution in urban areas and improving accessibility for all citizens, including those who cannot use a private vehicle. This is the rationale behind the demand-responsive transportation services, such as dial-a-ride and car-pooling services, and mobility management strategies [4]. Mobility management measures

* This work has been financed by the Italian Ministry of the Environment and the Municipality of Milan, and is also partially supported by PRIN 2005 (Sistemi di Infomobilità e Distribuzione Merci in Aree Metropolitane), a research program by the Italian Ministry of Education, University and Research (MIUR).
1 INDACO - Politecnico di Milano, via Durando 38/A, 20158 Milano, Italy, alberto.colorni@polimi.it
2 Poliedra - Politecnico di Milano, via Garofalo 39, 20133 Milano, Italy, [ciccarelli, lue, morrocchi]@poliedra.polimi.it
are generally “soft” measures that do not require new infrastructures but try to shift the users’ modal choice from the private car to more sustainable modes making use of techniques based on information, communication, coordination, and organization [4,7].

This paper presents a web-based software tool for the daily car pooling problem, applied for the area of Milan, Italy. The software has been developed within an ongoing research project, DREAMS (Demand Responsive Extended Area Mobility Services), whose main goal is to study and develop a web-based integrated information and management system for traditional and demand-responsive mobility services [3].

The direct objective of the DREAMS system is to enable single persons or groups to organize their trips using modes different from their own private vehicle, even in areas which are not intensively served by the public transportation service. An indirect objective of the system is to promote mobility management measures and to experiment them in the area of Milan.

2. The DREAMS system

The role of DREAMS is to connect the demand side and the supply side in passenger urban transportation. The core of the system is the travel planner module, which, like a travel agency: (1) receives the requests of the users, (2) looks into the available transportation services, and (3) proposes a set of travel solutions taking into account the users’ preferences.

As shown in Figure 1, the travel planner module, besides the car mode, takes into account the services embedded in DREAMS, the car-pooling and the dial-a-ride, and a set of existing external services, such as the local municipal taxi system, the local public transportation, a car sharing service (called Guidami), and a dial-a-ride night service (called Radiobus). While the car-pooling and the dial-a-ride services are already equipped with real-time data communication systems, real-time data on traffic and public transportation are not available at the moment.
The routing and scheduling algorithms behind each software tool inside the DREAMS system (car-pooling, dial-a-ride, and travel planner) use a time-dependent network model of the city of Milan. The time-dependency of the links represents different levels of congestion and is modeled assigning to the link different speeds during the day. Moreover, in the future we will integrate our model with the model developed by the transportation agency of the Municipality of Milan, which will collect real-time information and forecast the link travel times.

3. THE CAR-POOLING SOFTWARE TOOL

Car pooling is a collective transportation system based on a shared use of private cars (vehicles), whose objective is to reduce the number of cars in use by grouping people. Car pooling can be operated in two main ways: Daily Car Pooling Problem (DCPP) or Long-term Car Pooling Problem (LCPP) [10]. In the case of DCPP, each day a number of users (servers) declare their availability for picking up and later bringing back colleagues (clients) on that particular day. The problem is to assign clients to servers and to identify the routes to be driven by the servers in order to minimize service costs and a penalty due to unassigned clients, subject to user time window and car capacity constraints. In the case of LCPP, each user is available both as a server and as a client and the objective is to define crews where each user will in turn, on different days, pick up the remaining pool members.

Many car-pooling services are non-monitored and unrestricted (e.g., [1,5,9]). A main issue for these type of services is to guarantee the reliability of the information. Such services only rarely suggest a matching between the users, but operates as a “Post-it” board, where everybody can consult or add information about travel routes or origins and destinations. More recently, more structured services have been designed, e.g. [6], to improve the reliability of the system. The idea is to set up a car pooling service among a restricted set of users. Moreover, the spontaneous user matching is substituted by a solution found by means of an algorithm.

We developed a web-based software tool for the DCPP problem that focuses on the reliability of the overall system, because the users are provided with the expected schedule for their trip and they are informed immediately in case of delay or changes, and the use of the system is restricted to the employees of given companies.

The communication means between the CP software and the users are the following.

- **Web:** all the functionalities of the DREAMS system are accessible on the Internet, based on the profile of the user.
- **Call center:** the users who do not have access to the Internet or who are not familiar with computers can access some functionalities via call center.
- **Short Message Service (SMS):** the system is able to communicate unexpected events (delays, cancellations, …).

The company mobility manager (CMM) and the area mobility manager (AMM) are figures instituted by the Italian law and have a central role in the process. In the trial, the AMM of the city of Milan will identify some CMMs who will have access to DREAMS and will experiment the car-pooling in their companies.
At the moment, the CP software manages the following profiles of users:

- the employees of the companies that subscribe to the car-pooling services
- the CMM, which designs the company transportation plan for the home-work employees’ trips;
- the AMM, which organizes the mobility management initiatives at municipal or provincial levels.

Using the DREAMS tool, the AMM coordinates the CMMs of the single companies, fixing the general rules of the car-pooling service. The CMMs of the single companies manage the employees’ accounts and adapt the general rules to the specific situations of the companies.

Each employee can define his/her car-pooling preferences with the following information: origin and destination of the trip, the set of days when he/she is willing to carpool, the earliest departure time and the latest arrival time, the maximum travel time, and the availability of the car (Figure 2).

The employees have to state their preferences in advance, e.g. by Thursday for the next week. Afterwards, e.g. on Thursday night, the system computes a matching among the employees (for details on the algorithm, see [2]). The result is a set of routes starting from the driver’s house, arriving at the workplace and stopping by the houses of the non-drivers,

![FIGURE 2 The on-line form where the employees define their car-pooling preferences.](image-url)
respecting the time windows and the car capacity constraints. The schedule and the routes are available, e.g. from Friday morning, on the DREAMS web site (Figure 3) and are sent via email to the employees.

At present, car-pooling is not widely applied for many reasons. First of all, it can be difficult to group adequately the car-pooling users when they do not reach a “critical mass”. In order to overcome this problem, the service has to be proposed in a large area. Moreover, it is important that all the employees perceive a clear benefit in order to continue using the service. The system has to calculate and show in a transparent way the costs and benefits of all the carpoolers (Figure 4 shows an example of benefits related with car-pooling: monetary costs, time and length). Therefore, we developed an accounting system associated with the employees’ trips, that takes into account how much the passengers owe to the driver. The system can be used on the single trip (that is, each passenger pays the driver at the end of each trip), or for a period of time, e.g. a year, taking into account the credit and the debit of each employee. This accounting system can be used to assign to the drivers possible incentives provided by the AMM.

FIGURE 3 The car-pooling software showing on-line the results of the user requests (in the center) and the schedule of a day (on the right)
FIGURE 4 An example of benefits related with car-pooling: monetary costs (C), time (T) and length (L). On the left, a situation with two separate drivers; on the right, the car-pool trip.

References


Abstract. The paper illustrates methodology and preliminary results obtained in the project of basic research “Interaction between signal settings and traffic flow patterns on road networks”, funded by the Italian Ministry of University and Research. The object of the project is to develop a general procedure to study, model and solve the problem of the optimal road network signal settings, by taking into account the interaction between signal control systems and traffic flow patterns. Different macroscopic and microscopic modelling approaches are discussed and applied to test cases. Moreover, different signal control strategies are introduced and tested by numerical experiments on a wide area of the road network of Roma.

1. Introduction

The paper describes the objectives, the methodology and the preliminary results of the research project “Interaction between signal settings and traffic flow patterns on road networks”, granted by the Italian Ministry of University and Research with the Fund for Investments on Basic Research (FIRB). The project joins three research units, belonging respectively to the University of Rome “La Sapienza”, the University “Roma Tre” and the Institute for Information System Analysis (IASI) of the Italian National Council of Research.

1 Università degli Studi di Roma La Sapienza, via Eudossiana, 18, Roma, gaetano.fusco@uniroma1.it; guido.gentile@uniroma1.it; lorenzo.meschini@uniroma1.it
2 Istituto di Analisi dei Sistemi ed Informatica, CNR, viale Manzoni, 30, Roma, bielli@iasi.rm.cnr.it; felici@iasi.rm.cnr.it
3 Università degli Studi Roma Tre, Dipartimento di Scienze dell’Ingegneria Civile, Via Vito Volterra, 62, eciprian@uniroma3.it; sgori@uniroma3.it; mnigro@uniroma3.it
The object of the research is to develop a general procedure to study, model and solve the problem of optimal road network signal settings, by taking into account the interaction between signal control systems and traffic flow patterns.

This problem is interesting by both a theoretical and an application point of view, since several mathematical studies and experimental results have shown that usual signal setting policies, which simply adjust signal parameters according to the measured traffic, may lead to system unstable solutions and deteriorate network performances.

At core of the problem is the difference between a user equilibrium flow pattern, where individuals choose their paths in order to minimize their own travel time, and a system optimizing flow that minimizes total delay of all users. The problem is object of an intensive research activity by the scientific community by many years (Cfr. [4] and [27]).

Although the combined problem of signal settings and traffic assignment is a well known non convex problem, a systematic analysis of the objective function of the global optimization signal setting problem highlighted in several examples large quasi-convex intervals for even different levels and patterns of traffic demand [5]. Still open problems are recognizing which conditions give rise to non convexity of the objective functions as well as individuating possible real-time strategies to keep the state of the traffic network near system optimal conditions.

The global optimal signal settings on a road network is a complex problem that involves dynamic traffic patterns, users’ route choice and real-time application of suitable control strategies. With several noticeable exceptions (Cfr. [13], [26], [1] and [2]), the problem is usually tackled by following an equilibrium approach, that is, by searching for a possibly optimal configuration of mutually consistent traffic flows and signal variables. The equilibrium approach assumes, but does not reproduce, the existence of a process where drivers correct their route choice day-to-day, according to the modified network performances, until no further improvement could be achieved. The traffic flow is modeled by a stationary relationship between link travel time and link flow. Although the delay at node approaches can be taken into account as it may be added to the link cost, neither a realistic modeling of traffic congestion on a road network or an explicit representation of real-time traffic control is possible.

In order to investigate real-time applications, the research project aims at extending the static modeling framework usually adopted in the scientific literature to apply dynamic assignment models and deal with both coordinated arteries and traffic-flow responsive signal settings.

The study is being developed focusing on the following topics:

- Modelling of traffic flow along coordinated arteries and urban traffic networks;
- Integration of signal control and dynamic assignment problems;
- Advanced models and methods for signal settings and traffic control.

2. Modeling framework

Many models with increasing levels of complexity have been analyzed in order to identify the most suitable for different possible applications to arteries or networks, in static or dynamic contexts.
2.1. Analytical traffic model for synchronized arteries

A simple analytical traffic model that describes the average delay along a synchronized artery (see ref. [24] for further details) has been here extended to apply different hypotheses of drivers’ behavior. Such hypotheses concern the capability of drivers to accelerate and catch up the tail of a preceding platoon, if any. Specifically, the following assumptions have been introduced: a) all vehicles follow the same trajectory and, after having been stopped at a signal, accelerate at a given rate until they reach a given cruise speed; b) all vehicles can accelerate along a link and catch up the previous platoon; c) all vehicles can accelerate up to a given maximal acceleration rate, so that they catch up the previous platoon only if the link is long enough.

This model has been developed specifically to assess synchronization strategies along signalized arteries. Delays at every approach of the artery are computed by checking, for each arriving platoon, which condition occurs among the following: a) the platoon arrives during the red (or -if a queue already exists at that approach- before the queue is cleared, so that the first vehicle of the platoon is stopped) and ends during the green; b) the platoon arrives during the green (or after the queue, if any, has been cleared) and ends during the red; c) the platoon arrives during the green after the queue, if any, has been cleared and ends during the green of the same cycle; d) the platoon arrives during the green and ends after the green of the next cycle. A different analytical expression for the delay holds for each condition.

Since the existence and the length of a queue cannot be determined before all platoons have been analyzed, the delay computation requires an iterative procedure that classifies the different platoons progressively. It is worth noting that such a procedure involves few iterations, because the platoons can both catch up each other along the links and recompose themselves at nodes, when more platoons arrive during the red phase.

2.2. Dynamic traffic network loading

The cell transmission model [6] has been here extended to deal with coordinated arteries, by introducing the following changes:

- in less than critical conditions, vehicles can move forwards by a generic quantity, dependent on the traffic density on the link;
- capability of simulating even complex signalized intersections;
- apply dynamic shortest paths and follow different OD pairs in the network loading process.

The extended cell transmission model has been selected as the most suitable mathematical model to assess and calibrate real-time feedback signal control systems that use the relative link occupancy as control variable, like Traffic Urban Control does [7].
2.3. Microscopic simulation of traffic flow

A microscopic simulation model aimed at reproducing the behavior and interaction of vehicles on road networks has been developed by the IASI [11]. This microscopic simulator is able to manage different aspects of vehicles’ dynamic, that is: acceleration and deceleration, based on the car-following principles [6] and on the state of downstream signals; lane changes and overtakes, based on gap-acceptance rules and vehicles’ paths. Vehicles are generated according to the negative exponential distribution, where the number of vehicles to be generated every minute can be time varying in order to simulate the variation of vehicular flows during the day.

The system includes a graphic interface, developed in the UNIX environment in C language with the public domain graphic libraries X11. The micro-simulation is animated on a reproduction of the road network, where the signals and the vehicles are visualized.

2.4. Dynamic traffic assignment model

Dynamic traffic assignment models are the most advanced modeling framework to simulate traffic patterns in congested urban networks, as they can both simulate the flow progression along links and reproduce the complex interaction between traffic flow and route choice behavior of users.

Several software packages for solving dynamic traffic assignment are now available. Two of the most popular and advanced (Dynasmart [9] and Dynameq [8]) have been taken into consideration in this research and are object of a systematic comparison. Both have been applied to a selected area of the road network of Roma, having 51 centroids, 300 nodes, 870 links, 70 signalized junctions (18 of which are under control). The figure below shows the study network as displayed by the two software packages.

![Figure 1. The study network in Dynameq (on the left) and Dynasmart (on the right)](image)

Both models are simulation based, but they differ as far as: the modeling approach of traffic flow (Dynameq assumes a fundamental triangular diagram, Dynasmart a modified Greenshield model [11]); the route choice (two phases for Dynameq - a path generation phase in which one new path at each iteration is added until the maximum predefined number of paths is reached, and a convergence phase in which no new paths are added - one...
phase for Dynasmart, which does not introduce any distinction between the generation phase and the convergence phase; the simulation process (Dynameq uses an event based procedure, Dynasmart a macroparticle simulation model); convergence criterion (time based for Dynameq, flow based for Dynasmart).

Thus, Dynameq aims especially at providing a detailed representation of traffic flow by following a microsimulation approach. It reproduces the traffic behavior at junction approaches by an explicit lane changing model and resolves conflicts at nodes by applying a gap acceptance rule. It simulates pre-timed signal traffic control and ramp metering.

On the other hand, Dynasmart provides a simpler framework to model the traffic flow and aims rather at allowing simulating the impact of real-time strategies of traffic control and information systems (like actuated traffic signal, ramp metering, variable message signs and vehicle route guidance) on users’ behavior.

Even if the tests of the models are still in progress, the following remarks can be made.

Comparison between static and dynamic assignment models shows that they provide not so different patterns of traffic flows ($R^2=0.61$), but, as expected, the static model overestimates the traffic flow, as it assumes that the whole trip demand is assigned within a steady-state simulation period.

Dynasmart and Dynameq predict very similar trends of the overall outflow of the network. However, their results are quite different as far as the route choice and, consequently, the traffic flow patterns.

Finally, simulation experiments highlight that a relevant issue of dynamic simulation assignment models is their capability to reproduce a temporary gridlock, which is actually experienced on the network in the rush hour. In fact, if the dynamic assignment model is forced to prevent the gridlock, the goodness-of-fit of simulated flow with respect to observed values decreases from $R^2=0.54$ to $R^2=0.46$.

### 3. Signal Control strategies

The traditional approach to signal control assumes traffic pattern as given and signal parameters (cycle length, green splits and offsets) as design variables. However, the combined problem of signal settings and traffic assignment generalizes the signal settings problem to concern the interaction between the control action and the users’ reaction. Usually, it takes as design variables only green splits of signals at isolated junctions. In the Global Optimization of Signal Settings Problem (Gossp), the whole system of signals is set to minimize an objective function that describe the global network performances. In the Local Optimization of Signal Settings Problem (Lossp), flow-responsive signals are set independently each other either to minimize a local objective function or following a given criterion, like equisaturation [29] or Po policy [25].

In the research project the usual approach is being extended in two directions: taking into account signal synchronization and develop signal control strategies that can be suitably applied in real-time.
3.1. Pre-timed synchronization

Signal synchronization of two-way arteries can be applied by following two different approaches -maximal bandwidth and minimum delay-, although a solution procedure that applies the former problem to search for the solution of the latter one has been developed [23]. Specifically, it is well known that, given the synchronization speed and the vector of distances between nodes, the offsets that maximize the green bandwidth are univocally determined by the cycle length of the artery. Such a property of the maximal bandwidth problem has been exploited to facilitate the search for a sub-optimal solution of the minimum delay problem. Thus, a linear search of the sub-optimal cycle length is first carried out starting from the minimum cycle length for the artery and then a local search is performed starting from the offset vector corresponding to that cycle length.

The analytical model described in section 2.1 makes it possible a quick evaluation of the artery delay without involving any simulation. The performances of the solutions obtained by applying the analytical model are then assessed by dynamic traffic assignment.

Moreover, two different strategies for one-way signal synchronization are also introduced in this paper. The simplest strategy consists of starting the red of each node just after the end of the primary platoon (that is, the platoon with the most number of vehicles). In the second strategy, the red is checked to start after the end of the every platoon and the position giving the least delay is then selected.

Preliminary tests conducted by carrying out a systematic analysis of the optimal offsets on an artery of Roma underline that the latter strategy outperforms the former one, especially for large traffic flows. Such an occurrence highlights that the performances of the artery are heavily affected by the complex process of platoon re-combination at nodes, so that the strategy that favor the main platoon can be ineffective if it is stopped later, while secondary platoons can run without being delayed.

Numerical experiments concerning the analytical model have been carried out on Viale Regina Margherita, a 6-node artery situated into the study network of Rome, in the peak-off period. In the first experiment, the common cycle length of the artery has been set equal to the maximal cycle length of the junctions. The green times have been adjusted consequently to keep the green splits unchanged. Thus, the offsets are the only control vector. In the second experiment, also the cycle length has been varied from the minimum cycle for the most critical junction of the artery (namely, 32s) until its current value (108.5s). As before, the objective functions of the two problems are the maximal bandwidth and the minimum delay for the artery.

The corresponding values of maximum bandwidth and sub-optimal minimum delay offsets are reported in Table 1 and Table 2, respectively.
Table 1. Offset of maximum bandwidth.

<table>
<thead>
<tr>
<th>Node N.</th>
<th>32</th>
<th>42</th>
<th>52</th>
<th>62</th>
<th>72</th>
<th>82</th>
<th>92</th>
<th>102</th>
<th>108.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>14.9</td>
<td>19.5</td>
<td>24.2</td>
<td>28.8</td>
<td>33.5</td>
<td>38.1</td>
<td>42.8</td>
<td>47.4</td>
<td>104.7</td>
</tr>
<tr>
<td>3</td>
<td>14.9</td>
<td>40.5</td>
<td>50.2</td>
<td>59.8</td>
<td>33.5</td>
<td>38.1</td>
<td>42.8</td>
<td>47.4</td>
<td>50.4</td>
</tr>
<tr>
<td>4</td>
<td>30.9</td>
<td>19.5</td>
<td>50.2</td>
<td>59.8</td>
<td>69.5</td>
<td>79.1</td>
<td>88.8</td>
<td>47.4</td>
<td>50.4</td>
</tr>
<tr>
<td>5</td>
<td>15.5</td>
<td>20.4</td>
<td>9.2</td>
<td>61.1</td>
<td>34.9</td>
<td>39.8</td>
<td>44.6</td>
<td>100.5</td>
<td>52.6</td>
</tr>
<tr>
<td>6</td>
<td>17.7</td>
<td>23.3</td>
<td>2.9</td>
<td>3.4</td>
<td>40.0</td>
<td>45.5</td>
<td>51.1</td>
<td>5.6</td>
<td>6.5</td>
</tr>
</tbody>
</table>

Table 2. Sub-optimal offset of minimum delay

<table>
<thead>
<tr>
<th>Node N.</th>
<th>32</th>
<th>42</th>
<th>52</th>
<th>62</th>
<th>72</th>
<th>82</th>
<th>92</th>
<th>102</th>
<th>108.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>31.9</td>
<td>39.5</td>
<td>45.2</td>
<td>10.8</td>
<td>12.5</td>
<td>64.1</td>
<td>76.8</td>
<td>94.4</td>
<td>1.7</td>
</tr>
<tr>
<td>3</td>
<td>15.9</td>
<td>5.5</td>
<td>45.2</td>
<td>59.8</td>
<td>4.5</td>
<td>38.1</td>
<td>42.8</td>
<td>47.4</td>
<td>47.5</td>
</tr>
<tr>
<td>4</td>
<td>28.9</td>
<td>19.5</td>
<td>26.2</td>
<td>37.8</td>
<td>12.5</td>
<td>59.1</td>
<td>64.8</td>
<td>47.4</td>
<td>50.5</td>
</tr>
<tr>
<td>5</td>
<td>18.5</td>
<td>41.4</td>
<td>27.2</td>
<td>61.1</td>
<td>35.0</td>
<td>39.8</td>
<td>44.6</td>
<td>84.5</td>
<td>85.1</td>
</tr>
<tr>
<td>6</td>
<td>19.7</td>
<td>32.3</td>
<td>4.9</td>
<td>3.4</td>
<td>43.0</td>
<td>54.5</td>
<td>63.1</td>
<td>7.6</td>
<td>5.9</td>
</tr>
</tbody>
</table>

Figure 2 plots the value of delay at nodes along the artery for different values of the cycle length and the sub-optimal offsets. The analytical delay model described in section 2.1 has been applied, by assuming acceleration capabilities of vehicles and then allowing a platoon catching up the preceding one. As expected, the delay function is non convex. The lowest value is obtained for a cycle length of 42s, while other local minima are obtained for the cycle lengths of 72s and 102s.
3.2. Actuated signal control based on logic programming

The traffic actuated strategies are constructed on the basis of the current configuration of the traffic flows or of the current number of vehicles at the controlled junctions. Thus, both the measurement of data and the elaborations aimed at identifying the traffic signal setting at each instant are carried out in real time.

The literature proposes several control methods that use mathematical programming; amongst others, successful examples of adaptive systems are the SCOOT system [15], SCATS [18], UTOPIA [19], COP [14]. New approaches have been recently proposed with the development of models based on fuzzy logic, such [3], [17].

Traffic control based on logic programming have also recently appeared. The method adopted in this paper, described in [10], [11] and [12], is one of the first models and applications of this type. It is an adaptive method actuated by vehicles that adopts logic programming to model and solve the decision problems associated with traffic control. Such a method can be applied with success to urban intersections with high levels of traffic where many different and unpredictable events contribute to large fluctuations in the number of vehicles that use the intersection. The logic programming methods based on vehicle counts make it possible to design the traffic control strategies with a high degree of simplicity and flexibility. The system makes use of a very efficient logic programming solver, the Leibniz System [28], that is capable of generating fast solution algorithms for the decision problems associated with traffic signal setting.

This method has been applied to a system of two signalized and coordinated junctions in Roma, Italy. The junctions, which are 180 meters apart and lay on a urban arterial road used to access and egress the city, are characterized by high congestion peaks and flow conditions varying rapidly over time. The schema of the study area is reported in Figure 3.
In cooperation with the Mobility Agency of Rome, a preliminary analysis of technical aspects, signal plans, traffic flows and congestion levels has been performed. Some historic samples of traffic data have been therefore integrated with a data collection campaign with the use of temporary loop detectors. Collected data have been used to calibrate and validate the simulation of the system with the micro simulator described in section 2.3.

To take into account the problems related to the quantity and quality of available traffic data, both incomplete and at an aggregate level, this application required to integrate and refine the detected data with analytic processes based on mathematical traffic flow models. To this end, a suitable macroscopic model based on the simplified theory of kinematic waves [21] was developed to estimate the queue length of each stream at a given intersection, as a function of vehicles counts taken at an upstream section.

![Figure 3. Scheme of the two junctions for the application of the signal control based on logic programming](image)

3.3. **Real-time feedback network signal regulator**

Real-time network traffic control strategies usually follow a hierarchical approach that optimizes cycle length, green splits and offsets by evaluating their impacts on the network performances at different level of time and space aggregation.

The Traffic Urban Control (TUC) model formulates the real-time traffic urban control system as a Linear-Quadratic (LQ) optimal control problem based on a store-and-forward type of mathematical modeling [7]. The basic control objective is to minimize the risk of oversaturation and the spillback of link queues by suitably varying, in a coordinated manner, the green-phase durations of all stages at all network junctions around some nominal values.
We are now aiming at exploiting the results obtained in [5], where at least one local minimum of the objective function of the global optimization of signal settings was found out and then can be assumed as a desired state of the network.

We so extend the approach followed in TUC and we introduce a Linear-Quadratic regulator designed to control the distance to the desired state of the network, and specifically the green split vector and traffic flow patterns that are solution of the global optimization of signal settings problem.

The LQ optimal control can be then be formulated as follows:

\[
g(k) = g_d - L[x(k) - x_d]
\]

where the following notations are introduced:
- \( g \): green time vector (at time \( k \));
- \( g_d \): desired green time vector (solution of static Goss problem);
- \( x \): link flow vector (at time \( k \));
- \( x_d \): link flow vector (solution of static Goss problem);
- \( L \): gain matrix.

The generalized cell transmission model briefly described in section 2.3 has been recognized as the most suitable approach for simulating such a real-time control strategy in the short-term. However, a wider analysis involving the long-term rerouting behavior of drivers and the en-route diversion of drivers, which may occur when large increase of congestion is experienced on some links of the network, necessarily requires the application of a dynamic equilibrium traffic assignment model. It is worth noting that such diversions are taken by the control system as random disturbances of the process, unless a real-time route guidance system exists that provides the drivers with updated information on the current or, better, on the future traffic conditions of the network.

4. Signal control and dynamic assignment problems

The integration of signal control and dynamic assignment has a twofold goal. On one hand, we aim at assessing the effectiveness of the synchronization algorithm introduced in section 3.1 and, on the other hand, we aim at appraising the effect of signal control on users’ route choice.

In the first experiment the cycle length of 108.5 seconds, which better fits the current conditions, is used and the corresponding suboptimal offset vector (last column of Table 2) has been applied in simulation to the artery.

The equilibrium dynamic traffic assignment of the total off peak hour trip demand has been so carried out with Dynameq.

Results reported in Table 3 show a significant reduction of delay on the artery (about 13%), which is even more important if we consider that, due to its improved performances, the artery attracts 13.8% more traffic. The total travel time on the whole network decreases as about 2%, although the study area is much wider than the influence area of the artery and, more important, the objective function of the synchronization algorithm accounts only the travel time of the artery.
Current scenario Synchronization Difference

<table>
<thead>
<tr>
<th></th>
<th>Current scenario</th>
<th>Synchronization</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average arterial travel time [s]</td>
<td>176.5</td>
<td>154.0</td>
<td>-12.7%</td>
</tr>
<tr>
<td>Average total arterial travel length [veh·km]</td>
<td>41.4</td>
<td>47.1</td>
<td>+13.8%</td>
</tr>
<tr>
<td>Average total network travel time [veh·h]</td>
<td>87.6</td>
<td>85.5</td>
<td>-2.4%</td>
</tr>
</tbody>
</table>

Table 3. Comparison of arterial and network results pre and post synchronization.

A further investigation concerns the analytical traffic model used by the synchronization algorithm. Two issues are here dealt with. The first regards the capability of such a simple model to estimate the average travel time along the artery. The second issue concerns the impact of the interaction between traffic control and drivers’ route choice to the solution found. As signals have been set by assuming the traffic flows as independent variables and the dynamic traffic assignment provided the dynamic equilibrium traffic flows on the artery, the analytical delay model has been applied again by taking the average equilibrium traffic flows as input.

The travel times on the artery computed by Dynameq and by the analytical model are compared in Figures 5 and 6.

Although the analytical model is based on a much simpler theory that assumes stationary and homogeneous traffic flow, the two models provide very similar results. As far as the two assumptions considered in the analytical model, namely the possibility or not that platoons can always catch-up each other, the former provided a better correspondence with the microsimulation performed by Dynameq.

Figure 4. Comparison between simulation and analytical model (Arterial direction 1, Cycle length 108.5 s).
However, the differences between the two models become larger when shorter values of cycle are used. In such cases, the green is long just enough to allow the traffic flow being served. Even a small random increase of traffic produces a temporary over-saturation and a relevant increase of delay, which can not be predicted by a stationary model. Indeed, much larger differences have been recorded between the microsimulation and the well-known Webster model, which has been applied to compute the delay at the lateral approaches. Thus, although the analytical model predicted that shorter cycle lengths would reduce delays on the artery, the results obtained by the dynamic traffic assignment model indicate that, even if the travel times along the artery improve, the overall performances of the network are slightly worse.

5. Conclusions and further developments

The research project “Interaction between signal settings and traffic flow patterns on road networks” aims at extending the global signal settings and traffic assignment problem to a more realistic dynamic context and at developing effective signal control strategies that lead the network performances to a stable desired state of traffic.

Different approaches to model traffic flow are introduced and compared: the analytical model of delay on a synchronized artery; a generalized cell transmission model; a microsimulation model. Two dynamic traffic assignment models, Dynasmar and Dymameq, are applied to a real-size network.

the two-way synchronization algorithm that applies the analytical delay model to carry out a linear search of the minimum delay solution, which revealed to be effective in reducing the travel times along a synchronized artery even if, moreover, the traffic flow on the artery increased as it became more convenient.
Current research concerns a Linear-Quadratic regulator designed to keep signals settings and traffic flows stably near a desired state, namely the solution of the global optimization of signal settings problem.

References


Abstract. A first analysis of relations between urban land use, urban transportation and urban feature is carried out. The review of the literature permits to highlight the current lack of coordination and feedback between the different planning elements. The result is the development of automobile-oriented transportation systems so producing deleterious impacts. Critical aspects (density of residences and activities, door-to-door speed) in developing useful and attractive transit network are studied and new design criteria to adopt are proposed.

1. Introduction

During the International City Design Competition, held in 1988-1989 and conducted by the School of Architecture and Urban Planning at the University of Wisconsin (Milwaukee), the competitors were “challenged to create innovative and credible visions for Milwaukee in the year 2020”: it provided an opportunity for experts in urban design to present concepts for the pattern of cities in the future. The projects were analyzed by Beimborn et al. [2] to understand if they were done according to the most important criteria for transit-sensitive design (as concentrations of trip ends, easy pedestrian movements and compatibility with transit operations). The results showed that two different worlds can be individuated: a world of land developers, architects and planners who seldom think of transit and a world of transit and traffic professionals who seldom deal with land use issues. Only few competitors used mass transit in their design and the main part of those that included transit did not use it appropriately.

---

1 This paper presents the results of the first phase of the research project PRIN 2005 “Instruments for the evaluation and monitoring of strategic actions in large degraded urban areas”
2 Civil Engineering Department, Roma Tre University, Via Vito Volterra 62 Roma, labtras@uniroma3.it
That is, transit does not seem to be a strong part of the future’s vision by those who will make many of the design decisions for the urban areas: in a future with finite resources, an automobile-dominated world is assumed.

Now, after fifteen years, nothing seems to be changed: the cities change according to the decentralization criteria, assuming a feature as different as possible from the old compact urban center; new roads grow up; the buildings and the blocks are divided by streets of more than two lanes per direction and therefore the pedestrian movements are disturbed. All the trips, for every different activity, require the use of the car and, consequently, distances, travel times, congestion and air pollution increase. In a successive step, after the road network design, transit system is developed into an automobile-oriented network, with limited impacts to develop a sustainable mobility system because the urban features, the land use and the network design are “not able to receive it”.

These phenomena are visible in all the world. In the city of Rome the growth of suburban areas determined a density lower than the forecasts of planning document (P.R.G. 1962): a land planned for 5 million’s people has been “consumed” for only 2.800.000 people. New areas present superblocks and large roads: the effect is to limit pedestrian movements and to promote automobile transport.

At the same time new projects for transit are proposed: the PROIMO Plan proposes for Rome, in a indefinite future, a global 76 km metro-subway network with a cost of 10,000 million €. But the simulation of this scenario shows only a 1% of modal shift from private to public transport.

There is a clear gap between land use and transport system planning: these decisions are made without consideration for the potential role of other modes, especially transit and walking. Unless this situation is changed, serious problems will arise in the future in attempting to adapt and retrofit our urban communities to changes in future resource availability.

Schlossberg and Brown [3] face the problem at a micro-level. They study transit-oriented development sites with walkability indicators, identifying the trips between the transit stop and the immediately surrounding area as a key component to create walking, bicycle and transit as transportation options by using, in an adequate way, various elements of the street network. Moreover transit-land connection can improve the possibility of walking between the transit stops for distances up to 0.3/0.8 km.

Beimborn et al. [1] think that a land use pattern based on transit should incorporate the following principles: transit service market oriented, land use pattern with concentrated trip ends, quality access system, transit-oriented streets. They underline that transit-sensitive land use design could be developed through the definition of transit corridor districts, separated from arterial highway corridors by a distance of at least 0.4-0.8 km, that would divide transit and auto-oriented land uses.

Sinha [4] takes into account the impact of urban population density, or more specifically of activity density (population and jobs), on transit and private car demand: increased density is associated with less ownership and use of automobiles, less road, parking facilities and high transit use. Increased density can also be expected to induce higher walking and not motorized trips, contributing further to sustainability.
2. Object of the paper

The first phase of the research project (PRIN 2005) is addressed to develop a design guidelines for land use, transportation network and urban features, by means of four steps:
- the identification of significant elements to include in the design procedure;
- the definition of feasible actions to adopt in the design procedure;
- the realization of the design guidelines;
- the implementation of a quantitative analysis to evaluate the validity and effectiveness of planning design.

The integrated procedure of land use and transit design have to be applied to different set of variables and elements: land use planning; street network configuration; transit network topology and service characteristics. The final aim of the design guidelines is to develop urban areas in which the transit, walking and bicycles modes have the main role in providing mobility.

The implementation of useful and effective design guidelines requires the study and the definition of actions at all the different levels of possible detail. In particular, according to the dimension and the different representation of the study area, two analysis levels are adopted called as macro and micro. The macro scale involves the whole urban area represented with the classic tools of the transportation analysis: graph network, traffic zones characterized by a set of information about socioeconomic data and demand matrix for the definition of trips distribution and entity. The micro scale involves a more detailed level of representation of the urban area characteristics finalized to reproduce, in a realistic way, the pedestrian trips and the accessibility phase.

In this work, the main design elements related to the transit-based land use and network design are identified. Besides, different possible design actions are proposed taking into account the different typologies and features of urban areas and their possible temporal and spatial evolution. The results proposed in this paper are only related to the macro level.

3. Definition of design elements

The develop of a transit-sensitive design process needs the clear understanding of the basic requirements to define transit systems competitive with automobile in terms of access, convenience and comfort. Successful transit is mainly correlated to the need of trips demand concentration to make feasible the provision of an high level of transit services.

The demand concentration requires an adequate market size in terms of activities and residents located as close as possible to the transit stops. The second consequent requirement should be developed adopting adequate transit network design criteria. These involve the construction of direct services, the reduction of transfers and waiting times at the stops and an easy access to the transport service to obtain a competitive transit system respect to the automobile system in terms of door-to-door travel time.

On the basis of this consideration, two design elements are first of all involved: the speed of the generic trips expressed in terms of door-to-door speed and the density expressed in number of residents and/or jobs per unit of land.
The speed computed on the whole door-to-door trip, has to be evaluated for all the sequential steps of the trip because the overall speed, in an urban context, is dependent on every component of the multimodal trip (pedestrian and/or bicycle or car, transit and pedestrian to final destination). It is clear that similar approach (door-to-door trip as reference) explains very well the large preference of the users for private transport: when a user exits from the house, he has to choose between an infinite frequency service (the automobile) and a fixed, definite, not always high frequency service (the transit service). In addition, for the public transport, the pedestrian trips to reach the service or the final destination, the waiting times at the stop and the possible transfers disutility have to be added at the in-vehicle travel time while, for the private transport, the only additional costs are related to the parking phase.

Figure 1. Importance of the access phase.

An example of the importance of the access phase to the transit system is illustrated in figure 1. Let us consider a trip from A to D (9 km) in which 7 km by rail (B to C), 1 km from origin to rail (A to B) and another 1 km from rail to final destination (C to D). Adopting a speed for the access phases (A to B and C to D) respectively equal to 4 km/h and 42 km/h, the total travel time (TTT) of the trip is 40 minutes (case 1).

As follows, the results of two different actions finalized to reduce the travel time, are reported. In case 2, increasing the rail speed from 42 km/h to 84 km/h, the TTT decrease up to 35 minutes with a percentage reduction of about 13% respect to case 1. In case 3 the access speed are doubled from 4 km/h to 8 km/h obtained a percentage reduction of the TTT equal to 37.5%.

Other critical element, for the macro level, is the urban densities, expressed in number of residents and/or jobs per unit of land. This is one of the key indicators to explain the level of automobile use, especially but not only for the western countries in which the automobile is at disposal of almost all the population.

As showed in the following figure reported from [4], in which are presented statistical analyses from a database of the major cities around the world, the transit boardings per capita per year increase, more than proportionally, with the increase of urban population densities and the car kilometers of travel per capita per year decrease, also in this case more than proportionally, with the increase of urban population densities.
4. Definition of design criteria

Based on this preliminary consideration, it is important to underline the importance of a new design approach in which the planning of transit systems is not sequential but it is coordinated and integrated with the land use definition and the urban feature design.

In terms of land use elements, higher densities permit better results for transit services. Therefore the urban areas have to grow up always as condensed built area not sprawled and/or with the concentration along the transit network of the major urban densities of residents and activities.

The possible positive impact of this approach is showed by the analysis of the case of the city of Rome in Italy. In this example, two different evolution of the land use and mobility system are analyzed by means of an exhaustive transportation analysis.

In the first case, according to the planning choices of the local administration (PROIMO), the transport network evolution is based on the building of 2 additional metro lines providing a final network of about 76 km and 4 lines.

In the second case the evolution is based on the concentration of the low density residence zones in the available areas along the existing 2 metro lines, so exploiting the residual transport capacity at disposal in the existing metro system (Figure 3): 12 “compact islands” are defined with 0.35-0.5 km radius and an average density increased from 100 persons/ha to 300 persons/ha, so involving about 200,000 persons (the average population of each island is variable from 17,000 to 30,000 persons). The value of 300 persons/ha was chosen as the maximum typical value of the population density in some livable urban neighborhoods of Rome, like the Parioli or Bologna Square areas.
The results are reported in table 1 using as reference the data of the current situation.

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Number of metro lines [length in km]</th>
<th>Private demand</th>
<th>Transit demand</th>
<th>Impact on modal shift</th>
</tr>
</thead>
<tbody>
<tr>
<td>Current</td>
<td>2 [36]</td>
<td>428,100</td>
<td>141,800</td>
<td>75/25</td>
</tr>
<tr>
<td>PROIMO</td>
<td>4 [76]</td>
<td>421,200</td>
<td>148,700</td>
<td>74/26</td>
</tr>
<tr>
<td>Higher densities</td>
<td>2 [36]</td>
<td>373,800</td>
<td>196,100</td>
<td>65/35</td>
</tr>
</tbody>
</table>

Table 1. Comparison of results for different design approaches in Rome.

This first example shows the importance to consider a set of different elements, not only related to new infrastructures for transit, to develop sustainable mobility system. So an high density of activities and residences along rail line can drastically improve the modal split of transit system: in fact the third scenario (higher densities) produces an increment of 38% on transit demand, while the PROIMO Plan’s scenario involves only a change of 4.8%.

About the transit network configuration, the increase of the competitiveness of transit services should be obtained working to improve the accessibility to the rail systems more than to increase the speed of these systems or to build totally new rail lines. The latter observation is, of course, valid for existing rail system when the development of access systems should be more effective for the users needs producing, at the same time, relevant savings in financial budget and construction times. The optimal configuration of the network, in other words, has to come from an adequate balance among service coverage, accessibility and direct transit routes. The aim is to maximize the door-to-door trip speed.

In this case, it is possible to adopt two different measures:

1) the improvement of the speed of the access phase to the transit system also by means of new transit systems, alternatives to bus feeder services;
2) the increase of the coverage area of rail systems realizing longer rail lines respect to rigorously linear routes so linking in a better way high density areas.

Usually, the first measure is adopted for existing rail system while the second one could be applied for new system under design.

In other word, the transit network configuration should be defined, first of all, according to the spatial location of activities and residence. An effective example is offered.
by the city of Venice in Italy, populated fifty years ago by 180,000 persons with a density of 250 persons/ha. This city is characterized by a very particular mobility system based on pedestrians and public transport by boats. The construction of a metro line and its route represent one of the main question about mobility in these years. The more obvious answer is about a linear route dividing Venice in two parts. But it is easy to demonstrate that the metro line coverage area is less than the case in which the route follows the Canal Grande, the main city curvilinear canal. Similar level of coverage area are obtained only by realizing 3 different linear lines (see for the comparison figure 4). Therefore, about Venice, the more effective network should be a curvilinear profile to maximize the accessibility and reducing the door-to-door travel time. Longer trips on the metro line are more than balanced by the reduction of the access phase length to the metro: in fact the curvilinear route is about 4 km, against the 2 km of the singular linear route, but the access phase is reduced from 1 to 0.4 km. So, adopting an access speed and a rail speed respectively of 4 km/h and 42 km/h, the TTT is about 18 minutes for the Canal Grande route against the 33 minutes of the linear route.

The two previous examples permit to identify the transit stops of the rail network (or equivalent mode) as the basic elements for a real coordination and integration among transit system, land use and urban feature design. First of all, transit stops, working on land use and transit system, have to represent the aggregation point of necessary demand’s level to provide an adequate transit service.

So it is possible to identify a first scheme of land use, the “compact island”, as defined in Rome’s example. This can be realized concentrating activities and residences around the transit stop into an area of about 0.35-0.5 km radius (usual considered as the maximum walking distance). A second scheme could be defined as a “compact area” of possible different features with a maximum extension of about 300-400 ha in which the demand level for transit service is not dependent only on concentration of activities and residences, but can be reached with the introduction of feeder services also by means of non-conventional systems (people mover or cable system etc.).

The successive step concerns the connections among the transit stops. The result is the main transit network configuration that, as observed in the Venice’s case, derives by a satisfactory balance among service coverage, accessibility and direct transit routes. This
step requires not only the definition of the appropriate choice at route level, but also the identification of the proper technology solution, mainly dependent on vehicle’s capacity, geometric route characteristics and the separation grade from automobile’s infrastructures.

Finally it is important to remark that such approach respects the complexity and the variety of the interaction among urban feature, land use and transport network design providing a set of design elements and feasible actions that could be combined obtained always valid but diversified solutions. For instance, it is possible to consider the city evolution proposed in figure 3. In this case, the activities and residences concentration close to the transit stop make viable the use of the surrounding areas for green spaces and other services for citizens.

5. Conclusion and further developments

These preliminary observations permit to underline that urban feature and land use decide about transport network topology and, by the other side, the transport network decides about activities and residences on the urban areas. Therefore, an integrated and coordinated design process of all different families of elements involved is required to develop a really sustainable life in the urban areas.

Two main patterns are founded for transit based land use design: compact island (about 0.35-0.5 km radius, high density) and compact area (an extension of 300-400 ha, medium density).

Further study activities will focus a) on the optimal combination of the different elements before analyzed and b) on the other steps of the research project, starting by the analysis of the design process at the micro level (first of all, walkability and street network configuration). The analysis will also address on the assessment of the land use and transit network design process impact in terms of costs and evolution of communication systems (internet, ITS etc.).

References


DEFINING SHARE OF PARK AND RIDE TRIPS USING FUZZY INFERENCE SYSTEM

Dr Andrzej SZARATA

Abstract. This paper presents results of applying Fuzzy Inference System for estimation of the number of potential Park and Ride users. Fuzzy approach can be justified in the case of rough data and in the environment of decision making by the users (in this case, existing human factor is difficult to estimate using traditional mathematical approach). It is seemed, that fuzzy methodology can be treated as a proper way to describe choice of mode of transport, and especially that uncertainty accompanied of choosing process has rather fuzzy character. The proposed approach is based on the Mamdani Fuzzy Inference System, and for calculation there was used Matlab software, with Fuzzy Logic Toolbox.

1. Introduction

Congestion on roads penetrating central urban areas is one of the most important problems of major cities all over the world. Increasing of length and capacity of road network cannot be the only solution. Each new road helps drivers drive easily, but it is a matter of time, that new link generates more traffic. Park and Ride is one of the way how to solve traffic problems and mitigate the process of urban sprawl. The system tries to join the advantages of public transport in high-density areas and the advantages of the private car in the outskirts. The first Park- and- Ride facilities in Europe appeared in the end of the fifties (London, 1958). Since then growth has been spectacular. In the year 1970 there were 24 cities which had P&R facilities – number of car parks were close to 1200. After 20 years, in the year 1990 there were 76 cities with number of P&R exceeded 3700 [1]. Park and Ride facilities usually were situated near railway-stations aimed at long distance commuters. P&R facilities appear also at the outskirts of towns, linked to the local public transport system. In the run of time the quality of Park & Ride improves.

1 Chair of Transportation Systems, Cracow University of Technology, aszarata@transys.wil.pl.edu.pl
In Polish conditions Park and Ride practically do not operates. There are only two small parking lots (35 bays total) close to Warsaw. However P&R is present in transportation policies of Polish agglomerations. The decision makers are aware that P&R can be treated as a tool for improvement of downtown accessibility, but they can not decide whether implement the system or not. Main reason of such situation is caused by lack of knowledge about effects of P&R implementation. Furthermore, it is very important to determine number of users interested in P&R trips. This knowledge can help to define size of P&R lot and evaluate its economical efficiency.

Modal split can be treated as significant element of trip modelling. It is assumed, that knowledge of modal split can reduce total management costs of transport about 20-30%. Models of modal split have rather local character – they refers to selected city because they are based on Comprehensive Travel Study (CTS). Those survey take into consideration travel behaviour of inhabitants based on their activities (e.g. Home-Work, School-Home, Home-Shopping). CTS are very complicated, time-consuming and require big effort. Polish cities should have to organize CTS every 5 years, but practically due to high costs survey are taken every 7-8 years. Moreover many cities (especially small and middle-sized) have no such travel analysis and it implies difficulties during trip modelling.

2. Description of the model

In traditional approach of determining probability of mode choice, as an influence factor it is used quotient of travel times – private transport (PrT) and public transport (PuT). However it will not take into consideration subjective impression of the users connected with time of individual parts of trip chain (e.g. access time, waiting time, egress time, dwell time). That is why, within this approach as a base for further analysis there will be taken generalized cost of the trip [2]. Generalized cost is a total cost which user should cover to travel from origin to destination. This value consider not only operating costs (e.g. cost of the fuel, insurance, public transport ticket) but also monetary equivalent of travel time. Each element of trip chain has different weight for its time and to make this value additive with operating costs, subjective time of the trip is multiplied by monetary value of time. During first step of proposed approach it will be necessary to define value of generalized cost of the trip for selected modes of transport (private car, public transport and Park and Ride). It is worth to emphasize, that cost concerns separate origin-destination (O-D) trips, and it should be estimated for all possible O-D relations which can be realized through analyzed P&R lot (separately for each P&R lot).

This methodology can be divided into following stages:

Procedure of split estimation between private and public transport (1st Stage).

Comparison results with accessible logit models, calibrated for Polish cities. Positive comparison of proposed fuzzy approach with traditional logit models can be treated as a confirmation of chosen methodology.

Procedure of section off P&R trips from private and public modes of transport (2nd Stage). Share of analysed modes of transport can be estimated using following formulas:

\[ U_{PrT} = U'_{PrT} - U_{P&PR} \]  

(1)
where:

- $U_{PtT}$ – total share of private transport trips;
- $U'_{PtT}$ - share of private transport trips from bimodal approach;
- $U_{PuT}$ – total share of public transport trips;
- $U'_{PuT}$ - share of public transport trips from bimodal approach;
- $U_{P&R}^{PrT}$ - share of P&R trips among trips initially carried out by private transport;
- $U_{P&R}^{PuT}$ - share of P&R trips among trips initially carried out by public transport.


This approach is based on Mamdani Fuzzy Inference System [3]. In general to compute the output of FIS, the following six steps should be fulfilled:

- determining a set of fuzzy rules;
- fuzzifying the inputs using the input membership functions,
- combining the fuzzified inputs according to the fuzzy rules to establish a rule strength,
- finding the consequence of the rule by combining the rule strength and the output membership function,
- combining the consequences to get an output distribution,
- defuzzifying the output distribution (this step is only if a crisp output is needed).

For our case a Fuzzy Inference System with one input and one output set will be used. Share of PrT among non-pedestrian trips depends on value of quotient $\Delta$ (4):

$$\Delta = \frac{K_{PrT}}{K_{PuT}}$$

where:

- $K_{PrT}$ – generalized cost of the trip of private car;
- $K_{PuT}$ - generalized cost of the trip of public transport;

For further analysis as a symbol of quotient of generalized cost of the trip will be used symbol “quotient $PrT/PuT$”. Assumed input data (“quotient $PrT/PuT$”) requires choice of
linguistic variable [4]. Values of linguistic variables are called terms and can take values like “small”, “large”, etc.

In this case the linguistic variable is proposed as follow:

\[ \Delta \left(\text{quotient PrT/PuT}\right) = \{\text{"very small"}, \text{"small"}, \text{"equal"}, \text{"large"}, \text{"very large"}\} \]

For each term it is necessary to define shape of membership function (MF). Due to lack of functioning P&R system in Poland one has decided to use experts’ inquiries (sample size: 40). Due to significant value of standard deviation it was necessary to conduct inquiries second time – using Delphic methodology (the same group of experts received results of first questionnaire and had to create their own membership function which fit the area between upper and lower quartile).

As an output data there will be taken linguistic variable \( U \) (“share”), described by following terms:

\[ U(\text{"share"}) = \{\text{"small"}, \text{"medium"}, \text{"large"}, \text{"very large"}\} \]

Range for values of linguistic variable “share” were defined according to results of existing bimodal functions for Polish cities. Membership function for variable “share” was estimated in the same way as variable “quotient PrT/PuT”.

Last element of Fuzzy Inference System is to define semantic rules which are based on implication IF…THEN. E.g. IF “quotient PrT/PuT” is “very small” THEN “share” is “small”. It was created set of all possible rules. In next step the rules which are not possible or illogical were rejected. Then for some less important rules there were assumed weights. Defining process of whole set of rules was based on expert analysis. The scheme of applied Fuzzy Inference System is presented on Fig. 1:

![Scheme of FIS defining share of PrT](image)

The effect of running fuzzy inference process is presented as a two dimensional chart. To compare results of fuzzy process with existing logit models it was necessary to approximate it. Using Mathematica software [5] there was proposed empirical function – formula 5:

\[ U_{PrT}^n = 65.2 \cdot \Delta^{0.1} \] (5)
To verify obtained results it was necessary to compare “fuzzy” approach with existing logit models. For comparison there were used modal split formulas for chosen cities in Poland: Katowice, Warszawa, Siemianowice Śląskie and Poznan. All functions are based on Comprehensive Travel Studies conducted in those cities during last few years. Figure 2 presents functions of modal split for selected cities:

![Fig. 2 Comparison of existing modal split models with “fuzzy” model](image)

Obtained result of fuzzy approach (which was received with completely different way than other models) do not significant diverge from dominance of graph bundles – especially in the range of $\Delta 0.4-1.5$. It is impossible to use statistical tools (e.g. goodness-of-fit test) to compare those functions. Mostly they are logit ones but the way how the function were obtained, in spite of appearances, are different. There is no common procedure taking into consideration surveys of travel behaviour (differences concern e.g. definition of trip, formulating questionnaires or result analysis). Moreover dispersion of presented models can prove strong influence of local behaviours on travelling. However similar results of fuzzy approach with traditional models can be treated as a proof of chosen methodology. This assumption states base for further analysis.


Share of Park and Ride refers to two groups of non-pedestrian trips: private cars and public transport. For car drivers it is obvious – some of them will choose P&R as a way of travelling. For second case – public transport passengers – P&R can be choose only for those who have driving licence and car access. Both groups of users require special procedure to section off potential P&R users.
4.1. Section of P&R users among private car trips

As it was in bimodal approach, base for estimation share of P&R users will be quotient of generalized cost of the trip by car and P&R. However cost is not the only variable during process of choosing mode of transport. In the case of Park and Ride it is important to know how parking lot is localized, if it is comfortable and safe for users and if it is possible to arrange other business (e.g. shopping). All those elements can affect the decision process during mode choice. It is the reason, that it should be implemented additional variable “P&R attraction”. As the input data there will be used two linguistic variables:

\[ A\{\text{quotient PrT/P&R}\} = \{\text{very small}, \text{ small}, \text{ equal}, \text{ quite large}, \text{ large}, \text{ very large}\} \]

and

\[ A\{\text{P&R attractiveness}\} = \{\text{small}, \text{ medium}, \text{ large}\} \]

To estimate shape of membership function there were used results of questionnaire for linguistic variable “quotient PrT/PuT”. Initially there were used the same shapes but afterwards they were modified. During conducted questionnaire surveys it was added question concerning arduousness of changing modes of transport (from private car to P&R). This knowledge will help to determine proper shape of membership function. Because changing mode of transport is connected with arduousness, it means that function should part – left side for membership functions for terms “very small” and “small” (if it is more difficult to left car and choose P&R it means that mentioned terms should have less value) and right side for terms “large” and “very large”. In the result of function parting it was obtained empty area between terms “equal” and “large”. It is important from operational point of view to have crossing level similar to 0,5 (value for which adjacent functions are crossing) . According to this it is important to add new function describing term “quite large”.

For shape of “P&R attraction” it was chosen gauss function according to literature studies [6]. It was assumed, that level of attraction for each P&R lot will be assessed individually and range of values will be between 0-10 (0 – the worst location, 10 – the best one).

As the output data there was used the same shapes as in the case of bimodal approach. The only difference refers to assumed range of share. According to results of P&R study in western countries it was obtained, that share of P&R trips is not higher than 10% - it was assumption for the data range. Because share of P&R users is not higher than 10% it confirms meaning of “large” as a linguistic variable. It is worth to emphasize, that assumed values refers only to single O-D trip. In the scale of whole city P&R will have less significance.

After run fuzzy process there were obtained results which can be presented in three dimensional chart. To make result more accessible the approximation has been made. The results of FIS and its approximation are presented on figure 3.
The formula describing obtained surface can be presented as follows:
\[
U_{P\&R}^{PrT}(x, y) = -0.3x^2 + 2.4x + 0.17y + 0.82
\]  
where:
- \( x \) – quotient of generalized cost of the trip PuT/P&R;
- \( y \) – attractiveness of P&R location.

Obtained results can help to estimate potential share of P&R trips among private transport. This value refers only to analyzed OD relation and can not be directly treated as a total split. Share of P&R trips fluctuates between 1.5% - for extremely unfavourable conditions and 6.4% for the most favourable ones. It can be strangely high value of share for the worst conditions, but fuzzy procedure rejects extreme input values (in this case close to 0% and 10%). Thanks to this fact it brings closer to way of noticed of phenomenon by human mind.

4.2. Section off P&R users among public transport trips

Last element of presented approach refers to share estimation of P&R trips among public transport users. This refers only to group of persons who have possibility of mode choice, i.e. they have driving licence and car availability.

Proposed approach has the same character as presented above – the input and output data, rules and membership functions are the same. The only difference refers to shape of membership functions of linguistic variable “\( \text{quotient PuT/P&R} \)”. In this case it was necessary to difference “slenderness” of MF and its positioning. It refers to relationship between changing modes of transport – it is easier to change public mode of transport into P&R than private car into P&R. If it is easier, it means that MF should be more “slenderness” (“squeezed”). It can explained the fact, that if it is easier to change PuT into P&R the impulse to do it for the user can be smaller. Figure 7 presents proposed shapes of membership functions for input data “\( \text{quotient PuT/P&R} \)”.

After run of fuzzy procedure obtained results can be presented in three dimensional chart. The surface represent share of P&R trips among public transport (fig. 4).
The surface has different character in comparison with private transport approach. The main reason is completely different layout of membership functions of input data. In this case MF are concentrated on value 1,0 and it is the reason of significant growth of share. It can be explained, that according to taken assumption (it is easier to change mode of transport from PuT into P&R than from PrT into P&R) small changes of “quotient PuT/P&R” is enough to convince user to change mode into P&R. For higher/smaller values of “quotient PuT/P&R” the share of P&R has stable character.

Because surface has complex shape, one has decided to divide it into three parts. For each of them there were approximated functions using Mathematica software. Formulas for each part are presented below:

\[ \Delta_{PuT/P&R} \in (0.2;0.8) \]

\[ U_{PuT/P&R}^{PuT}(x,y) = -0.01y^2 + 0.15y - 0.1xy + 1.7x + 1.1 \]  \hspace{1cm} (7)

\[ \Delta_{PuT/P&R} \in (0.8;1.4) \]

\[ U_{PuT/P&R}^{PuT}(x,y) = -0.03y + 0.2xy + 72.6x - 12.9e^x - 34.9 \ln x - 33.5 \]  \hspace{1cm} (8)

\[ \Delta_{PuT/P&R} \in (1.4;2.5) \]

\[ U_{PuT/P&R}^{PuT}(x,y) = -0.1x^2 + 0.3x + 0.1xy - 0.2y^2 + 0.4y + 4.5 \]  \hspace{1cm} (9)

where:

- \( x \) – quotient of generalized cost of the trip PuT/P&R;
- \( y \) – attractiveness of P&R location.
5. Application of the model

To apply developed model it was chosen one of the planned P&R lots in Warsaw – P&R Włoczańska. For this parking it was calculated generalized cost of the trip with initial data referring to average conditions: existing parking cost in the centre of Warsaw and free of charge P&R system (user has to pay only for the public transport ticket). Values of the travel time from selected districts of the city to the P&R and from P&R to pre-arranged point of the city centre (assumed destination area) was estimated using transportation model of Warsaw worked out in VISUM. For P&R Włoczańska there was assumed average conditions of parking accessibility and connections with public transport. It refers to variable “P&R attraction” which was assumed as 5.0. For calculated generalized cost of the trip and assumed value of “P&R attraction”, there was estimated share of different modes of transport. Results of calculations for chosen districts (6 of 12) are presented in table 1:

<table>
<thead>
<tr>
<th>Assumed districts from where it is possible to use P&amp;R</th>
<th>( \Delta )</th>
<th>( \Delta_{\text{PrT}} \Delta_{\text{PuT}} \Delta_{\text{P&amp;R}} )</th>
<th>( U_{\text{PrT}} ) [%]</th>
<th>( U_{\text{PuT}} ) [%]</th>
<th>( U_{\text{P&amp;R}} ) [%]</th>
<th>( U_{\text{PuT}}^\text{P&amp;R} ) [%]</th>
<th>( U_{\text{PuT}}^\text{PuT} ) [%]</th>
<th>Total modal split [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boernerowo</td>
<td>0.62</td>
<td>2.13 1.33</td>
<td>62.2</td>
<td>37.8</td>
<td>5.4</td>
<td>5.51</td>
<td>58.8</td>
<td>35.7</td>
</tr>
<tr>
<td>Wiśniewo</td>
<td>0.49</td>
<td>1.90 0.94</td>
<td>60.8</td>
<td>39.2</td>
<td>5.1</td>
<td>4.66</td>
<td>57.6</td>
<td>37.4</td>
</tr>
<tr>
<td>Tarchomin</td>
<td>0.48</td>
<td>1.78 0.86</td>
<td>60.6</td>
<td>39.4</td>
<td>5.0</td>
<td>4.42</td>
<td>57.6</td>
<td>37.6</td>
</tr>
<tr>
<td>Choszczówka</td>
<td>0.54</td>
<td>1.72 0.93</td>
<td>61.3</td>
<td>38.7</td>
<td>4.9</td>
<td>4.64</td>
<td>58.3</td>
<td>36.9</td>
</tr>
<tr>
<td>Łomianki</td>
<td>0.49</td>
<td>1.75 0.85</td>
<td>60.7</td>
<td>39.3</td>
<td>4.9</td>
<td>4.41</td>
<td>57.7</td>
<td>37.6</td>
</tr>
<tr>
<td>Legionowo</td>
<td>0.49</td>
<td>1.57 0.77</td>
<td>60.7</td>
<td>39.3</td>
<td>4.7</td>
<td>4.28</td>
<td>57.8</td>
<td>37.6</td>
</tr>
</tbody>
</table>

Table 1. Modal Split for potential users of P&R Włoczańska.

For assumed assumptions (and for all analyzed districts) average share of P&R trips is equal to 4.8% (it refers to users which have possibility of mode choice). It seemed that obtained value is a little bit overstated, but it is worth to emphasize relationship between generalized costs of the trip of private car and P&R – in all cases private car has lowest values than P&R trip. However results are similar to share of P&R trips achieved in Western European countries.

To estimate number of potential users of P&R lot it is necessary to define size of traffic flow from taken districts to pre-arranged point of the trips destination. There were used OD matrices for PrT and PuT. It was necessary to aggregate origin zones (chosen external districts of the city) and destination zones (place close to pre-arranged destination point). As a result it was obtained number of cars and number of PuT users which can possibly use analysed P&R lot. For passengers flow it was assumed, that 10% of total has possibility of mode choice and average number of passenger in private car do not exceed 1.2. Thanks to those values it was possible to convert users of PuT which have possibility of choice into vehicle volume. Results of calculation are presented in table 2:
Table 2 Estimation of number of private cars which can possibly park on P&R lot.

<table>
<thead>
<tr>
<th>Number of potential P&amp;R users:</th>
<th>Volume obtained from OD matrices</th>
<th>Share of P&amp;R trips among:</th>
<th>Number of P&amp;R users</th>
</tr>
</thead>
<tbody>
<tr>
<td>Among PrT</td>
<td>3384 cars</td>
<td>4.8%</td>
<td>162</td>
</tr>
<tr>
<td></td>
<td>14 402 passengers</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Among PuT</td>
<td>1440 passengers have possibility of choice</td>
<td>4.2%</td>
<td>50</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1200 cars</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total:</td>
<td>212</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

6. Conclusions

Presented methodology section off P&R trips among private car users is circuitous and complicated. It requires interpretation of potential travel behaviours and its transmission into fuzzy procedure. Proposed formulas can be used as a tool for estimation of size of future parking lots and it can be also used to evaluate economical efficiency of the system (e.g. in the frame of Feasibility Study).

In Polish conditions problem of Park and Ride share in modal split was not investigated. There is no possibility of verification obtained results due to lack of functioning P&R system, but comparison of bimodal fuzzy approach with existing logit models of modal split can be treated as initial approval. However in near future it is planned to build P&R system in Warsaw. It will be possibility to test fuzzy models and compare them with real situation. It is future activity of author as well as sensitivity of the model.

References


ITS FRAMEWORK ARCHITECTURE IN A MULTIMODAL TRANSPORT SYSTEM

Mario BINETTI¹, Paola AMORUSO¹, Margherita MASCIA²

Abstract. Recently the development of communication and information technology has played an important role in increasing efficiency and safety in public transport. The wide spread of the Intelligent Transport Systems required a standardised approach, in order to fully exploit their functionalities. This has lead to the definition of the ITS Framework Architecture, which identifies a structured process and provides guidelines for high level ITS design. The adoption of the ITS Architecture allows creating an open system, ease to expand and communicating with other ITS via data exchange. In this context “PITAGORA” (Piattaforma Telematica per l’Informazione e la Gestione dei Sistemi di Trasporto Collettivo) project aims to develop an integrated public transport management system and multimodal traveller information service, compliant with the ITS Architecture.

1. Introduction

In order to tackle the rise of transport demand for passengers and goods, the White Book [1] highlights the need to recognise sustainable modes of transport that represent competitive alternatives to private car. In order to face the increasing number of motorists who share the same limited road resource with Public Transport systems, it is more and more evident that network expansion is not an option to tackle daily congestion. The way to follow consists in maximising the use of the existing roads and promoting multimodality across different transport modes, through the use of ITS.

¹ Highways and Transportation Department (D.V.T.) – Technical University of Bari, via E.Orabona, 4 70125 Bari – Italy, binetti@poliba.it, paolaamoruso@alice.it
² Faber Maunsell Ltd Knowles House ITS Services, High Street, Redhill, Surrey, RH1 1RN, UK, margherita.mascia@fabermaunsell.com
*This work has been financed by the Italian Ministry of Education, University and Research (MIUR) MIUR - FAR D.Lgs. N. 297/1999 (D.M. 8 agosto 2000, n. 593, art.5/6, prot. 7523/2002, http://sisinflab.poliba.it/pitagora
The various Intelligent Transport Systems need to be designed as a whole, integrated system, rather than isolated stand alone systems. This goal is achievable thanks to the ITS Architecture, a useful tool that supports ITS design in an interoperable perspective.

2. ITS Framework Architecture

The ITS Architecture was first developed in USA, then in Europe within the projects supported by European Commission: Karen and Frame [2]. The first project aimed to the recognition of the User Needs for different areas of ITS. “User” is referred to any potential user of the ITS architecture, including travellers and transport operators. The User Needs represent the requirements that the system has to meet in accordance with the scope of the Local Authority/private transport operator, who commissioned the ITS design (the Stakeholders”).

All User Needs are grouped in ten main categories according to the transport context where the technologies are applied. The User Need Groups are here reported:

- Group 1 General;
- Group 2 Management Activities;
- Group 3 Policing / Enforcing;
- Group 4 Financial Transaction;
- Group 5 Emergency Services;
- Group 6 Travel Information;
- Group 7 Traffic Management;
- Group 8 In-vehicle Systems;
- Group 9 Freight and Fleet Operations;
- Group 10 Public Transport.

The Frame project identified the main steps to follow in ITS design.

Public Authorities or private transport operators commissioning or interested in implementing the Intelligent Transport System highlight the high level desired characteristics, so called “Stakeholders Aspirations”. They represent the starting point for the procedure to follow in order to design a system compliant with ITS Architecture. The steps of the procedure are described in followings:

- From Stakeholders Aspirations to User Needs; the Stakeholder Aspirations are translated in detailed user needs, selected from the above groups.
- From User Needs to functions; each user need is associated to one or more functions that the system has to perform in order to meet the Stakeholder Aspirations. Hence the functions related to the system to be designed are selected. Functions have some data as input and output as well. At this stage the functional architecture is completed and the results can be drawn in diagrams called “Data Flow Diagrams” where functions and functional data flows are indicated.
- From the functional architecture to the physical architecture; each function has to be associated to a physical location where it will be carried out: in the control room, in the vehicle, roadside or handheld terminal. Generally more than one functions will be located in the same physical entity, creating a subsystem. At the end of the process several subsystems will be identified. They will exchange physical data, which are
made of several functional data flows. This stage is easily represented in diagrams called “Subsystem Diagrams”.

The results of the procedure are these subsystems with a clear identification of the functions and data to be exchanged. Two more steps can cover other aspects of the design. One consists in a detailed specification of capacity required for data transmission and communication channels. At last the organisational architecture aims to identify for each subsystem the staff in charge of the maintenance, reliability checks of data exchanged, the owner. This point of view is essential to assure that the whole system works efficiently.

The main benefits derived from the use of ITS Architecture are that the designed system:

- Satisfies the Stakeholders Aspirations;
- Is interoperable, an open system;
- Is modular, so easy to maintain and expand;
- Is designed with a standard approach, understandable in Europe.

Within the same project frame, a tool has been developed that makes the adoption of the ITS Architecture quite easier; most of the stages of the above procedure become automatic.

Each European country has been suggested to create his own architecture, taking into account different languages and peculiar aspects of the country or possible expansions of the ITS.

Italy created the Italian ITS Architecture (“Artist”) [3], where the most innovative issue is the introduction of another group of User Need, functions and data flow related to intermodality for freight transport.

In this work the Italian ITS Architecture has been adopted to design a multimodal transport system aimed to provide passengers with real time information across different modes of transports. This is the core of Pitagora project, described in the next section.

### 3. PITAGORA

The key words for Pitagora project are multimodal trips and multimodal traveller information. The city of Bari, with 300,000 inhabitants, is the test site of the project. One transport operator manages public transport at urban level, whereas different transport operators are in charge for road and railways transport services at extra urban level.

Integration between road public and private transport is performed through a Park&Ride system featuring an integrated fare for stop and travel.

In this context the project aims to develop a pre-trip and on-trip information system useful to all stakeholders, particularly multimodal travellers.

Travellers shall be able to get real time information about different modes and routes from the origin to their destination on multimodal network via static devices (Internet points, bus stop displays, kiosks,..) as well as mobile devices (mobile phones, PDAs,..). Also user profiles will be set up, in order to provide travellers with bespoke information about delays and disruptions that affect their planned travel times, costs and routes.

Automatic Vehicle Detection device will enable bus operators to manage their fleet more effectively through real time monitoring of vehicles running late or ahead of schedule.
The above described system is made of several sub-systems, obeying to different management rules and technical regulations. The effectiveness of the overall system is assured by a structured model that identifies functionalities and relationships between all subsystems and involved stakeholders. This triggered the adoption of National ITS Framework Architecture (Artist) in Pitagora applied to multimodal passengers information system.

4. Architectural model for the multimodal transport system in PITAGORA

As described in the paragraph 2, the starting point to adopt the ITS Architecture is the identification of Stakeholder Aspirations. They have been identified for the multimodal transport system to be developed within Pitagora; the system shall be able:
- to provide passengers with information about different modes of transport (timetable, routes, disruptions);
- to provide passengers with access to the system via different channels (internet, bus stop display, call centre, PDA, on board vehicle, kiosk, mobile phone);
- to collect and transmit data detected through on board equipment about vehicle location and status to the control room;
- to integrate communication for data transfer between control centres of various transport operators;
- to manage tariffs and control season tickets;
- to deliver information about points of interest;
- to manage multimodal network;
- to gather data about public transport demand.

The analysis of the Stakeholder Aspirations leads to the selection of detailed User Needs from those listed in ITS Architecture, which more represent the Stakeholder Aspirations for Pitagora. The User Need groups involved are reported in the Table 1.

<table>
<thead>
<tr>
<th>User Need Group</th>
<th>Group Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Management Activities</td>
<td>2</td>
</tr>
<tr>
<td>Financial Transaction</td>
<td>4</td>
</tr>
<tr>
<td>Travel Information</td>
<td>6</td>
</tr>
<tr>
<td>Traffic Management</td>
<td>7</td>
</tr>
<tr>
<td>Public Transport</td>
<td>10</td>
</tr>
</tbody>
</table>

Table 1. User Need Groups for Pitagora

In order to meet the selected User Needs, the system shall perform appropriate functionalities. Each user need is associated to one or more functions, already identified in ITS Architecture. The functions have a hierarchical structured: the highest level is made of Functional Areas, who refer to wide context of ITS applications; the lowest level functions describe system functionalities more in details. The Functional Areas selected for the Pitagora are reported in Table 2 together with their acronyms.
<table>
<thead>
<tr>
<th>Functional Areas</th>
<th>Acronyms</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Provide Electronic Payment Facilities</td>
<td>pepf</td>
</tr>
<tr>
<td>3. Manage Traffic</td>
<td>mt</td>
</tr>
<tr>
<td>4. Manage Public Transport Operations</td>
<td>mpto</td>
</tr>
<tr>
<td>6. Provide Traveller Journey Assistance</td>
<td>ptja</td>
</tr>
<tr>
<td>9. Provide Archive</td>
<td>pa</td>
</tr>
</tbody>
</table>

Table 2. Functional Areas for Pitagora

All the functions and the data flows exchanged within the system and with external entities are depicted in the following diagram, called Data Flow Diagram 0 or “DFD 0” where 0 refers to the highest functional level.

The following step of the procedure consists in locating each function in a physical site, so that all the functions associated to the same site will generate a subsystem. For Pitagora the subsystems identified are:

- On board vehicle; it is the vehicle used by travellers to move, equipped with location and communication devices.

Figure 1: Data Flow Diagram
Infrastructure where transport service is provided and a suitable communication system is installed.

Financial Organisation; it is the location of technical and operating offices in charge of financial operations.

Control Organisation; it is the location of technical and operating offices in charge of service management in conformity with regulations.

Traveller; it is the physical location where the traveller uses the system (e.g. home, office)

Management Organisation; it is the location of technical and operating offices in charge of service management.

Planning Organisation; it is the location of technical and operating offices in charge of service planning.

VASP (Value Added Service Provided); it is the location of technical and operating offices in charge of providing advanced and bespoke services to travellers.

Subsystems exchange physical data flows, made of one or more functional data. The subsystems identified and all the physical data exchanged between them can be drawn in a diagram here omitted for the sake of brevity.

5. Conclusions

This work has been carried out within the research project Pitagora supported by the Italian Minister of Research. The project focuses on the development of a telematic platform that enables transport operators to optimise fleet management and to provide real time seamless multimodal information to passengers. The integration between different modes of transport together with an improvement of the level of service in Public Transport makes a step forward the achievement of a sustainable mobility.

The implementation of a multimodal transport system, able to provide real time information to passengers, requires a structural approach in carrying out a design of an open and efficient system. ITS architecture is the key tool to design such a system. In this work it has been adopted allowing designing of a modular open platform, able to be easily integrated with legacy and future systems.

References


TOWARDS MOBILE SEMANTIC GRID FOR INFOMOBILITY

Mario BINETTI¹, Paola AMORUSO¹, Eugenio DI SCIASCIO², Michele RUTA², Agnese PINTO²

Abstract. Flexible and effective information services in transport call for novel and emerging telematics technologies. Such services require integration and processing of data coming from different information sources together with simple interfaces to make them available via static and mobile devices. A significant development for transport information services can be achieved exploiting networks featuring mobile nodes – with different characteristics and properties - represented by vehicles moving on road network. A semantic-based approach applied to such networks, allows building a high level infrastructure - mobile semantic grid - able to meet requirements of mobile grid with the interaction between connected devices and generic users provided with a simple interface, e.g. a handheld device. In this context the city of Bari, the test site for Pitagora project, represents the first step towards the development of a mobile semantic grid at a real scale. In this grid buses represent nodes as elements of intermodal transport system.

1. Introduction

In the last years a rapid development of telematic technologies has allowed the migration for wireless sensors from static first generation ones characterised by low data rate, towards dynamic second generation sensors with high data rate. This boosts new possible applications within Intelligent Transport Systems. Nowadays most of vehicles are equipped with diagnosis sensors; hence they can easily represent nodes in a network of mobile sensors. Also the widespread use of static and mobile terminals enables travellers to represent as well a mobile node within the network. He can hence be capable to get information in order to plan his trip and receive assistance during his trip. Moreover the

¹ Highways and Transportation Department (D.V.T.) –Technical University of Bari. via E.Orabona,4 70125 Bari – Italy, binetti@poliba.it, paolaamoruso@alice.it
² Electrical & Electronics Engineering Department (D.E.E.) Technical University of Bari. via E.Orabona,4 70125 Bari – Italy, {disciascio, m.ruta, agnese.pinto@poliba.it}
application of telematics technologies plays an important role in the management of an integrated transport system. With particular reference to a multimodal transport system, it allows an improvement in the level of service provided as well as an effect on modal split, increasing public transport in respect to private transport.

2. PITAGORA

The key words for Pitagora project are multimodal trips and multimodal traveller information. The city of Bari, with 300,000 inhabitants, is the test site of the project. The project aims to develop a pre-trip and on-trip information system useful to all stakeholders, particularly multimodal travellers. Travellers shall be able to get real time information about different modes and routes from the origin to their destination on multimodal network via static devices (Internet points, bus stop displays, kiosks,..) as well as mobile devices (mobile phones, PDAs,..). Also user profiles will be set up, in order to provide travellers with bespoke information about delays and disruptions that affect their planned travel times, costs and routes. Automatic Vehicle Detection device will enable bus operators to manage their fleet more effectively through real time monitoring of vehicles running late or ahead of schedule. In this context, communications between on board vehicle (bus equipped with location and communication devices), static devices (bus stops displays, kiosks, etc..) and control room as well as communications between traveller and a generic device used to acquire information can build two scenarios: the first is based on static sensors of the telematic grid; communication channels adopted are GSM, GPRS, MMS, etc.; the second one is based on a wireless communication system, with dynamic sensors that define a mobile grid. Physical extension of this grid on the land depends on the position of the buses (dynamic sensors).

3. Proposed Mobile Semantic Grid Infrastructure

Recent technological advances permit to evolve basic wireless networks in new ones where also sensors are involved and data rates are appreciably higher. Mobile Ad-hoc NETwork (MANET) infrastructures allow more and more powerful application as modern handheld devices integrate a good computational equipment making possible a strong decentralization of the network knowledge and a really ubiquitous computing. From this point of view, an emerging paradigm is the mobile grid one [2]. Although significant work has been done towards mobile grids, it has been basically aimed at extending wired paradigms and approaches to wireless infrastructures. But due to the deeply different nature of two contexts, real potentialities of mobile grids are still hidden and not exploited. Although grid computing offers significant enhancements to user capabilities for computation, information processing and collaboration, only a high level approach in data dissemination and resource discovery within the grid could allow to fully exploit these features [6]. We aim at building and experimenting a Mobile Semantic Grid infrastructure where Semantic Web techniques and technologies are seamlessly used in order to characterize grid nodes and actors, as well as to annotate information they manage in an infomobility scenario. Semantic Grids are an extension of the current Grids. There,
resources and services become machine understandable, better enabling computers and people to work in cooperation [2]. Here grid nodes are buses moving within a specified urban area, handheld devices of users as well as – considering stable nodes – survey centrals, data collectors and reasoning centres. Hence, by taking into account data received from various wireless sensors within the grid (traffic entity, environmental parameters, location values) as well as semantic annotations of grid resources, a generic user – external to the grid – could retrieve advanced information to this complex “black box” interacting with front end nodes. In particular the grid will combine data managed by different mobile nodes (conceived either as semantically annotated descriptions or as contextual parameters) in order to compute best matching resources w.r.t. user requests. We will use at lower layers communication protocols which allow the correct management of grid information, slightly modifying its application layer in order to enable the use of semantics. Data interchange among wireless sensors and survey centrals or data repositories within the grid will take place by means of ZigBee protocol standard, whereas grid hotspots (mobile vehicular information systems as well as networked service providers) are directly exposed towards external users by means of a Bluetooth interface. ZigBee is a high level protocol suite based on the IEEE 802.15.4 standard. Over the IEEE 802.15.4, the ZigBee Alliance has defined further layers of the protocol stack. Generally the allowed POS (Personal Operating Space) is up to one hundred metres circa, but more nodes can be combined to build a complex bridged infrastructure able to dynamically cover larger areas. ZigBee wireless networks are usually labelled as LR-WPAN (Low Rate WPAN). They are mobile networks simple and low cost permitting application oriented wireless connections with not too much high throughputs. Basically ZigBee allows to build general purpose auto-configuring wireless networks which integrate a variable number of different devices ranging from few ten to some thousand unities. Hence the basic protocol structure offers a good support for application requiring a high environmental covering level. It allows to monitor a heterogeneous and complex context as well as to interact with devices within the environment. For our purpose the high flexibility offered by the protocol will be used as basic network infrastructure in our Mobile Semantic Grid implementation. As previously hinted, together with ZigBee, in the proposed approach we will use Bluetooth to grant the interaction of grid front end nodes with the user. This is mostly due to a couple of reasons. First of all Bluetooth support is widespread in commercial handheld devices. Furthermore, granting a higher data rate, Bluetooth appears more effective in terms of interaction with the user. On the other hand the integration between Bluetooth and ZigBee involves only respective application layer (Service Discovery Protocols) and so it is easier to be implemented. Bluetooth is a short range, low power technology [4], which grants peer to peer interaction between hosts. In such a mobile infrastructure there is one or more devices providing and using services. Service Discovery (SD) protocol in original Bluetooth standard is based on a syntactic matching, which is largely inefficient in ad-hoc environments where there is not a common service interface. A more advanced usage of service discovery protocol is desirable, associating semantic descriptions to the services rather than simple numeric identifiers. For such purpose knowledge representation techniques shall be exploited and adapted to the ubiquitous environment [5], extending the SD protocol without troubling the basic one.
4. Research goal and application framework

By adopting the protocols briefly sketched above we will implement the communication infrastructure for our mobile semantic grid. A fundamental role in this approach is played by an adequate paradigm for data dissemination within the grid. In the proposed framework results and techniques practised for MANETs [3] will be adapted and extended. Basically, we aim to provide a general and common semantic layer over the application one, able to allow the usage of semantic based services also granting a backward compatibility with original standard functionalities. An effective resource discovery approach must take into account the limited bandwidth of mobile ad-hoc networks and their extremely volatile topology. Hence a hybrid solution appears more straightforward w.r.t. a completely decentralized one. Information about grid resources and services will be disseminated within the grid in a totally decentralized fashion allowing the rapid and simple identification of a node, whereas the resource or the services themselves will be mainly provided on-demand to the requester. The idea is that Knowledge is completely distributed within the grid, but thanks to a capillary data dissemination involving short information about it, the framework is able to easily retrieve semantically annotated services/resources so allowing the direct download from the provider. Notice that technologies designed for Semantic Web have to be adapted and integrated for an effective use in ubiquitous ad-hoc scenarios. In particular the resource discovery paradigm for MANETs has to be shaped taking into account peculiarities of such a system. We exploit an approach which uses a two step discovery. The first step concerns the advertisement of node resources within the grid and the second one allows to select the one best matching user requirements (according to semantic similarity criteria). Basically, the service advertisement is the necessary support to the subsequent semantic enabled discovery and selection procedure. We hence aim at adapting an existing data dissemination protocol [3] to mobile semantic grid applications with a twofold objective:

1. to build an adequate infrastructure for the application layer of the protocol in order to support advertisement and discovery of complex services/resources
2. to perform the tuning of the system for infomobility scenarios.

The first stage basically concerns the adaptation of the ZigBee resource discovery layer in order to implement an efficient system of data dissemination, and successive information retrieval. Obviously the protocol has to be modified to integrate the support for semantically annotated service/resource description. The mobile grid will contain two kind of nodes with different relevance according to their computational capabilities. The first class nodes play a passive role within the grid as they will be interested only in repeating information coming from nearby second class nodes. Those hosts, on the contrary, will also assume a computational function. The combination of ZigBee-Bluetooth protocols is likely to have a good attitude in semantic mobile grid applications although applications layer of both protocols must be modified in order to allow the use of semantics in annotation and discovery of services and resources. Furthermore a middleware for bridging two different application layer has to be implemented in order to allow a complete integration between mobile devices supporting either Bluetooth or ZigBee and a hybrid use of nodes within the grid. Both Zigbee and Bluetooth standards have a good versatility and present fair extension capabilities. We will exploit these features in order to overlap the traditional resource discovery of protocols with a semantic based micro-layer able to support data
dissemination as well as a semantic enabled service discovery. With reference to the second stage a fundamental question is the interaction user/grid. The system will have a common front-end interface where borderline nodes will be able to understand user requests also starting the complex discovery phase interfaced to the whole mobile grid. The grid structure must be able to grant reliability and a given Quality of Service (QoS). Hence minimum services have to be ensured, taking into account highly flexible and variable ubiquitous computing application scenario. Furthermore the ad-hoc infomobility context has to be semantically modelled. Grid services and resources have to be identified and characterized and sensor parameters (simple or combined among them) will be used as contextual values able to correct the semantic similarity criteria. The last phase is in identification of inference services useful for the application requirements and in implementation of utility functions taking into account the correlation existing between request/reply semantic based matchmaking procedures [1], numerical models involving field sensor data and contextual parameters. The scenario we envisage is one where a passenger, by means of her Bluetooth handheld device, will interact with the Mobile Semantic Grid requiring information about traffic, parking availability, possible lateness in public bus travels also planning her intermodal movements in real time. In addition she will be able to use further advanced services, we illustrate hereafter with an example scenario pictured in fig. 1.

Let us suppose a passenger is travelling towards downtown and willing to plan her evening. Via Bluetooth she may contact the micro-information server of the bus (vehicular hotspot) so interacting with the mobile grid. In this manner she can obtain, for example, the
theater programming in downtown or verify the opening hours of the downtown shops or finally she can submit to the system a complex request for “a premiere of an american action movie possibly showed in a Dolby surround downtown cinema”. The system shall be able to reply to this complex request correctly combining traffic data coming from grid sensor nodes with semantically annotated description of offered services. It will show results best matching the user request also taking into account an estimation of the arrival time (according to bus speed and traffic conditions – given the destination address). Furthermore the system shall also permit e-ticketing operations allowing to purchase a movie ticket on the bus.

5. Conclusions

In this paper we briefly introduced our vision of the Mobile Semantic Grid we plan to develop and implement in the framework of PITAGORA project. The objective is manifold: investigate the integration of protocols and technologies that are emerging to make them usable in practice, study user oriented semantic services and, provide a value-added infrastructure and a test-bed where new technologies and services can be experimented in a real environment.

References


