From Game Theory to Graph Theory: A Bilevel Journey

Ivana LJUBIC

ESSEC Business School, Paris

OR 2018, Brussels

EURO Plenary

Stackelberg Games

- Two-player sequential-play game: LEADER and FOLLOWER
- LEADER moves before FOLLOWER first mover advantage
- Perfect information: both agents have perfect knowledge of each others strategy
- Rationality: agents act optimally, according to their respective goals
- LEADER takes FOLLOWERS's optimal response into account
- Optimistic vs Pessimistic: when FOLLOWER has multiple optimal responses







• Optimistic vs Pessimistic: when FOLLOWER has multiple optimal responses

Stackelberg Games

- Introduced in economy by v. Stackelberg in 1934
- 40 years later introduced in Mathematical
 Optimization → Bilevel Optimization

A Convex Programming Model for Optimizing SLBM Attack of Bomber Bases

Jerome Bracken and James T. McGill

Institute for Defense Analyses, Arlington, Virginia

(Received July 30, 1970)

This paper formulates a convex programming model allocating submarinelaunched ballistic missiles (SLBMs) to launch areas and providing simultaneously an optimal targeting pattern against a specified set of bomber bases. Flight times of missiles from launch areas to bases vary and targets decrease in value over time. A nonseparable concave objective function is given for expected destruction of bombers. An example is presented.



Applications: Pricing

Two competitive agents act in a hierarchical way with different/conflicting objectives

- **Pricing:** operator sets tariffs, and then customers choose the cheapest alternative
- Tariff-setting, toll optimization (Labbé et al., 1998; Brotcorne et al., 2001)
- Network Design and Pricing (Brotcorne et al., 2008)
- Survey (van Hoesel, 2008)





Applications: Interdiction

Canada and the Transcontinental Drug Links Strategic Forecasting Inc go to original Email Page Print Page Email Us

Canadian police conducted several simultaneous raids on suspected drug traffickers in Newfoundland and Quebec provinces Oct. 11, arresting two dozen people and seizing marijuana, cocaine, weapons, cash and property. The drugtrafficking ring, which Canadian authorities believe was operated by the Quebec-based Hell's Angels motorcycle/crime gang, could have smuggled the cocaine into Canada from South America via Mexico and the United States.

More than 70 members of the Royal Newfoundland Constabulary and Quebec's Provincial Biker Enforcement Unit carried out the raids, which represented the culmination of an 18monthlong investigation dubbed Operation Roadrunner. The arrests were made near St. John's in Newfoundland and near the towns of Laval and La



The jungles of South America, where cocaine is produced, seem a long way from the St. Lawrence River. Using a sophisticated shipment and distribution network, however, criminal and militant organizations can cover the distance in a few days.

Tuque in Quebec. In Newfoundland, authorities seized \$300,000 in cash, 51 pounds of marijuana and 19 pounds of cocaine, as well as vehicles, weapons and computers. In Quebec, \$170,000 and four houses were seized.





source: banderasnews.com

Applications: Interdiction

Canada and the Transcontinental Drug Links Strategic Forecasting Inc go to original Email Page Print Page Email Us

Canadian police conducted several simultaneous raids on suspected drug traffickers in Newfoundland and Quebec provinces Oct. 11, arresting two dozen people and seizing marijuana, cocaine, weapons, cash and property. The drugtrafficking ring, which Canadian authorities believe was operated by the Quebec-based Hell's Angels motorcycle/crime gang, could have smuggled the cocaine into Canada from South America via Mexico and the United States.

More than 70 members of the Royal Newfoundland Constabulary and Quebec's Provincial Biker Enforcement Unit carried out the raids, which represented the culmination of an 18monthlong investigation dubbed Operation Roadrunner. The arrests were made near St. John's in Newfoundland and near the towns of Laval and La



The jungles of South America, where cocaine is produced, seem a long way from the St. Lawrence River. Using a sophisticated shipment and distribution network, however, criminal and militant organizations can cover the distance in a few days.

Tuque in Quebec. In Newfoundland, authorities seized \$300,000 in cash, 51 pounds of marijuana and 19 pounds of cocaine, as well as vehicles, weapons and computers. In Quebec, \$170,000 and four houses were seized.

 Monitoring / halting an adversary's activity on a network

- Maximum-Flow Interdiction
- Shortest-Path Interdiction
- Action:
 - Destruction of certain nodes / edges
 - Reduction of capacity / increase of cost on certain edges
- The problems are NP-hard! Survey (Collado and Papp, 2012)
- Uncertainties:
 - Network characteristics
 - Follower's response

source: banderasnews.com

Bilevel Optimization

General bilevel optimization problem



Both players may involve integer decision variables, functions can be non-linear, non-convex...

Bilevel Optimization

General bilevel optimization problem



Fakultät für Mathematik und Informatik

Hierarchy of bilevel optimization problems



About our journey

- With **sparse MILP formulations**, we can now solve to optimality:
 - Covering Facility Location (Cordeau, Furini, Ljubic, 2018): 20M clients
 - Code: <u>https://github.com/fabiofurini/LocationCovering</u>
 - Competitive Facility Location (Ljubic, Moreno, 2017): 80K clients (nonlinear)
 - Facility Location Problems (Fischetti, Ljubic, Sinnl, 2016): 2K x 10K instances
 - Steiner Trees (DIMACS Challenge, 2014): 150k nodes, 600k edges
- Common to all: Branch-and-Benders-Cut

Is there a way to exploit sparse formulations along with Branch-and-Cut for bilevel optimization?

Problems addressed today...

- Interdiction-Like Problems: LEADER "interdicts" FOLLOWER by removing some "objects". Both agents play pure strategies.
- FOLLOWER solves a combinatorial optimization problem (mostly, an NPhard problem!). One could build a payoff matrix (exponential in size!).
- We propose a generic Branch-and-Interdiction-Cuts framework to <u>efficiently solve these problems in practice</u>!
 - Assuming monotonicty property for FOLLOWER: interdiction cuts (facet-defining)
 - Computationally outperforming state-of-the-art
- Draw a connection to some problems in Graph Theory

Based on a joint work with...

- M. Fischetti, I. Ljubic, M. Monaci, M. Sinnl: A new general-purpose algorithm for mixed-integer bilevel linear programs, *Operations Research* 65(6): 1615-1637, 2017
- M. Fischetti, I. Ljubic, M. Monaci, M. Sinnl: Interdiction Games and Monotonicity, with Application to Knapsack Problems, *INFORMS Journal on Computing*, in print, 2018
- F. Furini, I. Ljubic, P. San Segundo, S. Martin: The Maximum Clique Interdiction Game, Optimization Online, 2018
- F. Furini, I. Ljubic, E. Malaguti, P. Paronuzzi: On Integer and Bilevel Formulations for the k-Vertex Cut Problem, submitted, 2018

Branch-and-Interdiction-Cut

A gentle introduction

Interdicting Communities in a Network



Critical Nodes: disconnect the network "the most" Survey: Lalou et al. (2018)

Defender-Attacker Game

LEADER: eliminates the nodes FOLLOWER: builds communities After studying the lives of 172 terrorists, Sageman found the most common factors driving them are the social ties. Communities in social networks are often characterized as **densely connected subgraphs**.



Research question: assuming that one can prevent k members of doing criminal activities, what is the size of the largest community that will remain?

Hamburg Cell: Max-Clique Interdiction





k=0

k=4

Hamburg Cell: Max-Clique Interdiction





k=8

Bilevel Integer Program

$$\begin{split} \min d^T y \\ b^T x \leq B_D \\ y \in \arg \max\{d^T y: \\ y_i \leq 1 - x_i, \quad i \in N \\ y \in Y\} \\ x_i \text{ binary}, \quad i \in N \end{split}$$
$$y_i = \begin{cases} 1 & \text{if node } i \text{ belongs to the community} \\ 0 & \text{otherwise} \\ x_i = \begin{cases} 1 & \text{if node } i \text{ is interdicted} \\ 0 & \text{otherwise} \end{cases} \quad i \in N. \end{split}$$

 $\min w$ $b^{T}x \leq B_{D}$ $w \geq \max\{d^{T}y:$ $y_{i} \leq 1 - x_{i}, \quad i \in I$ $y \in Y\}$ $x_{i} \text{ binary}, \quad i \in I$

Value Function Reformulation

$\min_{x \in \mathbb{R}^{ N }, w \in \mathbb{R}} w$	$\min_{x \in \mathbb{R}^{ N }} b^T x$
$w \ge \Phi(x)$	$K \ge \Phi(x)$
$b^T x \le B_D$	x_i binary, $i \in N$
x_i binary, $i \in N$	
INTERDICTION: Min-max	BLOCKING: Min-num or Min-sum

Value Function Reformulation



• To be solved in a branch-and-cut fashion

How to convexify the value function?

Convexification

Observation: Given x, for the optimal follower's response it holds:

$$x_j + y_j \le 1 \quad \Rightarrow \quad x_j y_j = 0 \qquad j \in N$$

Instead of solving:

$$\begin{split} \Phi(\boldsymbol{x}) &= \max_{y \in \mathbb{R}^{n_2}} d^T y & Y = \{ y \in \mathbb{R}^{n_2} : \quad Q \, y \leq q_0, \\ 0 \leq y_j \leq 1 - \boldsymbol{x_j}, \quad \forall j \in N & y_j \text{ integer } \forall j \in J_y \}. \\ & y \in Y \end{split}$$

Wood (2011) proposes to move x into the objective function and find the penalties M_j , such that we can equivalently solve:

$$\Phi(\boldsymbol{x}) = \max_{\boldsymbol{y} \in \mathbb{R}^{n_2}} \{ d^T \boldsymbol{y} - \sum_{j \in N} \boldsymbol{M}_j \boldsymbol{x}_j \boldsymbol{y}_j \qquad = \max_{\hat{\boldsymbol{y}} \in \operatorname{conv}(\boldsymbol{Y})} \{ d^T \hat{\boldsymbol{y}} - \sum_{j \in N} \boldsymbol{M}_j \boldsymbol{x}_j \hat{\boldsymbol{y}}_j \}$$
$$\boldsymbol{y} \in \boldsymbol{Y} \}$$

Convexification \rightarrow Benders-Like Reformulation

Benders-Like Reformulation

$\min_{x \in \mathbb{R}^{n_1}, w \in \mathbb{R}} w$	
$w \ge d^T \hat{y} - \sum_{j \in N} M_j x_j \hat{y}_j$	$orall \hat{y} \in \hat{Y}$
$Ax \le b$	
x_j integer,	$\forall j \in J_x$
x_j binary,	$\forall j \in N.$

The choice of M_j is crucial:

- If FOLLOWER solves an LP: Wood (2011), M_j is the upper bound of the dual variable.
- If FOLLOWER solves the KNAPSACK PROBLEM: Caprara et al. (2016), De Negre (2011), $M_j = d_j$.
- In general: **OPEN QUESTION**.

If the follower satisfies monotonicity property...



If \hat{y} is a feasible follower and y' satisfies integrality constraints and $0 \le y' \le \hat{y}$, then y' is *also feasible*.

If the follower satisfies monotonicity property...



If \hat{y} is a feasible follower and y' satisfies integrality constraints and $0 \le y' \le \hat{y}$, then y' is *also feasible*.

- max-knapsack (set packing)
- max-clique
- max-relaxed-clique (s-plex: degree, s-clique: distance, s-bundle: connectivity)

A Careful Branch-and-Interdiction-Cut Design

- Separation: finding the best FOLLOWER's response for a given x*. NPhard, in general.
- A good **balance** between "lazy cut separation" (integer points only) and "user cut separation" (fractional points).



 \rightarrow Branch-and-Interdiction-Cut

A Careful Branch-and-Interdiction-Cut Design

- Separation: finding the best FOLLOWER's response for a given x*. NPhard, in general.
- A good **balance** between "lazy cut separation" (integer points only) and "user cut separation" (fractional points).
- **Crucial: specialized procedures/algorithms** for FOLLOWER's subproblem (if available).
- Combinatorial algorithms for LOWER and UPPER BOUNDS.
- Efficient **PREPROCESSING** techniques.



 \rightarrow Branch-and-Interdiction-Cut

A Careful Branch-and-Interdiction-Cut Design

- Separation: finding the best FOLLOWER's response for a given x*. NPhard, in general.
- A good **balance** between "lazy cut separation" (integer points only) and "user cut separation" (fractional points).
- **Crucial: specialized procedures/algorithms** for FOLLOWER's subproblem (if available).
- Combinatorial algorithms for LOWER and UPPER BOUNDS.
- Efficient **PREPROCESSING** techniques.
- Under monotonicity property: Interdiction cuts are facet-defining or could be lifted, otherwise.
- Resulting in general in **strong LP-relaxation bounds**.



 \rightarrow Branch-and-Interdiction-Cut

Max-Clique-Interdiction on Large-Scale Networks

	 Max-Clique Solver San Segundo et al. (2016) 			$k = \lceil 0.0 \rangle$				
	V	E	ω [s]	t [s]	$ V_p $	t [s]	$ V_p $	eliminated by preprocessing
socfb-UIllinois	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	
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	
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	
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	

Furini, Ljubic, Martin, San Segundo (2018)

Max-Clique-Interdiction on Large-Scale Networks

	 Max-Clique Solver San Segundo et al. (2016) 			$k = \lceil 0.0$				
	V	E	ω [s]	t [s]	$ V_p $	t [s]	$ V_p $	eliminated by preprocessing
socfb-UIllinois	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	
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	
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	
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	

#variables

Furini, Ljubic, Martin, San Segundo (2018)

B&IC Ingredients



τ

Comparison with the state-of-the-art MILP bilevel solver

		Branch-and- Interdiction-Cut				Generic B&C for Bilevel MILPs (Fischetti, Liubic, Monaci, Sinnl, 2017)				
V	#	# 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	

Slide "NOT TO BE SHOWN"



Downward Monotonicity: Assoc

"if $\hat{y} = (\hat{y}_N, \hat{y}_R)$ is a feasible follower for a given x and $y' = (y'_N, \hat{y}_R)$ satisfies integrality constraints and $0 \le y'_N \le \hat{y}_N$, then y' is **also feasible** for x".

The result can be further generalized

Relevant Operations Research applications. Two companies competing at the market for customers.

- LEADER: established on the market,
- FOLLOWER: a newcomer who wants to disrupt the market.

LEADER wants to keep the customers by providing them coupons, vouchers. FOLLOWER is solving a profit-maximization problem:

- NETWORK DESIGN: prize-collecting Steiner tree
- LOGISTICS: orienteering problems
- FACILITY LOCATION: profit maximization variant



And what about Graph Theory?

A weird example...

• **Property:** A set of vertices is a **vertex cover** if and only if its complement is an independent set

• Vertex Cover as a Blocking Problem:

- LEADER: interdicts (removes) the nodes.
- FOLLOWER: maximizes the size of the largest connected component in the remaining graph.
- Find the smallest set of nodes to interdict, so that FOLLOWER's optimal response is at most one.



The k-Vertex-Cut Problem

- A set of vertices is a **vertex** *k***-cut** if upon its removal the graph contains at least *k* components.
- **The** *k* **Vertex–Cut Problem:** Find a vertex *k*-cut of minimum cardinal-ity/weight.



Open question: Is there an ILP formulation in the natural space of variables?

Furini, Ljubic, Malaguti, Paronuzzi (2018)

The k-Vertex-Cut Problem

- A set of vertices is a **vertex** *k***-cut** if upon its removal the graph contains at least *k* components.
- **The** *k* **Vertex–Cut Problem:** Find a vertex *k*-cut of minimum cardinal-ity/weight.
- Influential nodes in a diffusion model for social networks, Kempe et al. (2005)
- Decomposition method for linear equation systems, e.g. GCG solver (Bastubbe, Lübbecke, 2017)

Open question: Is there an ILP formulation in the natural space of variables?



Furini, Ljubic, Malaguti, Paronuzzi (2018)

K-Vertex-Cut

Property: A graph G has at least k (not empty) components if and only if any cycle-free subgraph of G contains at most |V| - k edges.

Example with |V| = 9 and k = 3:



K-Vertex-Cut

Property: A graph G has at least k (not empty) components if and only if any cycle-free subgraph of G contains at most |V| - k edges.

Example with |V| = 9 and k = 3:



Stackelberg game:

- LEADER: searches the smallest subset of nodes to delete;
- FOLLOWER maximizes the size of the cycle-free subgraph on the remaining graph.

k-Vertex-Cut: Benders-like reformulation

$$\begin{split} \min \sum_{v \in V} x_v \\ \Phi(x) \leq |V| - \sum_{v \in V} x_v - k \\ x_v \in \{0, 1\} \end{split}$$

The value function reformulation

 $v \in V$.

The following **Natural Space Formulation**, is a valid model for the *k*-vertex cut problem (Furini et al. 2018):

$$\min \sum_{v \in V} x_v$$

$$\sum_{v \in V} [\deg_T(v) - 1] x_v \ge k - |V| + |E(T)| \qquad T \in \mathcal{T},$$

$$x_v \in \{0, 1\} \qquad v \in V.$$

Furini, Ljubic, Malaguti, Paronuzzi (2018)

k-Vertex-Cut: Benders-like reformulation



Furini, Ljubic, Malaguti, Paronuzzi (2018)

Conclusions.

And some directions for the future research.

Takeaways

- Bilevel optimization: very difficult!
- Branch-and-Interdiction-Cuts can work very well in practice:
 - Problem reformulation in the natural space of variables ("thinning out" the heavy MILP models)
 - Tight *"*interdiction cuts" (monotonicity property)
 - **Crucial:** Problem-dependent (combinatorial) separation strategies, preprocessing, combinatorial poly-time bounds
- Many graph theory problems (node-deletion, edge-deletion) could be solved efficiently, when approached from the bilevel-perspective



Possible directions for future research

- **Bilevel Optimization:** a better way of **integrating customer behaviour** into decision making models
- Generalizations of **Branch-and-Interdiction-Cuts** for:
 - Non-linear follower functions
 - Submodular follower functions
 - Interdiction problems under uncertainty
 - ...
- Extensions to **Defender-Attacker-Defender (DAD)** Models (trilevel games)
- Benders-like decomposition for general mixed-integer bilevel optimization

Thank you for your attention!

References:

 M. Fischetti, I. Ljubic, M. Monaci, M. Sinnl: A new general-purpose algorithm for mixedinteger bilevel linear programs, *Operations Research* 65(6): 1615-1637, 2017
 SOLVER: https://msinnl.github.io/pages/bilevel.html

- M. Fischetti, I. Ljubic, M. Monaci, M. Sinnl: Interdiction Games and Monotonicity, with Application to Knapsack Problems, *INFORMS Journal on Computing*, in print, 2018
- F. Furini, I. Ljubic, P. San Segundo, S. Martin: The Maximum Clique Interdiction Game, *Optimization Online*, 2018
- F. Furini, I. Ljubic, E. Malaguti, P. Paronuzzi: On Integer and Bilevel Formulations for the k-Vertex Cut Problem, submitted, 2018

Literature

- Bastubbe, M., Lübbecke, M.: A branch-and-price algorithm for capacitated hypergraph vertex separation. Technical Report, Optimization Online (2017)
- L. Brotcorne, M. Labbé, P. Marcotte, and G. Savard. A Bilevel Model for Toll Optimization on a Multicommodity Transportation Network, Transportation Science, 35(4): 345-358
- L. Brotcorne, M. Labbé, P. Marcotte, and G. Savard. Joint design and pricing on a network. Operations Research, 56 (5):1104–1115, 2008
- Caprara A, Carvalho M, Lodi A, Woeginger GJ (2016) Bilevel knapsack with interdiction constraints. INFORMS Journal on Computing 28(2):319–333
- R.A.Collado, D. Papp. Network interdiction models, applications, unexplored directions, Rutcor Research Report 4-2012, 2012.
- J.F. Cordeau, F. Furini, I. Ljubic. <u>Benders Decomposition for Very Large Scale Partial Set Covering</u> and <u>Maximal Covering Problems</u>, submitted, 2018
- S. Dempe. Bilevel optimization: theory, algorithms and applications, TU Freiberg, ISSN 2512-3750. Fakultät für Mathematik und Informatik. PREPRINT 2018-11
- DeNegre S (2011) Interdiction and Discrete Bilevel Linear Programming. Ph.D. thesis, Lehigh University
- M. Fischetti, I. Ljubic, M. Sinnl: Redesigning Benders Decomposition for Large Scale Facility Location, Management Science 63(7): 2146-2162, 2017

Literature, cont.

- R.G. Jeroslow. The polynomial hierarchy and a simple model for competitive analysis. Mathematical Programming, 32(2):146–164, 1985
- Kempe, D., Kleinberg, J., Tardos, E.: Influuential nodes in a diffusion model for social networks. In: L. Caires, G.F. Italiano, L. Monteiro, C. Palamidessi, M. Yung (eds.) Automata, Languages and Programming, pp. 1127-1138., 2005
- M. Labbé, P. Marcotte, and G. Savard. A bilevel model of taxation and its application to optimal highway pricing. Management Science, 44(12):1608–1622, 1998
- I. Ljubic, E. Moreno: Outer approximation and submodular cuts for maximum capture facility location problems with random utilities, European Journal of Operational Research 266(1): 46-56, 2018
- M. Sageman. Understanding Terror Networks. ISBN: 0812238087, University of Pennsylvania Press, 2005
- San Segundo P, Lopez A, Pardalos PM. A new exact maximum clique algorithm for large and massive sparse graphs. Computers & OR 66:81–94, 2016
- S. van Hoesel. An overview of Stackelberg pricing in networks. European Journal of Operational Reseach, 189:1393–1492, 2008
- R.K. Wood. Bilevel Network Interdiction Models: Formulations and Solutions, John Wiley & Sons, Inc., http://hdl.handle.net/10945/38416, 2010