The Maximum Clique Interdiction Game

Fabio Furini, Ivana Ljubić, Sébastien Martin, Pablo San Segundo

Université Paris-Dauphine
The Maximum Clique Node-Interdiction Game (CIG)

- We study the two player zero-sum Stackelberg game in which the leader interdicts (removes) a limited number of vertices from a simple graph (interdiction budget), and the follower searches for the maximum clique in the interdicted graph.

- The goal of the leader is to derive an interdiction policy which will result in the worst possible outcome for the follower.

**Definition**

Given a graph $G$ and an interdiction budget $k (k \geq 1)$, the maximum clique interdiction game is to find a subset of at most $k$ nodes to delete from $G$ so that the size of the maximum clique in the remaining graph is minimized.

The set of interdicted nodes is called the optimal interdiction policy.
Example: $\omega(G) = 5$ and $k = 1$

Maximum Clique $\tilde{K} = \{v_3, v_4, v_7, v_8, v_9\}$

Optimal interdiction policy $\{v_8\}$
Example: $\omega(G) = 5$ and $k = 2, k = 3$

Optimal interdiction policy $\{v_4, v_8\}$

Optimal interdiction policy $\{v_4, v_7, v_8\}$
Motivation

- In the context of **terrorist networks** (Chen 04 and Sampson 89), **cliques** are used to model terrorist cells (tightly knit groups of people).
- In the context of **crime detection and prevention**, **large cliques** are potential origins of catastrophic events:
  - terrorist or hacker attacks (Berry 04 and Sageman 04)
  - sources of outbreaks of sexually transmitted diseases (Rothenberg 96).

For these reasons we study the problem on how to efficiently reduce the size of the largest clique of a network, given a predefined number of vertices that can be interdicted (**most vital clique nodes of a graph**).

- centralized control of Software Defined Networks (SDNs)
- In the context of **graph theory**, we can analyze the resilience of the graphs with respect to vertex-interdiction deletion (**Clique-Interdiction curve of a graph**).
Literature Overview

- No exact specialized algorithms for CIG exit in the literature

- CIG belongs to a larger family of Interdiction Games under Monotonicity (Fischetti et al. 16; focus on knapsack interdiction games).

- Games where the follower subproblem satisfies a monotonicity (or hereditary) property, exploited to derive a single-level integer linear programming formulation.

Related problems

- **Minimum Vertex Blocker Clique Problem** (Mahdavi Pajouh et al. 16), they tackle graphs with at most 200 vertices and most of the instances are unsolved

- **Edge Interdiction Clique Game** (Tang et al. 16), they tackle graphs with 15 vertices and most of the instances are unsolved

SPOILER: we can solve graph with 100k nodes and 3M edges!
**Complexity**

**Decision Version of CIG (d-CIG):** Given a graph $G$ and two integers $k$ and $\ell$, can we remove (at most) $k$ vertices from $G$ such that the resulting graph does not contain a clique of size $\ell$?

- Observe that the answer to the decision problem is YES if only if the optimal CIG solution is $\leq \ell - 1$.

- d-CIG is not in NP, to test whether the resulting graph does not contain a clique of size $\ell$ requires answering the decision version of:
  - the maximum clique problem (NP-complete).

- d-CIG has been also called **Generalized Node Deletion (GND)** problem.

**Proposition (Rutenburg1991, Rutenburg1994)**

The decision version of CIG is $\Sigma_2^P$-complete.
Single-Level ILP Reformulation

\[ w_u = \begin{cases} 
1, & \text{if vertex } u \text{ is interdicted by the leader,} \\
0, & \text{otherwise} 
\end{cases} \quad u \in V \]

\[ x_u = \begin{cases} 
1, & \text{if vertex } u \text{ is used in the maximum clique of the follower,} \\
0, & \text{otherwise} 
\end{cases} \quad u \in V \]

Let \( \mathcal{W} \) be the set of all feasible interdiction policies of the leader:

\[ \mathcal{W} = \left\{ w \in \{0, 1\}^n : \sum_{u \in V} w_u \leq k \right\} \quad (0.1) \]

Let \( \mathcal{K} \) be the set of incidence vectors of all cliques in the graph \( G \):

\[ \mathcal{K} = \left\{ x \in \{0, 1\}^n : x_u + x_v \leq 1, uv \in \bar{E} \right\} \quad (0.2) \]

Property

CIG can be restated as follows:

\[ \min_{w \in \mathcal{W}} \max_{K \in \mathcal{K}} \left\{ |K| - \sum_{u \in K} w_u \right\}. \quad (0.3) \]
Single-Level ILP Reformulation

For every feasible interdiction policy \( \tilde{w} \in \mathcal{W} \), the follower’s problem becomes:

\[
\max_{x \in \mathcal{K}} \left\{ \sum_{u \in V} x_u : x_u \leq 1 - \tilde{w}_u, \; u \in V \right\} = \max_{x \in \mathcal{K}} \sum_{u \in V} x_u(1 - \tilde{w}_u)
\]

- the set of feasible solutions of the follower does not depend on the actions of the leader anymore.
- One can enumerate all cliques in \( G \) and optimize over the set \( \mathcal{K} \).

Proposition

The following is a valid ILP formulation for CIG:

\[
\begin{align*}
\text{min} & \quad \theta \\
\theta + \sum_{u \in K} w_u & \geq |K| & K \in \mathcal{K} \\
\sum_{u \in V} w_u & \leq k \\
w_u & \in \{0, 1\} & u \in V.
\end{align*}
\]
Exact Solution Framework – CLIQUE-INTER

(i) Effective separation procedure of the Clique Interdiction (CI) cuts:

- Specialized combinatorial branch-and-bound algorithm (IMCQ) for solving the maximum clique problem once the nodes of an interdiction policy have been removed from the graph $G$.
- Make the separated cliques maximal

(ii) Tight CIG upper and lower bounds ($\ell_{\text{min}}$ and $\ell_{\text{max}}$):

- To initialize the lower bound value of the variable $\theta$ we used the global lower bound $\ell_{\text{min}}$ using node-disjoint maximum cliques
- To determine a high-quality feasible CIG solution of value $\ell_{\text{max}}$, we apply a battery of effective sequential greedy heuristics.

(iii) The graph Reduction Technique:

- For large-scale real-world graphs the ILP formulation unless the input graph can be safely reduced to a smaller one.
Facial study

- the following Proposition provides necessary and sufficient conditions under which the CI cuts are facet defining.
- major theoretical result! it allows to characterize the strength of the ILP formulation upon which our solution framework is built on.

**Theorem**

Let $K \in \mathcal{K}$ be a maximal clique. Inequality (0.5) associated with $K$ defines a facet of $P(G, k)$ if and only if

- $|K| \geq \ell_{\text{opt}} + 1$
- for all $v \in K$, there exists a subset $V' \subseteq V$ such that $v \in V'$, $|V'| \leq k$ and $\omega(G[V \setminus V']) + |V' \cap K| \leq |K|$.

It is NP-hard to down-lift coefficients of a clique interdiction cut

- Heuristic lifting procedure! by underestimating the left-hand-side and overestimating the right-hand-side of the condition.
Separating the Clique Interdiction Cuts with IMCQ

The separation problem requires solving the MCP in a number of induced subgraphs $G[V \setminus V_w]$, where $V_w$ is a feasible interdiction policy.

- We have designed a combinatorial branch-and-bound (B&B) algorithm inspired by the ideas described in (Li 17) and (San Segundo 16).

- Using tight lower based on the **infrachromatic bounding functions** (potentially stronger than the fractional chromatic number!)

- **Main Idea!** Given a valid lower bound on MCP of value $q$, we can partition $V$ into two disjoint sets of vertices $P$ and $B = V \setminus P$ such that $\omega(G[P]) \leq q$

  Branching is necessary on the vertices in $B$ only!

- **Plus!** Compact bitstring representation both for vertex sets and the adjacency matrix and peeling procedures.
Computing the global lower bound $\ell_{\text{min}}$

**Proposition**

*Given a subgraph $G' = (V, E')$ with $E' \subset E$, the optimal CIG solution on $G'$ provides a valid lower bound for the optimal CIG solution on $G$.***

- rather counter-intuitive! reducing the input graph, instead of obtaining a valid upper bound for a minimization problem, we obtain a valid lower bound (the feasibility space of the follower is reduced)*

**Corollary**

*Given a set $Q_{p+1} = (K_1, \ldots, K_{p+1})$ of vertex-disjoint cliques of $G$, such that $|K_1| \geq \cdots \geq |K_{p+1}|$, a valid lower bound $\ell_{\text{min}}$ for the CIG can be obtained by computing*

\[
\ell_{\text{min}} = \begin{cases} 
\max \left\{ |K_{p+1}|, |K_p| - 1 - \left\lfloor \frac{k-k(Q_p)}{p} \right\rfloor \right\}, & \text{if } k < k(Q_p^{*+1}) \\
|K_{p+1}| - 1 - \left\lfloor \frac{k-k(Q_{p+1})}{p+1} \right\rfloor, & \text{otherwise}
\end{cases}
\]

(0.8)

Where $k(Q_q)$ denote the size of an optimal interdiction policy necessary to reduce the size of all cliques in $Q_q$ to $|K_q| - 1$.

\[
k(Q_q) = q + \sum_{i=1}^{q-1} i \cdot (|K_i| - |K_{i+1}|).
\]
Reducing the input graph

- The **clique number** of \( v \) is the size of the largest clique with \( v \) (\( \omega_G(v) \)).

- The **coreness-number** of a vertex \( v \), is equal to \( \kappa \) if \( v \) belongs to a \( \kappa \)-core but not to any \((\kappa + 1)\)-core.

\[
\omega_G(v) \leq \text{coreness}(v) + 1 \leq |N(v)| + 1 \quad v \in V. \tag{0.9}
\]

The following result identifies redundant vertices in the input graph \( G \)

**Proposition**

*Let \( v \) be an arbitrary vertex from \( V \). If \( \omega_G(v) \leq \ell_{opt} \), then \( v \) cannot be part of a minimal optimal interdiction policy.*

- instead of using the (unknown) value of \( \ell_{opt} \), we use the lower bound \( \ell_{\text{min}} \)
- instead of using \( \omega_G(v) \) (NP-hard), we use \( \text{coreness}(v) + 1 \) (polynomial)
Test-bed Instances

- Set A – Random Erdős-Rényi random $G(n, p)$ – 220 instances:
  - $n = |V| \in \{50, 75, 100, 125, 150\}$
  - $p \in \{0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 0.95, 0.98\}$
  - $k \in \{\lceil 0.05 \cdot |V| \rceil, \lceil 0.1 \cdot |V| \rceil, \lceil 0.2 \cdot |V| \rceil, \lceil 0.4 \cdot |V| \rceil\}$

- Set B – Synthetic graphs – 32 instances:
  - Instances with $|V| = 200$ from the 2nd DIMACS challenge on Maximum Clique, Graph Coloring, and Satisfiability;
  - $k \in \{20, 40\}$

- Set C – Real-world (sparse) networks – 60 instances.
  - instances with up to $\approx 100,000$ nodes and $\approx 3,200,000$ edges.
  - $k \in \{\lceil 0.005 \cdot |V| \rceil, \lceil 0.01 \cdot |V| \rceil\}$
## Comparison with state-of-the-art generic bilevel solver (BILEVEL)

| $|V|$ | # | CLIQUE-INTER | BILEVEL |
| --- | --- | --- | --- | --- | --- | --- | --- |
|   |   | # solved | time | exit gap | root gap | # solved | time | exit gap | root gap |
| 50 | 44 | 44 | 0.01 | - | 0.16 | 28 | 68.58 | 6.44 | 8.50 |
| 75 | 44 | 44 | 1.45 | - | 0.41 | 14 | 120.19 | 9.47 | 10.91 |
| 100 | 44 | 37 | 9.30 | 1.00 | 0.98 | 7 | 164.42 | 12.65 | 13.11 |
| 125 | 44 | 35 | 13.43 | 1.33 | 1.20 | 2 | 135.33 | 13.88 | 14.73 |
| 150 | 44 | 33 | 27.23 | 1.91 | 1.43 | 1 | 397.52 | 16.42 | 16.39 |

## Results on Real-world (sparse) networks

| Network         | \(|V|\)    | \(|E|\)     | \(\omega\) [s] | \(k = 0.005 \cdot |V|\) [s] | \(|V_p|\)  | \(k = 0.01 \cdot |V|\) [s] | \(|V_p|\)  |
|-----------------|-----------|-------------|----------------|----------------|--------|----------------|--------|
| socfb-UlIllinois | 30,795    | 1,264,421   | 0.5            | 24.4           | 10,456 | 41.6           | 8290   |
| ia-email-EU     | 32,430    | 54,397      | 0.0            | 0.6            | 30,375 | 0.5            | 29,212 |
| rgg_n_2_15_s0   | 32,768    | 160,240     | 0.0            | 2.2            | 27,791 | 0.2            | 30,848 |
| ia-enron-large  | 33,696    | 180,811     | 0.0            | 2.2            | 27,791 | 29.5           | 26,651 |
| socfb-Uf        | 35,111    | 1,465,654   | 0.3            | 17.8           | 14,264 | 87.8           | 10,708 |
| socfb-Texas84   | 36,364    | 1,590,651   | 0.3            | 24.6           | 10,706 | 74.3           | 8,704  |
| tech-internet-as| 40,164    | 85,123      | 0.0            | 1.4            | 31,783 | -              | -      |
| fe-body         | 45,087    | 163,734     | 0.1            | 1.8            | 2,259  | 1.8            | 2259   |
| sc-nasasrb      | 54,870    | 1,311,227   | 0.1            | -              | -      | 145.5          | 1,195  |
| soc-themarker_u | 69,413    | 1,644,843   | 2.1            | T.L.           | 35,678 | T.L.           | 31,101 |
| rec-eachmovie_u | 74,424    | 1,634,743   | 0.7            | -              | -      | 367.3          | 13669  |
| fe-tooth        | 78,136    | 452,591     | 0.5            | 18.9           | 7      | 19.0           | 7      |
| sc-pkustk11     | 87,804    | 2,565,054   | 1.1            | 70.7           | 2,712  | 57.1           | 2,712  |
| soc-BlogCatalog | 88,784    | 2,093,195   | 11.7           | T.L.           | 51,607 | T.L.           | 46,240 |
| ia-wiki-Talk    | 92,117    | 360,767     | 0.2            | 49.2           | 72,678 | 87.4           | 72,678 |
| sc-pkustk13     | 94,893    | 3,260,967   | 1.3            | 724.9          | 2,360  | 879.2          | 2,354  |
Clique Interdiction Game Structural Properties, Modeling and Exact Algorithms

Clique-Interdiction curve of a graph

OPT / omega
k [%] 
brock200_2
brock200_3
brock200_4
c-fat200-1
c-fat200-2
c-fat200-5
san200_0.7_1
san200_0.7_2
san200_0.9_1
san200_0.9_9
Clique Interdiction Game Structural Properties, Modeling and Exact Algorithms Computational Results

Clique-Interdiction curve of a graph

OPT / omega vs k [%]

- netscience
- power
- hep-th
- PGPgiantcompo
- astro-ph
- cond-mat
- memplus
- as-22july06
- cond-mat-2003
- cond-mat-2005
Conclusions

- We developed the first study on how to find the **most vital** $k$ **vertices of a graph**, so as to reduce its **clique number**

- We derived tight combinatorial lower and upper bounds

- We derive a single-level reformulation based on an exponential family of **Clique-Interdiction Cuts**

- We provide necessary and sufficient conditions under which these cuts are facet defining and we propose a fast lifting procedures

- We developed a **state-of-the-art algorithm for finding maximum cliques in interdicted graphs**

- Social Networks are “vulnerable” to vertex-deletion attacks!

THANKS FOR YOUR ATTENTION!!!!
Convex hull of feasible solutions of the CIG formulation (0.4)-(0.7)

$$
\mathcal{P}(G, k) = \text{conv} \left\{ w \in \{0, 1\}^{|V|}, \theta \geq 0 : \theta + \sum_{u \in K} w_u \geq |K|, \sum_{u \in V} w_u \leq k, K \in \mathcal{K} \right\}.
$$

Proposition

The polytope $\mathcal{P}(G, k)$ is full dimensional.

Proposition

Let $u \in V$. The trivial inequality $w_u \leq 1$ defines a facet of $\mathcal{P}(G, k)$ if and only if $k \geq 2$.

Proposition

Let $u \in V$. The trivial inequality $w_u \geq 0$ defines a facet of $\mathcal{P}(G, k)$.

Lemma

Let $K \in \mathcal{K}$ be an arbitrary clique in $G$. If $|K| \leq \ell_{\text{opt}}$, then the associated clique interdiction inequality (0.5) cannot define a facet.
Facial study

**Lemma**
Let $K \in \mathcal{K}$ be an arbitrary clique in $G$. The inequality $\theta + \sum_{u \in K} w_u \geq |K|$ defines a facet only if $K$ is maximal.

**Lemma**
Let $K$ be a maximal clique and $v \in K$. If

$$\omega(G[V \setminus V']) \geq |K| - |V' \cap K| + 1 \quad \forall V' \subseteq V \text{ where } v \in V' \text{ and } |V'| \leq k, \quad (0.1)$$

then there exists $\alpha_v \leq 0$ such that the associated clique interdiction cut (0.5) can be down-lifted to

$$\theta + \sum_{u \in K \setminus \{v\}} w_u + \alpha_v w_v \geq |K|.$$

**Corollary**
Let $K \subset V$ be a clique. If there exists $v \in K$ satisfying (0.1) then the inequality (0.5) cannot define a facet.
Performance Profile – Set A

1. CLIQUE-INTER: this is the benchmark setting of our exact algorithm, fully exploiting all its components.

2. CLIQUE-INTER (no bounds): in this configuration we remove the use of CIG upper and lower bounds ($\ell_{\text{min}}$ and $\ell_{\text{max}}$).

3. CLIQUE-INTER (no maximality): in this configuration we did not make maximal the cliques separated using IMCQ before adding the corresponding CIC.

4. Basic CLIQUE-INTER with IMCQ: in this configuration all components are removed, except the use of IMCQ to separate CICs.

5. Basic CLIQUE-INTER with CPLEX: this configuration corresponds to the basic branch-and-cut approach in which CICs are separated using CPLEX as a black-box clique solver applied to the classical clique ILP formulation.
Performance Profile – Set A

\[ \text{CLIQUE-INTER} \]
\[ \text{CLIQUE-INTER (no bounds)} \]
\[ \text{CLIQUE-INTER (no maximality)} \]
\[ \text{Basic CLIQUE-INTER with IMCQ} \]
\[ \text{Basic CLIQUE-INTER with CPLEX} \]
CPU times group by the graph density

![Graph showing CPU times group by the graph density](image-url)
CPU times group by the size and the interdiction budget

<table>
<thead>
<tr>
<th></th>
<th>#OPT</th>
<th>#PREP</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0</td>
</tr>
</tbody>
</table>
### Table 1: Computational results obtained by the CLIQUE-INTER on the instances with $|V| = 200$ from the 2nd DIMACS Challenge.

<table>
<thead>
<tr>
<th>Instance</th>
<th>$\mu$</th>
<th>$\omega(G)$</th>
<th>time $\omega$</th>
<th>CLIQUE-INTER $k = 20$</th>
<th>CLIQUE-INTER $k = 40$</th>
</tr>
</thead>
<tbody>
<tr>
<td>brock200_1</td>
<td>0.75</td>
<td>21</td>
<td>0.2</td>
<td>LB UB</td>
<td>time $\ell_{\min}$ $\ell_{\max}$</td>
</tr>
<tr>
<td>brock200_2</td>
<td>0.50</td>
<td>12</td>
<td>0.0</td>
<td>9 9</td>
<td>0.1</td>
</tr>
<tr>
<td>brock200_3</td>
<td>0.61</td>
<td>15</td>
<td>0.0</td>
<td>12 12</td>
<td>1.0</td>
</tr>
<tr>
<td>brock200_4</td>
<td>0.66</td>
<td>17</td>
<td>0.0</td>
<td>14 14</td>
<td>2421.8</td>
</tr>
<tr>
<td>c-fat200-1</td>
<td>0.08</td>
<td>12</td>
<td>0.0</td>
<td>10 10</td>
<td>-</td>
</tr>
<tr>
<td>c-fat200-2</td>
<td>0.16</td>
<td>24</td>
<td>0.0</td>
<td>20 20</td>
<td>-</td>
</tr>
<tr>
<td>c-fat200-5</td>
<td>0.43</td>
<td>58</td>
<td>0.0</td>
<td>52 52</td>
<td>0.0</td>
</tr>
<tr>
<td>san200_0.7_1</td>
<td>0.70</td>
<td>30</td>
<td>0.0</td>
<td>17 17</td>
<td>5.4</td>
</tr>
<tr>
<td>san200_0.7_2</td>
<td>0.70</td>
<td>18</td>
<td>0.0</td>
<td>14 14</td>
<td>16.7</td>
</tr>
<tr>
<td>san200_0.9_1</td>
<td>0.90</td>
<td>70</td>
<td>0.0</td>
<td>50 50</td>
<td>-</td>
</tr>
<tr>
<td>san200_0.9_2</td>
<td>0.90</td>
<td>60</td>
<td>0.1</td>
<td>41 41</td>
<td>3.2</td>
</tr>
<tr>
<td>san200_0.9_3</td>
<td>0.90</td>
<td>44</td>
<td>0.0</td>
<td>33 34</td>
<td>T.L.</td>
</tr>
<tr>
<td>sanr200_0.7</td>
<td>0.70</td>
<td>18</td>
<td>0.1</td>
<td>15 15</td>
<td>29.2</td>
</tr>
<tr>
<td>sanr200_0.9</td>
<td>0.90</td>
<td>42</td>
<td>1.9</td>
<td>33 35</td>
<td>T.L.</td>
</tr>
<tr>
<td>gen200_p0.9_44</td>
<td>0.90</td>
<td>44</td>
<td>0.1</td>
<td>34 34</td>
<td>674.4</td>
</tr>
<tr>
<td>gen200_p0.9_55</td>
<td>0.90</td>
<td>55</td>
<td>0.1</td>
<td>38 38</td>
<td>62.4</td>
</tr>
</tbody>
</table>