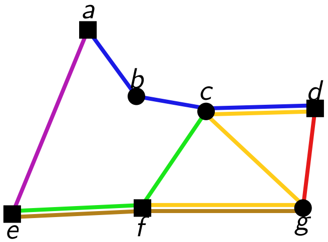
Approximation and Min-Max Results for the Steiner Connectivity Problem

Marika Karbstein
joint work with Ralf Borndörfer

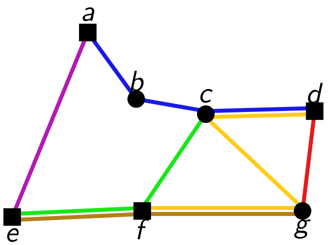
Zuse Institute Berlin

DFG Research Center MATHEON
Mathematics for key technologies

- undirected graph $G = (V, E)$
- set of terminal nodes $T \subseteq V$
- set of (simple) paths $\mathcal{P}$
- nonnegative cost $c \in \mathbb{R}_+^{\mathcal{P}}$

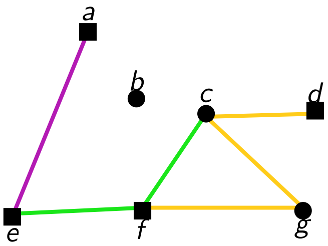


- undirected graph $G = (V, E)$
- set of terminal nodes $T \subseteq V$
- set of (simple) paths $\mathcal{P}$
- nonnegative cost $c \in \mathbb{R}_+^{\mathcal{P}}$



Steiner connectivity problem (SCP)

Find a set of paths $\mathcal{P}' \subseteq \mathcal{P}$ such that $c(\mathcal{P}') := \sum_{p \in \mathcal{P}'} c_p$ is minimal and all terminals T are connected.



Line Planning

- Input: public transport network, demands (OD-Matrix), operating cost, travel times
- Output: lines in network and frequencies s.t. demand is satisfied
- Objective: minimize operating cost, travel time, number transfers

Line Planning

- Input: public transport network, demands (OD-Matrix), operating cost, travel times
- Output: lines in network and frequencies s.t. demand is satisfied
- Objective: minimize operating cost, travel time, number transfers

SCP in line planning (one case study)

- ignore travel times
- assume "unlimited" capacities
- connect OD-nodes by choosing a set of lines with minimal cost

Outline

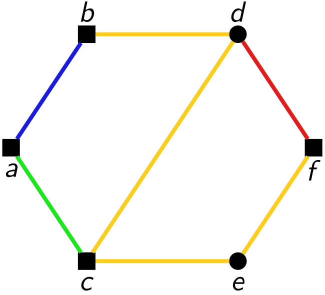
1. Relation to Steiner Trees and $|T| = \text{constant}$
2. Relation to Set Cover Problems and $T = V$
3. Primal Dual Approximation Algorithm

1. Relation to Steiner Trees and $|T| = \text{constant}$
2. Relation to Set Cover Problems and $T = V$
3. Primal Dual Approximation Algorithm

The SCP is a generalization of the (undirected) Steiner tree problem (STP)
($|p| = 1$ for all $p \in \mathcal{P}$).

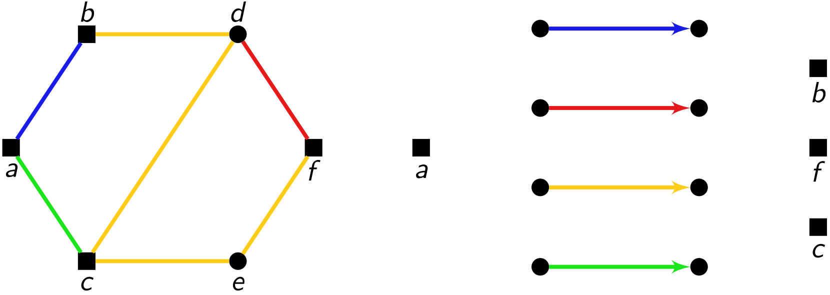
The SCP is a generalization of the (undirected) Steiner tree problem (STP)
($|p| = 1$ for all $p \in \mathcal{P}$).

The SCP can be transformed to the directed Steiner tree problem (DSTP):



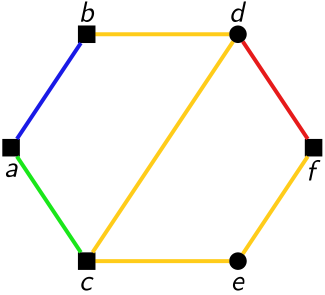
The SCP is a generalization of the (undirected) Steiner tree problem (STP)
($|p| = 1$ for all $p \in \mathcal{P}$).

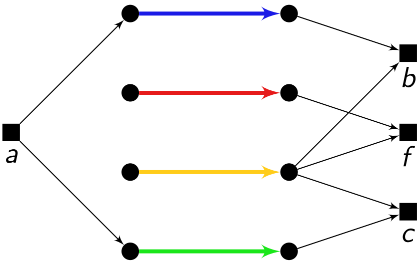
The SCP can be transformed to the directed Steiner tree problem (DSTP):



The SCP is a generalization of the (undirected) Steiner tree problem (STP)
($|p| = 1$ for all $p \in \mathcal{P}$).

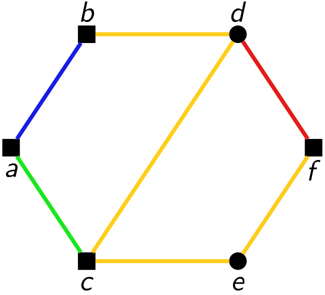
The SCP can be transformed to the directed Steiner tree problem (DSTP):

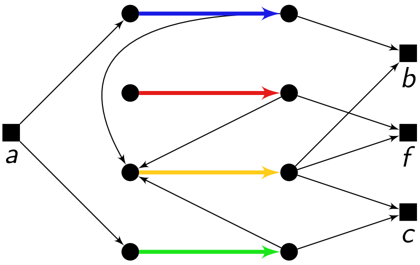




The SCP is a generalization of the (undirected) Steiner tree problem (STP)
($|p| = 1$ for all $p \in \mathcal{P}$).

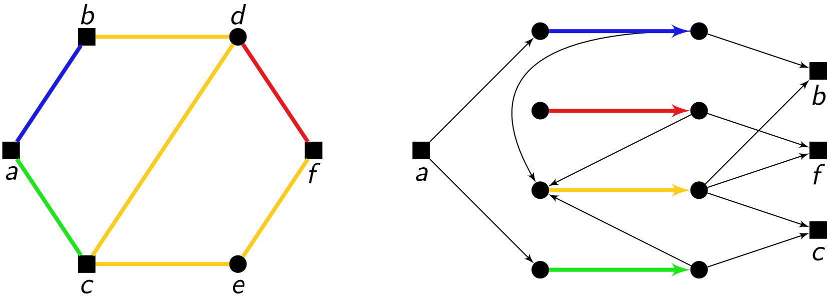
The SCP can be transformed to the directed Steiner tree problem (DSTP):



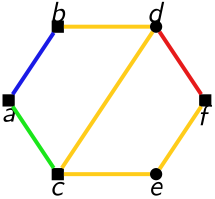


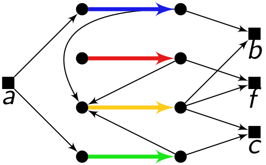
The SCP is a generalization of the (undirected) Steiner tree problem (STP)
($|p| = 1$ for all $p \in \mathcal{P}$).

The SCP can be transformed to the directed Steiner tree problem (DSTP):



all "path"-arcs receive cost of the corresponding path
all other arcs receive cost 0





Proposition

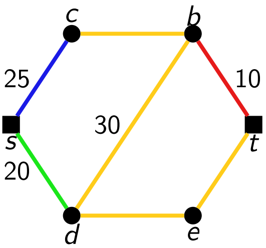
For each solution of one problem exists a solution of the other problem with the same objective value. The optimal objective value is independent of the choice of the root node.

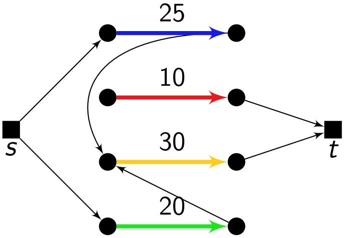
Corollary

SCP is solvable in polynomial time for $|T| = k$, k constant.

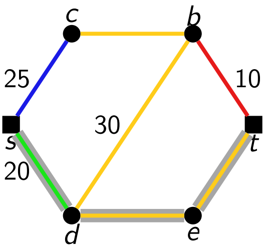
This follows from the complexity results for the directed Steiner tree problem, e.g., Feldman and Ruhl (1999)

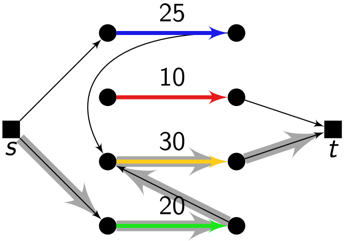
An *st*-connecting set of paths connects two nodes s and t.
An *st*-disconnecting set of paths breaks all *st*-connecting sets.





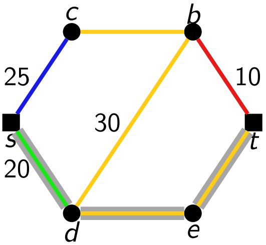
An st-connecting set of paths connects two nodes s and t.
An st-disconnecting set of paths breaks all st-connecting sets.

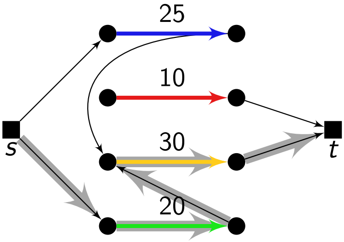




$\rightsquigarrow$ directed shortest path problem in associated directed graph

An *st*-connecting set of paths connects two nodes *s* and *t*.
An *st*-disconnecting set of paths breaks all *st*-connecting sets.





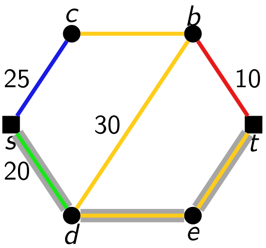
$\rightsquigarrow$ directed shortest path problem in associated directed graph

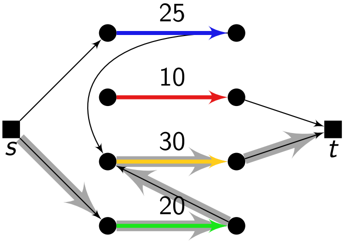
Proposition (Version of Menger's theorem)

The minimum cardinality of an st-disconnecting set is equal to the maximum number of path-disjoint st-connecting sets.

Follows from hypergraph theory.

An *st*-connecting set of paths connects two nodes s and t.
An *st*-disconnecting set of paths breaks all *st*-connecting sets.





$\rightsquigarrow$ directed shortest path problem in associated directed graph

Proposition (Max-flow-min-cut theorem w.r.t. paths)

The minimum weight of an st-disconnecting set is equal to the maximum flow w.r.t. paths.

Follows from general max-flow-min-cut theorem (Hoffman).

$|T| = 2$: problem can also be solved in original graph via an adapted shortest path algorithm
Advantages:

- no transformation (directed graph can have $O(|\mathcal{P}|^2)$ arcs)
- better complexity
- extended to a primal dual algorithm it can be used to prove the companion theorem to Menger

Proposition (companion to Menger's theorem)

The minimum cardinality of an st-connecting set is equal to the maximum number of path-disjoint st-disconnecting sets.

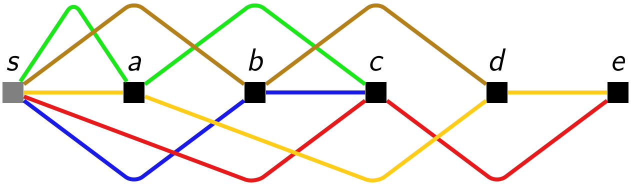
seems to be natural (for hypergraphs) but found no proof

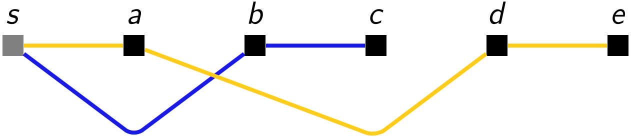
Proposition

The Steiner connectivity problem is NP-hard for $T = V$.

Reduction from set covering problem.

$S = \{a, b, c, d, e\}, \quad (\{a, c\}, \{b, d\}, \{b, c\}, \{c, e\}, \{a, d, e\})$





Let $N = \{1, \ldots, n\}$ and $z : 2^N \to \mathbb{R}$ be a nondecreasing, submodular function. Then

$$\min_{S \subseteq N} \{\sum_{j \in S} c_j \ : \ z(S) = z(N)\}$$

is a *submodular set covering problem*. It is integer-valued if $z : 2^N \to \mathbb{Z}$.

Observation

The SCP with $T = V$ can be interpreted as an integer-valued submodular set covering problem. Here, $z(\mathcal{P}')$, $\mathcal{P}' \subseteq \mathcal{P}$, is the maximum number of edges in $(V, E(\mathcal{P}'))$ containing no cycle. ($z(p) = |p|$, $p \in \mathcal{P}$, $z(\mathcal{P}) = |V| - 1$, $z(\emptyset) = 0$)

Theorem (Wolsey, 1982)

A greedy heuristic gives an $H(k) = \sum_{i=1}^{k} \frac{1}{i}$ *approximation guarantee for integer-valued submodular set covering problems where* $k = \max_{j \in N} z(\{j\}) - z(\emptyset)$.

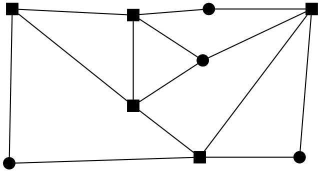
Corollary

A greedy heuristic gives an $H(k) = \sum_{i=1}^{k} \frac{1}{i}$ *approximation guarantee for the Steiner connectivity problem where* k *is the maximum number of edges in a path.*

This logarithmic bound is asymptotically optimal (Feige, 1998).

General approximation technique of Goemans and Williamson (1995)

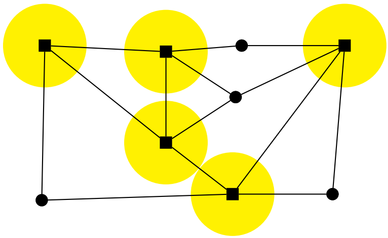
Example for the Steiner tree problem



- init: Steiner tree $F = \emptyset$
  $y_S = 0$, $S \subseteq V$,
  $\emptyset \neq S \cap T \neq T$

General approximation technique of Goemans and Williamson (1995)

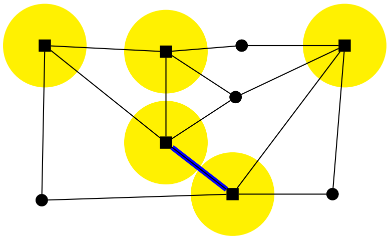
Example for the Steiner tree problem



- init: Steiner tree $F = \emptyset$
  $y_S = 0$, $S \subseteq V$,
  $\emptyset \neq S \cap T \neq T$
- raise y_S (S minimal with $\delta(S) \cap F = \emptyset$) uniformly until some edge e goes "tight" $F = F \cup \{e\}$

General approximation technique of Goemans and Williamson (1995)

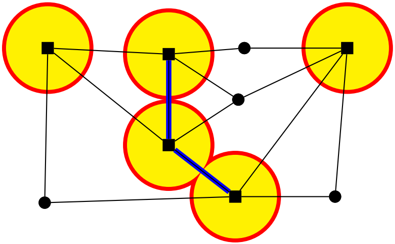
Example for the Steiner tree problem



- init: Steiner tree $F = \emptyset$
  $y_S = 0$, $S \subseteq V$,
  $\emptyset \neq S \cap T \neq T$
- raise y_S (S minimal with $\delta(S) \cap F = \emptyset$) uniformly until some edge e goes "tight" $F = F \cup \{e\}$

General approximation technique of Goemans and Williamson (1995)

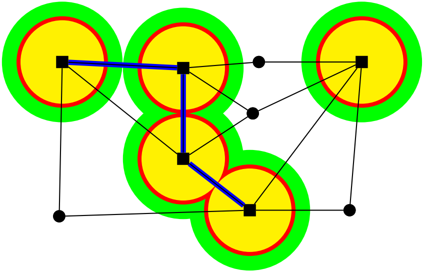
Example for the Steiner tree problem



- init: Steiner tree $F = \emptyset$
  $y_S = 0$, $S \subseteq V$,
  $\emptyset \neq S \cap T \neq T$
- raise y_S (S minimal with $\delta(S) \cap F = \emptyset$) uniformly until some edge e goes "tight" $F = F \cup \{e\}$

General approximation technique of Goemans and Williamson (1995)

Example for the Steiner tree problem

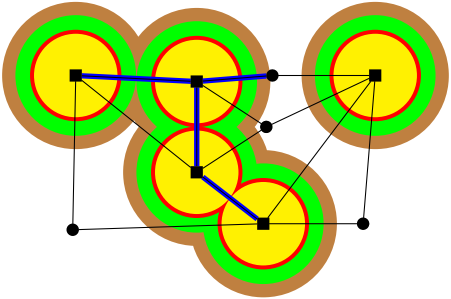


- init: Steiner tree $F = \emptyset$
  $y_S = 0$, $S \subseteq V$,
  $\emptyset \neq S \cap T \neq T$
- raise y_S (S minimal with $\delta(S) \cap F = \emptyset$) uniformly until some edge e goes "tight" $F = F \cup \{e\}$

General Approximation Technique – STP

General approximation technique of Goemans and Williamson (1995)

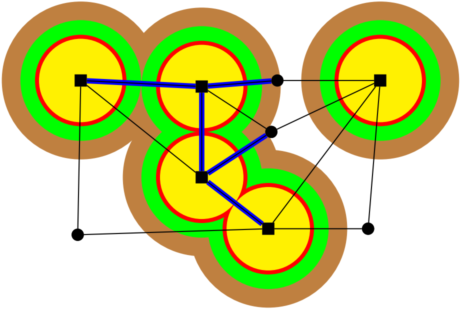
Example for the Steiner tree problem



- init: Steiner tree $F = \emptyset$
  $y_S = 0$, $S \subseteq V$,
  $\emptyset \neq S \cap T \neq T$
- raise y_S (S minimal with $\delta(S) \cap F = \emptyset$) uniformly until some edge e goes "tight" $F = F \cup \{e\}$

General approximation technique of Goemans and Williamson (1995)

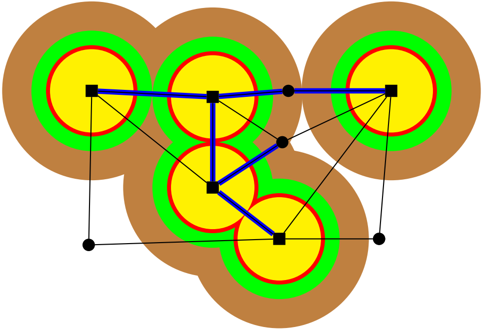
Example for the Steiner tree problem



- init: Steiner tree $F = \emptyset$
  $y_S = 0$, $S \subseteq V$,
  $\emptyset \neq S \cap T \neq T$
- raise y_S (S minimal with $\delta(S) \cap F = \emptyset$) uniformly until some edge e goes "tight" $F = F \cup \{e\}$

General approximation technique of Goemans and Williamson (1995)

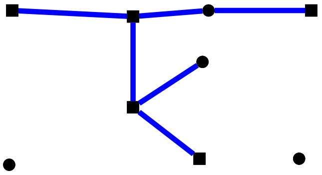
Example for the Steiner tree problem



- init: Steiner tree $F = \emptyset$
  $y_S = 0$, $S \subseteq V$,
  $\emptyset \neq S \cap T \neq T$
- raise y_S (S minimal with $\delta(S) \cap F = \emptyset$) uniformly until some edge e goes "tight" $F = F \cup \{e\}$

General approximation technique of Goemans and Williamson (1995)

Example for the Steiner tree problem

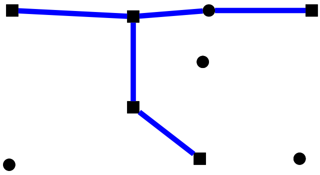


- init: Steiner tree $F = \emptyset$ $y_S = 0$, $S \subseteq V$, $\emptyset \neq S \cap T \neq T$
- raise y_S (S minimal with $\delta(S) \cap F = \emptyset$) uniformly until some edge e goes "tight" $F = F \cup \{e\}$
- reverse delete: consider edges in F in reverse order: delete e if $F = F \setminus \{e\}$ is feasible

General Approximation Technique – STP

General approximation technique of Goemans and Williamson (1995)

Example for the Steiner tree problem



- init: Steiner tree $F = \emptyset$
  $y_S = 0$, $S \subseteq V$,
  $\emptyset \neq S \cap T \neq T$
- raise y_S (S minimal with $\delta(S) \cap F = \emptyset$) uniformly until some edge e goes "tight" $F = F \cup \{e\}$
- reverse delete: consider edges in F in reverse order: delete e if $F = F \setminus \{e\}$ is feasible

Primal Dual Programs – STP

Let $\mathcal{S} = \{S \subseteq V : \emptyset \neq S \cap T \neq T\}$

$$\begin{array}{llll} \min & \sum_{e \in E} c_e x_e & & \\ \text{s.t.} & \sum_{e \in \delta(S)} x_e & \geq 1 & \forall S \in \mathcal{S} \\ & x_e & \geq 0 & \forall e \in E \end{array}$$

$$\begin{array}{llll} \max & \sum_{S \in \mathcal{S}} y_S & & \\ \text{s.t.} & \sum_{S \in \mathcal{S} : e \in \delta(S)} y_S & \leq c_e & \forall e \in E \\ & y_S & \geq 0 & \forall S \in \mathcal{S}. \end{array}$$

Proposition (Goemans and Williamson (1995))

The general approximation technique of Goemans and Williamson yields a 2-approximation algorithm for the Steiner tree problem.

Proof:

$$\sum_{e \in F} c_e = \sum_{e \in F} \sum_{S \in \mathcal{S} : e \in \delta(S)} y_S = \sum_{S \in \mathcal{S}} \sum_{e \in \delta(S) \cap F} y_S = \sum_{S \in \mathcal{S}} \deg_F(S) y_S \leq 2 \sum_{S \in \mathcal{S}} y_S$$

The last inequality follows since the average degree on a terminal node in a Steiner tree is 2.

Let $\mathcal{S} = \{S \subseteq V : \emptyset \neq S \cap T \neq T\}$

$$\begin{array}{llll} \min & \sum_{p\in\mathcal{P}} c_p\, x_p & & \\ \text{s.t.} & \sum_{p\in\mathcal{P}_{\delta(S)}} x_p & \geq 1 & \forall\, S \in \mathcal{S} \\ & x_p & \geq 0 & \forall\, p \in \mathcal{P} \end{array} \qquad \begin{array}{llll} \max & \sum_{S\in\mathcal{S}} y_S & & \\ \text{s.t.} & \sum_{S\in\mathcal{S}: p\in\mathcal{P}_{\delta(S)}} y_S & \leq c_p & \forall\, p \in \mathcal{P} \\ & y_S & \geq 0 & \forall\, S \in \mathcal{S}. \end{array}$$

Proposition

The general approximation technique of Goemans and Williamson applied to the SCP yields a (k+1)-approximation algorithm with k being the minimum of

(a) the maximal number of edges in a path,

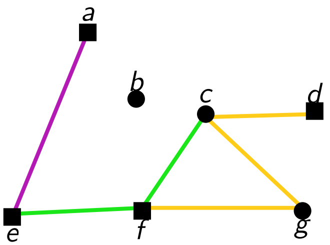
(b) the maximal number of terminal nodes in a path.

(a) the maximal number of edges in a path is k

Fujito (1999) observed: idea of proof for STP can be generalized to get a $(k+1)$ appr. for hypergraphs

$$\sum_{p \in \mathcal{P}'} c_p = \sum_{p \in \mathcal{P}'} \sum_{\substack{S \in \mathcal{S} \\ p \in \delta(S)}} y_S = \sum_{S \in \mathcal{S}} \sum_{p \in \delta(S) \cap \mathcal{P}'} y_S = \sum_{S \in \mathcal{S}} \deg_{\mathcal{P}'}(S) y_S \overset{(!)}{\leq} (k+1) \sum_{S \in \mathcal{S}} y_S$$

Assumption: the average degree of a terminal node in a (inclusion wise minimal) solution for the SCP is at most $k+1$
$\Leftrightarrow$ the average degree of a node $t \in T$ in a minimal T-connecting hypergraph is $k+1$

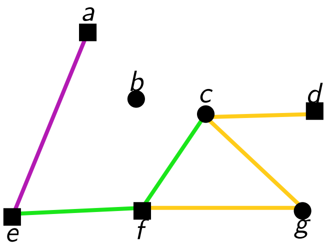


(a) the maximal number of edges in a path is k

Fujito (1999) observed: idea of proof for STP can be generalized to get a $(k+1)$ appr. for hypergraphs

$$\sum_{p\in\mathcal{P}'} c_p = \sum_{p\in\mathcal{P}'} \sum_{\substack{S\in\mathcal{S}\\ p\in\delta(S)}} y_S = \sum_{S\in\mathcal{S}} \sum_{p\in\delta(S)\cap\mathcal{P}'} y_S = \sum_{S\in\mathcal{S}} \deg_{\mathcal{P}'}(S) y_S \overset{(!)}{\leq} (k+1) \sum_{S\in\mathcal{S}} y_S$$

Assumption: the average degree of a terminal node in a (inclusion wise minimal) solution for the SCP is at most $k+1$
$\Leftrightarrow$ the average degree of a node $t \in T$ in a minimal T-connecting hypergraph is $k+1$ (seems to be natural but found no proof)



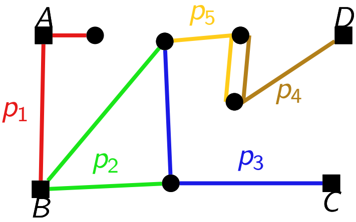
Lemma

The average degree of a terminal node in an inclusion wise minimal solution for the SCP is at most $(k+1)$ *with* k *being the minimum of*

(a) the maximal number of edges in a path,

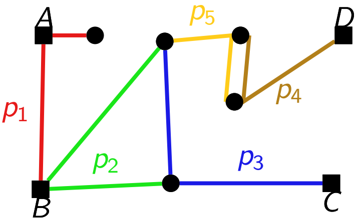
(b) the maximal number of terminal nodes in a path.

(Equivalent to:
The average degree of a node $t \in T$ *in a minimal* T*-connecting hypergraph is at most* $k+1$*.)*



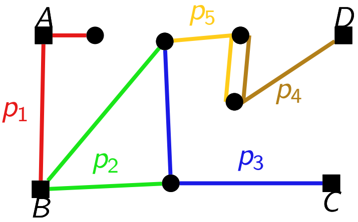
- $\{p_1, p_2, p_3\}\{p_4\}$

- $\mathcal{P}'$ minimal; consider only paths that intersect terminal nodes
  order paths according to connected components



- $\{p_1, p_2, p_3\}\{p_4\}$
- $T_1 = T_2 = \{A, B\}$
- $T_3 = \{A, B, C\}$
- $T_4 = \{A, B, C, D\}$

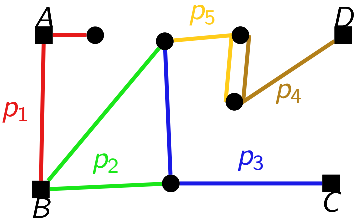
- $\mathcal{P}'$ minimal; consider only paths that intersect terminal nodes order paths according to connected components
- T_i set of terminal nodes that are covered by $p_1, \ldots, p_i$



- $\{p_1, p_2, p_3\}\{p_4\}$
- $T_1 = T_2 = \{A, B\}$
- $T_3 = \{A, B, C\}$
- $T_4 = \{A, B, C, D\}$

- $\mathcal{P}'$ minimal; consider only paths that intersect terminal nodes order paths according to connected components
- T_i set of terminal nodes that are covered by $p_1, \ldots, p_i$
- $r_i = |T_i \setminus T_{i-1}|$ additional terminal nodes covered by p_i

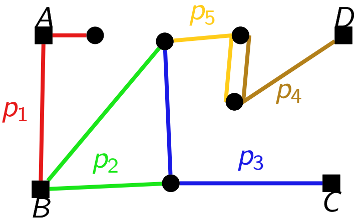
Average Degree – Proof



- $\{p_1, p_2, p_3\}\{p_4\}$
- $T_1 = T_2 = \{A, B\}$
- $T_3 = \{A, B, C\}$
- $T_4 = \{A, B, C, D\}$

- $\mathcal{P}'$ minimal; consider only paths that intersect terminal nodes
  order paths according to connected components
- T_i set of terminal nodes that are covered by $p_1, \ldots, p_i$
- $r_i = |T_i \setminus T_{i-1}|$ additional terminal nodes covered by p_i

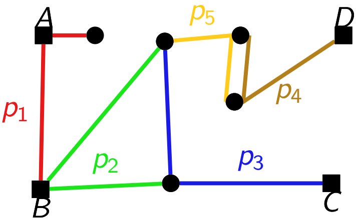
  $r_i \geq 1$: p_i increase degree on all terminal nodes by
  (a) $r_i + \min(|T_{i-1}| - 1, k + 1 - r_i)$ (path length $\leq k$)
  (b) $r_i + \min(|T_{i-1}| - 1, k - r_i)$ (path contains $\leq k$ terminals)



- $\{p_1, p_2, p_3\}\{p_4\}$
- $T_1 = T_2 = \{A, B\}$
- $T_3 = \{A, B, C\}$
- $T_4 = \{A, B, C, D\}$

- $\mathcal{P}'$ minimal; consider only paths that intersect terminal nodes
  order paths according to connected components
- T_i set of terminal nodes that are covered by $p_1, \ldots, p_i$
- $r_i = |T_i \setminus T_{i-1}|$ additional terminal nodes covered by p_i

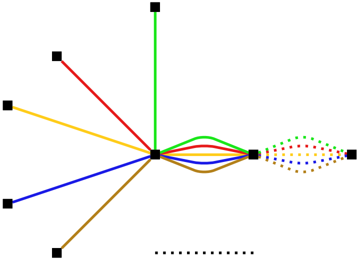
  $r_i \geq 1$: p_i increase degree on all terminal nodes by
  (a) $r_i + \min(|T_{i-1}| - 1, k + 1 - r_i)$ (path length $\leq k$)
  (b) $r_i + \min(|T_{i-1}| - 1, k - r_i)$ (path contains $\leq k$ terminals)
  $r_i = 0$: $\exists p_h, h > i$ with $V(p_i) \cap V(p_h) \neq \emptyset$, $T_i \cap V(p_h) = \emptyset$ and p_h adds
  $r_h \geq 1$ "new" terminals; move p_h at position $i + 1$
  p_i, p_h increase degree by $r_h + \min(|T_{i-1}| - 1, k)$

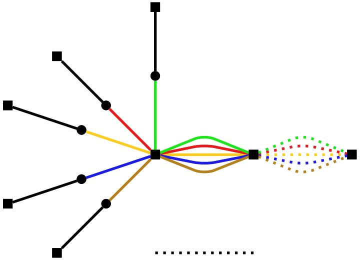


- $\{p_1, p_2, p_3\}\{p_4\}$
- $T_1 = T_2 = \{A, B\}$
- $T_3 = \{A, B, C\}$
- $T_4 = \{A, B, C, D\}$

- final order: set of paths and pairs of paths that increase degree on terminal nodes by at most $r_i + \min(|T_{i-1}| - 1, k)$

$$\begin{aligned}
&\sum_{t \in T} \deg_{\mathcal{P}'}(t) \\
&= \quad r_1 + r_2 + \min\{k, |T_1| - 1\} + \ldots + r_m + \min\{k, |T_{m-1}| - 1\} \\
&\leq \quad r_1 + r_2 + \min\{k, r_1 - 1\} + \ldots + r_m + \min\{k, (\textstyle\sum_{i=1}^{m-1} r_i) - 1\} \\
&\leq \quad |T| + (r_1 - 1) + \ldots + (r_1 + \ldots + r_{j-1} - 1) + (m - j)k \\
&\qquad \ldots \\
&\leq \quad (|T| - 1)(k + 1). \qquad \square
\end{aligned}$$





- n nodes in the rim, k nodes in the middle; all nodes are terminal nodes (plus n nodes in the inner rim)
- n paths: each path contains one node in the (inner) rim and all nodes in the middle
- all paths are minimal T-connecting set
- total degree: $n \cdot 1 + k \cdot n = n(k+1)$
- average degree: $\frac{n(k+1)}{n+k} \xrightarrow{n \to \infty} k+1$

Comparison: STP vs. SCP

	STP	SCP
general case	$\mathcal{NP}$-hard	
$\|T\| = k$	polynomial	
$\|T\| = 2$	version of Menger's and companion theorem holds	
$T = V$	polynomial	NP-hard
	min. spanning tree	Greedy: $H(k)$-appr. k max. number edges in paths
primal-dual alg.	2-approximation	$(k+1)$-approximation k minimum of (a) max. number edges/path, (b) max. number terminals/path

Polyhedral Aspects (Borndörfer, K., Pfetsch, 2012):

- Can generalize structures such as partition inequalities to the SCP.
- Can also obtain a directed formulation dominating natural formulation.