ISCO 2018

5th International Symposium on Combinatorial Optimization

BOOK OF ABSTRACTS

Marrakesh, Morocco
April 11-13, 2018

Organizing Institutions

Ecole nationale supérieure d'informatique et d'analyse des systèmes (ENSIAS)

Lecture Notes in Computer Science

Université Mohammed V de Rabat

Laboratory for Analysis and Modelling of Decision Support Systems

Paris Dauphine University

ISCO 2018 Conference

5th International Symposium on Combinatorial Optimization

April 11-13 2018, Marrakesh, Morocco

BOOK OF ABSTRACTS

ISCO 2018 Committees

Conference Chairs

Abdellatif El Afia (Mohammed V University, Rabat, Morocco)
Jon Lee (University of Michigan, USA)
A. Ridha Mahjoub (Paris Dauphine University, France)
Giovanni Rinaldi (IASI, Roma, Italy)

Steering Committee

M. Baïou (CNRS, Clermont-Auvergne University, France)
P. Fouilhoux (Pierre and Marie Curie University, Paris, France)
L. Gouveia (University of Lisbon, Portugal)
N. Maculan (Federal University of Rio de Janeiro, Brasil)
A. R. Mahjoub (Paris Dauphine University, France)
V. Paschos (Paris Dauphine University, France)
G. Rinaldi (IASI, Roma, Italy)

Program Committee

M. Anjos (Polytechnique Montreal, Canada)
F. Barahona (IBM, T. J. Watson, New York, USA)
A. Basu (Johns Hopkins University, USA)
Y. Benadada (ENSIAS, Mohammed V University, Rabat, Morocco)
V. Bonifaci (IASI, Roma, Italy)
F. Bonomo (University of Buenos Aires, Argentine)
F. Carrabs (University of Salerno, Italy)
R. Cerulli (University of Salerno, Italy)
K. Chandrasekaran (University of Illinois, Urbana-Champaign, USA)
J. Correa (University of Chile, Santiago, Chile)
F. Eisenbrand (EPFL, Lausanne, Switzerland)
I. Elhallaoui (Polytechnique Montreal, Canada)
A. Elhilali Alaoui (Sidi Mohamed Ben Abdellah University, Fez, Morocco)

M. Ettaouil (Sidi Mohamed Ben Abdellah University, Fez, Morocco)
M. Fampa ((Federal University of Rio de Janeiro, Brazil)
S. Fujishige (RIMS, Kyoto University, Japan)
R. Fukasawa (University of Waterloo, Canada)
T. Fukunaga (JST, PRESTO, Japan)
B. Gendron (University of Montreal, Canada)
C. Gentile (IASI, Roma, Italy)
M. Gentili (University of Salerno, Italy)
E. Gourdin (Orange Labs, Paris, France)
V. Goyal (Columbia University, USA)
A. Gupte (Clemson University, USA)
M. Haouari (University of Qatar, Qatar)
I. Kacem (LCOMS, University of Lorraine, Metz, France)
V. Kaibel (Otto-von-Guericke Universität, Magdeburg, Germany)
N. Kamiyama (Kyushu University, Japan)
Y. Kobayashi (University of Tsukuba, Japan)
M. Labbé (Free University, Brussels, Belgium)
M. Lampis (Paris-Dauphine University, France)
A. Letchford (Lancaster University, UK)
L. Liberti (LIX, Ecole Polytechnique, Paris, France)
I. Ljubic (ESSEC, Paris, France)
M. Lübbecke (RWTH Aachen University, Germany)
S. Mattia (IASI, Roma, Italy)
S.T. McCormick (UBC, Vancouver, Canada)
V. Nagarajan (University of Michigan, USA)
P. Pesneau (University of Bordeaux, France)
F. Rendl (University of Klagenfurt, Graz, Austria)
J.J. Salazar Gonzalez (Universidad de La Laguna, Tenerife, Spain)
A. Toriello (Georgia Institute of Technology, USA)
P. Toth (University of Bologna, Italy)
E. Uchoa (UFF Rio de Janeiro, Brasil)
P. Ventura (IASI, Roma, Italy)
S. Weltge (ETHZ Zürich, Switzerland)
H. Yaman (Bilkent University, Ankara, Turkey)

Organizing Committee

Rdouan Faizi (Mohammed V University, Rabat, Morocco) (Chair)
Abdellatif El Afia (Mohammed V University, Rabat, Morocco)
Raddouane Chiheb (Mohammed V University, Rabat, Morocco)
Sanaa El Fkihi (Mohammed V University, Rabat, Morocco)
Pierre Fouilhoux (Pierre and Marie Curie University, Paris, France)
Rohaifa Khaldi (Mohammed V University, Rabat, Morocco)
Ridha Mahjoub (Paris Dauphine University, France)
Sébastien Martin (Lorraine University, Metz, France)
Fatima Ouzayd (Mohammed V University, Rabat, Morocco)

Additional Reviewers

Gary Au
Rafael Oliveira Colares
Mehdi Khosravian Ghadikolaei
Burak Kocuk
Georgios Stamoulis
Marc Goerigk
Jose Neto
Boshi Yang
Herve Kerivin
Okan Arslan
Cristina Bazgan
Omar El Housni
Uéverton Souza
Themis Gouleakis
Luidi Simonetti
Claudson Bornstein
Tiziano Bacci
Roland Grappe
Cédric Bentz
Tobias Mömke
Guilherme D. Da Fonseca
Vedat Bayram
Jerome Galtier
Michael Poss
Teobaldo Leite Bulhões Júnior
Danny Segev
Jesper Nederlof
Mayssoun Messaoudi
Fabrizio Marinelli
Gabriella Stecco
Clément Mommessin
Pawel Pralat
Fernanda Raupp
Gianpaolo Oriolo
Andrea Raiconi
Joseph Paat
Joshua Ignatius
Manish Bansal
Claudio Arbib
Cid C. De Souza
Julia Lange
Mario Ruthmair
Antoine Deza
Ana Flavia Macambira
Mohamed Ouzineb

Preface

Dear Participants,
welcome to ISCO 2018 the 5th International Symposium on Combinatorial Optimization!
After the previous editions held in Hammamet, Athens, Lisbon and Vietri sul Mare, we are happy to welcome you in Marrakesh, in the red city, one of the famous city of Morocco.
ISCO is nowadays a highly anticipated biannual meeting for the combinatorial optimization research community.
We are grateful to all authors who contributed to our high-level scientific program. Overall, we received about 120 submissions from researchers in many different countries. Among them, about 75 full papers were submitted for the LNCS post-conference proceedings book, and 36 of them were accepted. We would like to also thank all PC members, the members of the organizing committee and the external reviewers for their excellent work, within demanding time constraints. Our 21 contributed session span many different topics of combinatorial optimization. Together with the 4 invited lectures by internationally renowned researchers such as Marcia Fampa, Bernard Gendron, Fritz Eisenbrand and Franz Rendl, we are sure that they will be a source of insights and fruitful discussions for all participants.

We hope you will all enjoy the conference and your stay in Marrakesh!

Abdellatif El Afia
Jon Lee
A. Ridha Mahjoub
Giovanni Rinaldi
ISCO 2018 Co-Chairs

Plenary Lectures

Challenges in MINLP - The Euclidean Steiner Tree Problem in n-dimensional space
Marcia Fampa (Federal University of Rio de Janeiro, Brazil)

Proximity results and faster algorithms for Integer Programming using the Steinitz Lemma
Friedrich Eisenbrand (EPFL, Lausanne, Switzerland)

Lagrangian Relaxations and Reformulations for Network Design
Bernard Gendron (University of Montreal, Canada)

Order through partition: A semidefinite Programming Approach
Franz Rendl (University of Klagenfurt, Graz, Austria)

Challenges in MINLP - The Euclidean Steiner Tree Problem in n-dimensional space

Marcia Fampa
Federal University of Rio de Janeiro, Brazil, fampa@cos.ufrj.br

The Euclidean Steiner tree problem asks for a network of minimum length interconnecting a given set of points in n-dimensional space. We present a historical background for this problem, discuss existing algorithms to solve it, and identify characteristics of the problem that make its solution a big challenge when n is greater than 2, focusing on the application of MINLP solvers to different formulations. We present ideas that have been applied to handle the main difficulties in the solution of the problem, and point out others as future research directions.

Proximity results and faster algorithms for Integer Programming using the Steinitz Lemma

Friedrich Eisenbrand
EPFL, Lausanne, Switzerland, friedrich.eisenbrand@epfl.ch

We consider integer programming problems in standard form $\max\{cTx : Ax = b, x \geq 0$ $, x \in Z^n\}$ where $A \in Z^{m \times n}$, $b \in Z^m$ and $c \in Z^n$. We show that such an integer program can be solved in time $(m\delta)O(m) \cdot ||b||_\infty^2$, where δ is an upper bound on each absolute value of an entry in A. This improves upon the longstanding best bound of Papadimitriou (1981) of $(m \cdot \delta)^{O(m^2)}$, where in addition, the absolute values of the entries of b also need to be bounded by δ. Our result relies on a lemma of Steinitz that states that a set of vectors in R^m that is contained in the unit ball of a norm and that sum up to zero can be ordered such that all partial sums are of norm bounded by m. We also use the Steinitz lemma to show that the ℓ_1-distance of an optimal integer and fractional solution, also under the presence of upper bounds on the variables, is bounded by $m \cdot (2m \cdot \delta + 1)^m$. Here δ is again an upper bound on the absolute values of the entries of A. The novel strength of our bound is that it is independent of n. We provide evidence for the significance of our bound by applying it to general knapsack problems where we obtain structural and algorithmic results that improve upon the recent literature.

Lagrangian Relaxations and Reformulations for Network Design

Bernard Gendron
University of Montreal, Canada, bernard.gendron@cirrelt.ca

We consider a general network design model for which we compare theoretically different Lagrangian relaxations. Fairly general assumptions on the model are proposed, allowing us to generalize results obtained for special cases. The concepts are illustrated on the fixed-charge multicommodity capacitated network design problem, for which we present three different Lagrangian relaxations: the well-known shortest path and knapsack relaxations, and a new one, called the facility location relaxation. Dantzig-Wolfe reformulations are derived for each of these Lagrangian relaxations and bundle methods are proposed for solving these reformulations, along with Lagrangian heuristics and branch-and-price algorithms.

Order through partition: A semidefinite Programming Approach

Franz Rendl
University of Klagenfurt, Graz, Austria, franz.rendl@aau.at

Ordering Problems on n objects involve pairwise comparison among all objects. This typically requires $\binom{n}{2}$ decision variables.

In this talk we investigate the idea of partitioning the objects into k groups (k-partition) and impose order only among the partition blocks.

We demonstrate the efficiency of this approach in connection with the bandwidth minimization on graphs. We consider relaxations of the partition model with the following characteristics:

1) The weakest model is formulated in the space of symmetric $n \times n$ matrices and has the Hoffman-Wielandt theorem in combination with eigenvalue optimization as a theoretical basis.

2) We also consider semidefinite relaxations in the space of $n \times n$ matrices, involving k semidefinite matrix variables. The idea here is to linearize the quadratic terms using eigenvalue decompositions.

3) Finally, the strongest model is formulated in the space of symmetric $nk \times nk$ matrices. It is based on the standard reformulation-linearization idea.

We present theoretical results for these relaxations, and also some preliminary computational experience in the context of bandwidth minimization.

Co-authors: Renata Sotirov (Tilburg, Netherlands) and Christian Truden (Klagenfurt, Austria)

Sessions

WEA1 : Polyhedral Approaches I

- Refining Column Generation Subproblems Using Classical Benders' Cuts
  Jonas Witt, Marco Lübbecke, Stephen J. Maher.

- An exact column generation-based algorithm for Bi-Objective Vehicle Routing Problems
  Estèle Glize, Nicolas Jozefowiez, Sandra Ulrich Ngueveu.

- The Stop Number Minimization Problem: complexity and polyhedral analysis
  Mourad Baiou, Rafael Colares, Hervé Kerivin.

- The Next Release Problem: Complexity, Exact Algorithms and Computations
  José Carlos Almeida Jr., Felipe De C. Pereira, Marina V. A. Reis, Breno Piva.

Refining Column Generation Subproblems Using Classical Benders' Cuts

Marco E. Lübbecke[1], Stephen J. Maher[2], and Jonas T. Witt[1]

[1] Lehrstuhl für Operations Research, RWTH Aachen University
{marco.luebbecke,jonas.witt}@rwth-aachen.de
[2] Department of Management Science, Lancaster University
s.maher3@lancaster.ac.uk

Column generation is an iterative solution algorithm that can be applied to solve the linear programming (LP) relaxation of large-scale mixed-integer programs (MIPs). In each iteration the restricted master problem (RMP) is updated with the addition of promising variables, or columns, that are identified by solving subproblems. Unfortunately, it is possible that columns are generated by the subproblems that are not necessary to express an integer optimal solution of the MIP. Such columns are called *redundant* [5]. Since redundant columns are not necessary for an integer optimal solution, and the dual bound obtained by solving the LP relaxation is potentially stronger if redundant columns are not generated, it is expected to be advantageous to avoid their generation. This can be achieved by refining the feasible region of the column generation subproblems in order to eliminate (some) redundant columns. So far, only domain propagation techniques were applied to obtain tighter variable bounds in the column generation subproblems [2, 5]. In this work, we will address the refinement of column generation subproblems by adding inequalities to avoid the generation of redundant columns.

Consider a MIP that has been reformulated using Dantzig-Wolfe reformulation. Column generation is then employed to solve the LP relaxation of the reformulated problem. To evaluate the redundancy of a column, it must be determined whether there exists an integer optimal solution that can be expressed without using this column. Instead of testing this, we introduce a sufficient condition for redundancy, which is based on the LP relaxation of the original MIP. We fix all integer variables corresponding to a partial solution described by the evaluated column in the LP relaxation of the original MIP. The resulting LP is called the *redundancy LP*. If the redundancy LP is infeasible, the evaluated column is redundant since it is not part of any integer feasible solution to the reformulated problem. The Farkas proof of infeasibility of the redundancy LP can be used to generate classical Benders' feasibility cuts that eliminate the evaluated column: Potentially other redundant columns from the column generation subproblem are also eliminated. Additionally, the redundancy check can be strengthened by incorporating a primal bound on the optimal solution value. Inequalities that are generated using the redundancy LP are called *subproblem cuts*.

Column generation algorithms can then be adjusted as follows: After solving the column generation subproblems, we check each generated column for redundancy by solving the corresponding redundancy LP. If the redundancy

LP is infeasible for some column, we do not add the redundant column to the column generation master problem, but refine the corresponding column generation subproblem by adding subproblem cuts. Hence, the redundant column is forbidden in the subproblem and will not be generated again. If, on the other hand, the redundancy LP is feasible, we add the column to the column generation master problem. Since only a sufficient condition redundancy is checked, generated columns may still be redundant. In the case that all columns in a particular column generation iteration are forbidden, no columns will be added to the column generation master problem. For the exactness of the algorithm, in this case it is required that all refined column generation subproblems are resolved.

The column generation subproblem refinement algorithm is implemented in the branch-price-and-cut (BP&C) solver GCG and evaluated using various publicly available test instances arising from many application areas. Unfortunately, almost no subproblem cuts are generated when solving instances from "classical" applications of BP&C algorithms, such as bin packing, capacitated p-median or coloring problems. We explain that for several of these "classical" applications the underlying problem structure makes it difficult to identify redundant columns. Nevertheless, several subproblem cuts are found on capacitated lot-sizing and linearized unit commitment problem instances using *temporal* decompositions [1, 3, 4]. On several of these instances, the subproblem cuts strengthen the dual bound in the root node and improve the performance when solving the instances to optimality.

References

1. Fragkos, I., Degraeve, Z., Reyck, B.D.: A horizon decomposition approach for the capacitated lot-sizing problem with setup times. INFORMS Journal on Computing 28(3), 465–482 (2016), https://doi.org/10.1287/ijoc.2016.0691
2. Gamrath, G., Lübbecke, M.: Experiments with a generic Dantzig-Wolfe decomposition for integer programs. In: Festa, P. (ed.) Proceedings of the 9th. vol. 6049, pp. 239–252. Springer-Verlag, Berlin (2010)
3. Kim, K., Botterud, A., Qiu, F.: Temporal decomposition for improved unit commitment in power system production cost modeling. Tech. rep. (2017), http://www.mcs.anl.gov/papers/P7073-0717.pdf
4. Pimentel, C.M.O., e Alvelos, F.P., de Carvalho, J.M.V.: Comparing Dantzig-Wolfe decompositions and branch-and-price algorithms for the multi-item capacitated lotsizing problem. Optimization Methods and Software 25(2), 299–319 (2010), https://doi.org/10.1080/10556780902992837
5. Vanderbeck, F., Savelsbergh, M.W.: A generic view of Dantzig-Wolfe decomposition in mixed integer programming. Operations Research Letters 34(3), 296 – 306 (2006), http://www.sciencedirect.com/science/article/pii/S0167637705000659

An exact column generation-based algorithm for Bi-Objective Vehicle Routing Problems

Estle Glize[1(✉)], Nicolas Jozefowiez[2], and Sandra Ulrich Ngueveu[1]

[1] CNRS, LAAS, INSA, INP Toulouse, Toulouse, France,
glize,ngueveu@laas.fr,
[2] LCOMS, Université de Lorraine, Metz, France,
nicolas.jozefowiez@univ-lorraine.fr

We propose a new exact method for bi-objective vehicle routing problems where edges are associated with two costs. The method generates the minimum complete Pareto front of the problem by combining the scalarization of the objective function [2] and the column generation technique. The aggregated objective allows to apply the exact algorithm for the mono-objective vehicle routing problem of Baldacci et al. [1]. The algorithm is applied to a bi-objective VRP with time-windows. Computational results are compared with a classical bi-objective technique [3]. The results show the pertinence of the new method, especially for clustered instances.

References

1. Baldacci, R., Christofides, N., Mingozzi, A.: An exact algorithm for the vehicle routing problem based on the set partitioning formulation with additional cuts. Mathematical Programming 115, 351–385 (2008)
2. Ulungu, E.L., Teghem, J., Fortemps, P.H., Van Nieuwenhuyze, K.: The two phases method: An efficient procedure to solve bi-objective combinatorial optimization problems. Foundations of Computing and Decision Sciences 20, 149–165 (1995)
3. Haimes, Y., Lasdon, L., Wismer, D.: On a bicriterion formulation of the problems of integrated system identification and system optimization. IEEE Transactions on systems, man and cybernetics 1, 296–297 (1971)

The Stop Number Minimization Problem: complexity and polyhedral analysis

Mourad Baïou, Rafael Colares, and Hervé Kerivin

LIMOS, CNRS and Université Clermont Auvergne, Clermont-Ferrand, France
{baiou,colares,kerivin}@isima.fr

Keywords: autonomous vehicles, branch-and-cut, complexity

1 The Unit Stop Number Minimization Problem

The Stop Number Minimization Problem (SNMP) arises in the management of a dial-a-ride system [1] with autonomous vehicles. In such a system, a fleet of capacitated vehicles travels along a closed circuit network with predefined stations in a clockwise direction. Customers request for a ride, expressed as a load (*i.e.*, a number of seats) from an origin station to some destination station of their choice. The SNMP consists of assigning the customer requests to the vehicles such that no vehicle gets overloaded and the total number of pick-up/drop-off operations is minimized.

In this talk we focus on a constrained version of SNMP where each demand can only request a single seat in a vehicle and the fleet must respond to all requests in a single tour. This constrained problem is called Unit SNMP (USNMP) and is formally defined as follows. Let $V = \{1, \ldots, n\}$ denote the set of stations sequentially ordered as they appear in the circuit network and $E = \{e_1, \ldots, e_m\}$ denote the set of unit load dial-a-ride demands such that each demand $e \in E$ is specified by an origin station $o_e \in V$ and a destination station $d_e \in V$, that is, $e = (o_e, d_e)$. Also, let K denote the set of available identical vehicles each with capacity $C \in \mathbb{Z}_+$.

Let $\Delta_E(v) = \{e \in E : o_e \leq v \text{ and } d_e \geq v + 1\}$ be the set of demands that cross or starts at station v. Demand $e \in E$ *intersects* station $v \in V$ if $e \in \Delta_E(v)$. Then a feasible solution to USNMP is a partition of E into $|K|$ subsets $\{E_1, \ldots, E_{|K|}\}$, such that $|\Delta_{E_i}(v)| \leq C$ for any $i \in K$ and $v \in V$. Given a feasible solution $\{E_1, \ldots, E_{|K|}\}$, vehicle $i \in K$ stops in every station of $V(E_i)$. Therefore, the cost of this solution is $\sum_{i=1}^{|K|} |V(E_i)|$, and the USNMP is to find a feasible solution of minimum cost.

The USNMP was formulated in [2] as the following integer linear problem, where the variable x_e^i expresses the fact that demand e is assigned or not to vehicle i (that is, $x_e^i = 1$ if $e \in E_i$, $x_e^i = 0$ otherwise) and the variable y_v^i expresses whether or not vehicle i stops at station v (that is, $y_v^i = 1$ if $v \in V(E_i)$, $y_v^i = 0$ otherwise).

$$\min \quad \sum_{v \in V} \sum_{i \in K} y_v^i \tag{1}$$

$$
\begin{aligned}
\text{s.t.} \quad & \sum_{i \in K} x_e^i = 1 && \forall e \in E, && (2)\\
& \sum_{e \in \Delta_E(v)} x_e^i \leq C && \forall v \in V, i \in K, && (3)\\
& x_e^i \leq y_v^i && \forall i \in K, e \in E, v \in \{o_e, d_e\}, && (4)\\
& x_e^i \in \{0,1\} && \forall e \in E, i \in K, && (5)\\
& y_v^i \in \{0,1\} && \forall v \in V, i \in K. && (6)
\end{aligned}
$$

2 Our Contribution

In [2], SNMP was shown to be weakly NP-Hard using a reduction from Partition Problem and USNMP was conjectured to be NP-Hard. In this talk, we answer affirmatively to this conjecture by showing that for any fixed capacity $C \geq 2$, USNMP is strongly NP-Hard and APX-Complete.

An interesting particular case arises when there exists some station v' wherein all demands intersect, that is, $\Delta_E(v') = E$. We show that this can be solved in polynomial time when C=2 but is NP-Hard for $C \geq 3$ even when $G = (V, E)$ is restricted to the class of planar bipartite graphs.

From primal-dual relations, we show that the linear relaxation of (1)-(6) always provides the trivial dual bound of value $|V|$. To reinforce the formulation we introduce the following family of *k-tree inequalities*:

$$
\sum_{e \in T} x_e^i \leq \sum_{v \in V(G[T])} (d_{G[T]}(v) - 1) y_v^i \qquad \forall i \in K, j \in V, T \subseteq \Delta_E(j), \tag{7}
$$

such that the graph $G[T]$ induced by edge set T is a tree of C+1 edges, and $d_{G[T]}(v)$ denote the degree of vertex $v \in V$ in $G[T]$.

We give necessary and sufficient conditions for inequalities (7) to be facet-defining. We also prove that their separation problem is NP-Hard.

Finally we show that adding inequalities (7), even separated heuristically, and some symmetry-breaking constraints, significantly improves the performances of a branch-and-cut algorithm based on (1)-(6).

References

1. Cordeau, J-F., Laporte, G.: Dial-a-Ride: models and algorithms. Annals of Operations Research 153, 29-46 (2007).
2. Pimenta, V., Quilliot, A., Toussaint, H., Vigo, D.: Models and algorithms for reliability-oriented Dial-a-Ride with autonomous electric vehicles. European Journal of Operational Research 257 601–613 (2017).

The Next Release Problem: Complexity, Exact Algorithms and Computations

José Carlos Almeida Jr., Felipe de C. Pereira, Marina V. A. Reis, and Breno Piva *

Universidade Federal de Sergipe, Departamento de Computação, Av. Marechal Rondon, s/n, Jd. Rosa Elze, 49100-000, São Cristóvão, Sergipe, Brasil.
{fcpereira,marinavar,brenopiva}@dcomp.ufs.br

1 Introduction

The Next Release Problem (NRP) was first formalized by Bagnall et al. in [1] as a description of a company's next release plan considering its involvement in the development and maintenance of large, complex systems to a set of clients that have different needs and different values for the company.

The input of NRP can be formalized as been composed by a set R of requirements, a set of clients C, a budget $B \in \mathbb{Z}^+$ and a directed graph $D = (R \cup C, A)$ indicating the association between requirements and between requirements and clients. The set of arcs in D is $A \subseteq R \times (R \cup C)$. There are also two functions $\omega : R \to \mathbb{Z}^+$ and $\delta : C \to \mathbb{Z}^+$ indicating, respectively, the cost of each requirement and the value of each client.

The NRP can be formalized using the following integer linear program (IP) formulation due to Bagnall et al. [1]:

$$(\text{NRP}) \qquad z = \max \sum_{c \in C} \delta(c) y_c \tag{1}$$

$$\text{subject to} \qquad \sum_{r \in R} \omega(r) x_r \leq B \tag{2}$$

$$x_{r'} \leq x_r \qquad \forall\ (r, r') \in A \tag{3}$$

$$y_c \leq x_r \qquad \forall\ (r, c) \in A \tag{4}$$

$$x_r, y_c \in \{0, 1\}; r \in R, c \in C \tag{5}$$

Our contribution. In this work we prove the strong $\mathcal{NP}$-hardness of NRP and, therefore, does not admit an FPTAS. We present a new family of valid inequalities for the NRP IP formulation from [1]. We present a separation heuristic routine for this family of valid inequalities and use this routine to construct a Branch-and-Cut (B&C) algorithm. We create and make available a new set of bigger instances for the NRP based on instances from the literature. Finally, the performance of the B&C algorithm is compared against that of a Branch-and-Bound (B&B) algorithm that is currently the faster algorithm in the experiments found in the literature.

* The authors would like to thank PIBIC/UFS for the support.

2 The Complexity of NRP

In [1], Bagnall et al. showed that the NRP is $\mathcal{NP}$-hard. The proof is based on the fact that NRP generalizes the well-known 0-1 knapsack problem and, therefore, there is an obvious reduction from this problem to the NRP.

Proposition 1. *The* NRP *is strongly $\mathcal{NP}$-hard.*

The proof of Proposition 1 is obtained through a reduction from the k-clique problem to NRP.

Corollary 1. *The* NRP *does not admit an* FPTAS*, unless $\mathcal{P} = \mathcal{NP}$.*

3 Integer Programming Approach

Cover Inequalities. Given that the weight limit in the knapsack problem and the budget in the NRP have similar roles, it is possible to consider cover inequalities for the NRP. Therefore, it is easy to see that the inequalities $\sum_{i \in H} x_{r_i} \leq |H| - 1$ and $\sum_{i \in E(H)} x_{r_i} \leq |H| - 1$ are valid for (NRP) where $H \subseteq \{1, ..., |R|\}$ such that $\sum_{i \in H} \omega(r_i) > B$ and $E(H) = H \cup \{j \in \{1, ..., |R|\} \setminus H | \omega(r_j) \geq \omega(r_i) \forall i \in H\}$.

Let $E(H)$ be an extended cover for NRP. Let $\chi(C')$ denote an ordering of a set of clients C' according to some criterion. And denote by $\chi(C')[0]$ the index of the first client in this ordering. Now, let $F(c, H') = \{i \in \{1, ..., |R|\} \setminus H' | (r_i, c) \in A\}$ denote the set of requirements whose indexes are not in H' that are prerequisites of client c. Finally, a client cover over sets H' of requirements indexes and C' of clients can be defined as follows: $CC(H', C') = \chi(C')[0] \cup CC(H' \cup F(\chi(C')[0], H'), C' \setminus \{\chi(C')[0]\})$ if $\sum_{j \in F(c,H')} \omega(r_j) \geq \omega(r_i)$ for all $i \in E(H)$ or $CC(H', C') = CC(H', C' \setminus \{\chi(C')[0]\})$ otherwise, moreover, $CC(H', \emptyset) = \emptyset$. With these definitions we can now define the extended client cover inequalities as $\sum_{i \in E(H)} x_{r_i} + \sum_{j \in CC(E(H),C)} y_{c_j} \leq |H| - 1$.

A separation heuristic was divised in order to find extended cover client inequalities that are violated by fractional solutions found at each node the B&B search and, that way, generate a B&C algorithm.

4 Conclusions and Future Works

NRP is strongly $\mathcal{NP}$-hard and does not admit an FPTAS, however, it is still not clear whether there is a PTAS or if there is a limit on its approximability. Despite having exponential time complexities, IP based algorithms have shown their usefulness. Furthermore, the B&C algorithm suggests that there is still room for improvement, since this algorithm was able to solve bigger instances quicker than a simple B&B algorithm.

References

1. A.J. Bagnall, V.J. Rayward-Smith, and I.M. Whittley. The next release problem. *Information and Software Technology*, 43(14):883 – 890, 2001.

WEA2 : Approximation Algorithms I

- Efficient Algorithms for Measuring the Funnel-likeness of DAGs
  Marcelo Garlet Millani, Hendrik Molter, Rolf Niedermeier, Manuel Sorge.

- On bounded pitch inequalities for the min-knapsack polytope
  Yuri Faenza, Igor Malinovic, Monaldo Mastrolilli, Ola Svensson.

- Approximating the Caro-Wei Bound for Independent Sets in Graph Streams
  Graham Cormode, Jacques Dark, Christian Konrad.

- A PTAS for the Time-Invariant Incremental Knapsack problem
  Yuri Faenza, Igor Malinovic.

Efficient Algorithms for Measuring the Funnel-likeness of DAGs*

Marcelo Garlet Millani[1], Hendrik Molter[1], Rolf Niedermeier[1], and Manuel Sorge[1,2]

[1] Institut für Softwaretechnik und Theoretische Informatik, TU Berlin, Germany
{m.garletmillani, h.molter, rolf.niedermeier}@tu-berlin.de
[2] Department of Industrial Engineering and Management, Ben-Gurion University of the Negev, Beer Sheva, Israel, sorge@post.bgu.ac.il

Abstract. Funnels are a new natural subclass of DAGs. Intuitively, a DAG is a funnel if every source-sink path can be uniquely identified by one of its arcs. Funnels are an analog to trees for directed graphs that is more restrictive than DAGs but more expressive than in-/out-trees. Computational problems such as finding vertex-disjoint paths or tracking the origin of memes remain NP-hard on DAGs while on funnels they become solvable in polynomial time. Our main focus is to determine the algorithmic complexity of finding out how funnel-like a given DAG is. To this end, we focus on the NP-hard problem of computing for a given DAG the arc-deletion distance to a funnel. We develop efficient exact and approximation algorithms for the problem and test them on synthetic random graphs and real-world graphs.

Introduction. Directed acyclic graphs (DAGs) are finite directed graphs (digraphs) without directed cycles and appear in many applications, including the representation of precedence constraints in scheduling, data processing networks, causal structures, or inference in proofs. From a more graph-theoretic point of view, DAGs can be seen as a directed analog of trees; however, their combinatorial structure is much richer. Thus a number of directed graph problems remain NP-hard even when restricted to DAGs. This motivates the study of subclasses of DAGs. We study *funnels* which are DAGs where each source-sink path has at least one *private arc*, that is, no other source-sink path contains this arc.

Funnels are both of combinatorial and graph-theoretic as well as of practical interest: First, funnels are a natural compromise between DAGs and trees as, similarly to in- or out-trees, the private-arc property guarantees that the overall number of source-sink paths is upper-bounded linearly by the funnel's number of arcs, yet multiple paths connecting two vertices are possible. Second, we show that funnels, in a divide & conquer spirit, allow for a vertex partition into a set of *forking* vertices with indegree one and possibly large outdegree and a set of *merging* vertices with outdegree one and possibly large indegree. This

* This is a two-page summary of the same-titled submission to ISCO 2018. A full version is available on arXiv: https://arxiv.org/abs/1801.10401.

partitioning helps in designing our algorithms. Third, in terms of applications, due to the simpler structure of funnels, problems such as DAG Partitioning, Vertex Disjoint Paths, (also known as k-Linkage), or a variation of the problem Network Inhibition become tractable on funnels while they are NP-hard on DAGs. Altogether, we feel that funnels are one of so far few natural subclasses of DAGs.

Contribution and Results. The focus of this paper is on investigating the complexity of turning a given DAG into a funnel by a minimum number of arc deletions. The motivation for this is twofold. First, due to the noisy nature of real-world data, we expect that graphs from practice are not pure funnels, even though they may adhere to some form of funnel-like structure. To test this hypothesis we need efficient algorithms to determine funnel-likeness. Second, as mentioned above, natural computational problems become tractable on funnels (e.g., k-Linkage). Thus it is promising to try and develop fixed-parameter algorithms for such NP-hard DAG problems with respect to distance parameters to funnels. This approach is known as exploiting the "distance from triviality". A natural way to measure the distance of a given DAG D to a funnel is the *arc-deletion distance to a funnel*, the minimum number of arcs that need to be deleted from D to obtain a funnel. The problem of computing this distance parallels the well-studied NP-hard Feedback Arc Set problem where the task is to turn a given digraph into a DAG by a minimum number of arc deletions.

Formally, we study the problem Arc-Deletion Distance to a Funnel (ADDF) where, given a DAG D, we want to find its arc-deletion distance d to a funnel. Our main results are as follows.

- We show that ADDF is NP-hard.
- We give a linear-time factor-two approximation algorithm for ADDF.
- We present an $\mathcal{O}(3^d \cdot |D|)$ time algorithm for ADDF, where $|D|$ is the size of the input DAG.[3]

We empirically evaluated the practical usefulness of our algorithms through experiments. The approximation algorithm always found solutions which where within a factor of 1.16 of the optimal ones. The exact algorithm solved most instances with $50 \leq d \leq 175$ within ten minutes. Our data sets contain artificial and real-world instances with between 250 and 29810 vertices. From the experiments we conclude that our algorithms are of practical interest.

Conclusion. Our results add to the relatively small list of fixed-parameter tractability results for directed graphs and introduce a novel interesting structural parameter for directed (acyclic) graphs. In particular, our approximation and fixed-parameter algorithms could help to establish the arc-deletion distance to a funnel as a useful "distance-to-triviality measure" for designing fixed-parameter algorithms for NP-hard problems on DAGs. We leave open whether computing a DAG's arc-deletion distance to a funnel is APX-hard. Finally, funnels might provide a basis for defining some useful digraph width or depth measures.

[3] There is also a simple $\mathcal{O}(5^d \cdot |D|^2)$ time algorithm for general digraphs.

On bounded pitch inequalities for the min-knapsack polytope

Yuri Faenza[1], Igor Malinović[2], Monaldo Mastrolilli[3], Ola Svensson[2]

[1] Columbia University (USA)
[2] École Polytechnique Fédérale de Lausanne (Switzerland)
[3] IDSIA Lugano (Switzerland)

The min-knapsack problem (MinKnap)

$$\min c^T x \quad \text{s.t.} \quad p^T x \geq 1,\ x \in \{0,1\}^n \tag{1}$$

is the variant of the max knapsack problem (MaxKnap) where, given a cost vector c and a profit vector p, we want to minimize the total cost given a lower bound on the total profit. MinKnap is known to be NP-Complete, even when $p = c$, while its complexity is settled with the classical FPTAS.

However, in applications one aims at developing techniques that remain valid when less structured constraints are added on top of the original knapsack one. This can be achieved by providing *strong* linear relaxations for the problem, providing a good starting point for any branch-and-bound procedure. The most common way to measure the strength of a linear relaxation is by measuring its *integrality gap*, and this is the point where MinKnap and MaxKnap seem to be very different. Indeed, the standard linear relaxation for MaxKnap has integrality gap 2, and this can be boosted to $(1+\epsilon)$ by an extended formulation with $n^{\bar{O}(1/\epsilon)}$ many variables and constraints, for $\epsilon > 0$ [2]. Conversely, the standard linear relaxation for MinKnap has unbounded integrality gap, and this remains true even after $\Theta(n)$ rounds of the Lasserre hierarchy [6]. It is an open problem whether there exists an extended LP formulation for MinKnap with polynomially many constraints and constant integrality gap.

Recent results showed the existence [1] and gave an explicit construction [5] of a linear relaxation for MinKnap of quasi-polynomial size with integrality gap $2+\epsilon$. This is obtained by giving an approximate formulation for *Knapsack Cover inequalities* (KC) [4]. Adding those exponentially many inequalities that can be approximately separated [4] gives an integrality gap of 2. This bound is tight even in the simpler case when $p = c$. A well-behaved candidate for further reducing the gap are so called *bounded pitch* inequalities [3]. Intuitively, the *pitch* is a parameter measuring the complexity of an inequality, and the associated separation problem is NP-Hard already for pitch-1 (also known as *unweighted cover inequalities* [1]).

In this paper, we study structural properties and separability of bounded pitch inequalities for MinKnap, and the strength of linear relaxations for MinKnap when they are added. Let $\mathcal{F}$ be the set given by pitch-1, pitch-2, and inequalities from the linear relaxation of (1). We first show that, for any arbitrarily small precision, we can solve in polynomial time the *weak separation problem* for the set $\mathcal{F}$.

Theorem 1. *Given a* MinKnap *instance* (1), *for each fixed* $\epsilon > 0$*, there exists an algorithm that takes as input a point* $\bar{x}$ *and, in time polynomial in* n*, either outputs an inequality from* $\mathcal{F}$ *that is violated by* $\bar{x}$*, or outputs a point* $\bar{y}$, $\bar{x} \leq \bar{y} \leq (1+\epsilon)\bar{x}$ *that satisfies all inequalities in* $\mathcal{F}$.

It is then a natural question whether bounded pitch inequalities can help to reduce the integrality gap below 2.

Theorem 2. *Consider an instance of* MinKnap (1) *with* $p = c$*. Denote by* K *the linear relaxation of* (1) *to which all pitch-1 and pitch-2 inequalities have been added. The integrality gap of* K *is at most* $3/2$.

However, this is false in general. Indeed, we also prove that KC plus bounded pitch inequalities do not improve upon the integrality gap of 2.

Theorem 3. *For any fixed* $k \in \mathbb{N}$ *and* $n \in \mathbb{N}$ *sufficiently large, there exists a* MinKnap *instance such that its standard linear relaxation, with added all valid KC and inequalities of pitch at most* k*, has integrality gap* $\frac{2}{1+\frac{k}{n}} \approx 2$.

Moreover, bounded pitch alone can be much weaker than KC: we show that, for each fixed k, the integrality gap may be unbounded even if all pitch-k inequalities are added. Using the relation between bounded pitch and Chvátal-Gomory (CG) closures established in [3] we obtain the following.

Theorem 4. *For a fixed* $q \in \mathbb{N}$*, let* $CG^q(K)$ *be the* q*-th CG closure of the* MinKnap*. There exists an instance* K *such that the integrality gap of* $CG^q(K)$ *is* $\Omega(\sqrt{n})$.

References

1. Abbas Bazzi, Samuel Fiorini, Sangxia Huang, and Ola Svensson. Small extended formulation for knapsack cover inequalities from monotone circuits. In *Proc. of SODA '17*, pages 2326–2341, 2017.
2. Daniel Bienstock. Approximate formulations for 0-1 knapsack sets. *Oper. Res. Lett.*, 36(3):317–320, May 2008.
3. Daniel Bienstock and Mark Zuckerberg. Approximate fixed-rank closures of covering problems. *Mathematical Programming*, 105(1):9–27, Jan 2006.
4. Robert D. Carr, Lisa K. Fleischer, Vitus J. Leung, and Cynthia A. Phillips. Strengthening integrality gaps for capacitated network design and covering problems. In *Proc. of SODA '00*, pages 106–115, 2000.
5. Samuel Fiorini, Tony Huynh, and Stefan Weltge. Strengthening convex relaxations of 0/1-sets using boolean formulas. *arXiv preprint arXiv:1711.01358*, 2017.
6. Adam Kurpisz, Samuli Leppänen, and Monaldo Mastrolilli. On the hardest problem formulations for the 0/1 lasserre hierarchy. *Math. Oper. Res.*, 42(1):135–143, 2017.

Approximating the Caro-Wei Bound for Independent Sets in Graph Streams

Graham Cormode, Jacques Dark, and Christian Konrad

Department of Computer Science, Centre for Discrete Mathematics and its Applications (DIMAP), University of Warwick, Coventry, UK
{g.cormode,j.dark,c.konrad}@warwick.ac.uk

1 Introduction

In the model of streaming graph analysis, an algorithm observes the edges of an input graph one by one while maintaining some representative summary. We seek to understand how well different problems can be solved in this model, in terms of the size of the summary and the accuracy of any approximation obtained.

Independent Sets and the Caro-Wei Bound. Given an n-vertex graph $G = (V, E)$, an *independent set* $I \subseteq V$ is a set of nodes such that there is no edge between any pair, and the *independence number* $\alpha(G)$ of G is the size of a *maximum independent set*, i.e., one of maximum cardinality. Approximating $\alpha(G)$ is a difficult task in the streaming model: Halldórsson et al. [ICALP 2012] proved that space $\Omega(\frac{n^2}{c^2})$ is necessary for computing a c-approximation to $\alpha(G)$. Due to these extensive space requirements, in this paper, we address the problem of approximating the *Caro-Wei bound*, a well-known lower bound on $\alpha(G)$. Caro [1979] and Wei [1981] proved that every graph G contains an independent set of size

$$\beta(G) := \sum_{v \in V} \frac{1}{\deg_G(v) + 1} \ .$$

The quantity $\beta(G)$ is an attractive bound. For example, the sequential min-degree Greedy algorithm produces an independent set of size at least $\beta(G)$, and $\beta(G)$ approximates $\alpha(G)$ within a poly-logarithmic factor in many interesting graph classes, such as graph of polynomially bounded-independence [Halldórsson, Konrad, DISC 2015].

Starting Point: -1 Frequency Moment. Approximating $\beta(G)$ is essentially the same as approximating the -1 (negative) frequency moment (or the harmonic mean) of a frequency vector derived from the vertex degrees in a graph stream. Braverman and Chestnut [APPROX 2015] showed that computing a $(1+\epsilon)$-approximation to the harmonic mean in one pass requires $\Omega(n)$ space if the length of the input sequence is $\Omega(n^2)$. While this lower bound is designed for arbitrary frequency vectors, it can be embedded into a graph with $\Theta(n^2)$ edges so that frequencies correspond to vertex degrees. This implies we cannot find an algorithm to approximate the Caro-Wei bound within a factor of $1+\epsilon$ which guarantees that the space used will always be sublinear.

2 Our Results

Despite these lower bounds, we are able to provide upper and lower bounds that improve on those stated above. Since in our setting the frequency vector is derived from a graph stream, we can exploit the properties of the underlying graph. In our first result, we relate the space complexity of our algorithm to a given lower bound γ on $\beta(G)$. A meaningful lower bound γ is easy to obtain: It is easy to see that the Turán bound, which shows that $n/(\overline{d}+1)$ is a lower bound on $\alpha(G)$, is also a lower bound on $\beta(G)$, where $\overline{d}$ is the average degree of the input graph. Our first result is then a one-pass randomized streaming algorithm with space $\mathrm{O}(\frac{n \log n}{\gamma c^2})$ that approximates $\beta(G)$ within a factor of c with high probability. Using $\gamma = \frac{n}{\overline{d}+1}$, the space becomes $\mathrm{O}(\frac{\overline{d} \log n}{c^2})$, which is polylogarithmic for graphs of constant average degree such as planar graph or bounded arboricity graphs. The algorithm can also give a $(1+\epsilon)$-approximation using $\mathrm{O}(\frac{n \log n}{\gamma \epsilon^2})$ space.

We prove that our algorithm is best possible (up to a log factor). Via a reduction from a hard problem in communication complexity, we show that every p-pass streaming algorithm for computing a c-approximation to $\beta(G)$ requires $\Omega(\frac{n}{\beta(G)c^2 p})$ space. This lower bound also holds in the *vertex arrival order*, where vertices arrive one by one together with those incident edges that connect to vertices that have previously arrived. Our lower bound is more general than the lower bound from Braverman and Chestnut, since their lower bound only holds for $(1+\epsilon)$-approximation algorithms and does not establish a dependency on the output quantity, i.e., the -1-negative frequency moment.

Our lower bound shows that the promise that the input stream is in vertex arrival order is not helpful for approximating $\beta(G)$. However, if we regard the task of approximating $\beta(G)$ as obtaining a (hopefully large) lower bound on the size of a maximum independent set of the input graph, then any value sandwiched between $\beta(G)$ and the maximum independent set size would be equally suitable (or even superior). In the vertex arrival setting, we give a randomized one-pass streaming algorithm with space $\mathrm{O}(\log^3 n)$, which outputs a value β' with $\beta' = \Omega(\beta(G)/\log n)$ and β' is at most the maximum independent set size.

3 Conclusion

From a technical perspective, we leverage this problem to advance the study of the degree moments in the streaming model. The fact that the frequencies are derived from the degrees of the input graph adds an additional dimension to the frequency moments problem, since, as illustrated by our two algorithms, the arrival order of edges can now be exploited. Furthermore, it seems plausible that exploiting additional graph structure could reduce the space complexity even further. One of the objectives of this work was the popularization of the Caro-Wei bound, and we thus only addressed the -1-negative frequency moment. Generalizing our approach to other frequency moments is left for future work.

A PTAS for the Time-Invariant Incremental Knapsack problem

Yuri Faenza[1], Igor Malinovic[2]

[1] Columbia University (USA)
[2] École Polytechnique Fédéral de Lausanne (Switzerland)

In this work, we study a generalization of the classical Maximum Knapsack problem to a discrete multi-period setting, known as *Time-Invariant Incremental Knapsack problem* (IIK). In IIK, we are given a set of items $[n]$ with profits $p : [n] \to \mathbb{R}_{>0}$ and weights $w : [n] \to \mathbb{R}_{>0}$, and a knapsack with non decreasing capacities b_t over time $t \in [T]$. We can add items at each time as long as the capacity constraint is not violated, and once inserted, an item cannot be removed from the knapsack. The goal is to maximize the total profit, which is defined to be the sum, over $t \in [T]$, of profits of items in the knapsack at time t.

IIK models a scenario where available resources (e.g. money, labour force) augment over time in a predictable way, allowing to grow our portfolio. Take e.g. a bond market with an extremely low level of volatility, where all coupons render profit only at their common maturity time T (*zero-coupon* bonds) and an increasing budget over time that allows buying more and more (differently sized and priced) packages of those bonds. A different application arises in government-type decision processes, where items are assets of public utility (schools, parks, etc.) that can be built at a given cost and give a yearly benefit (both constant over the years), and the community will profit each year those assets are available.

Previous work on IIK. Although the first publication on IIK appeared just very recently [2], this problem has been introduced by Bienstock et al. in [1] and studied in several PhD theses [3,5,6]. In [1], IIK is shown to be strongly NP-hard and an instance showing that the natural LP relaxation has unbounded integrality gap is provided. In the same paper, a PTAS is designed for $T = O(\log n)$. This improves over [5], where a PTAS for the special case $p = w$ is given when T is a constant. Again when $p = w$, a 1/2-approximation algorithm for generic T is provided in [3]. Results from [6] can be adapted to give an algorithm that solves IIK in time polynomial in n and of order $(\log T)^{O(\log T)}$ for a fixed approximation guarantee ε [4]. The authors in [2] provide an alternative PTAS for IIK with constant T, and a 1/2-approximation for arbitrary T with under the assumption that every item alone fits into the knapsack at $t = 1$.

Our contributions. In this paper, we give an algorithm for computing a $(1 - \varepsilon)$-approximated solution for IIK that depends polynomially on the number n of items and, for any fixed ε, also polynomially on the number of times T. In particular, our algorithm provides a PTAS for IIK, regardless of T.

Theorem 1. *There exists an algorithm that, when given as input $\varepsilon \in \mathbb{R}_{>0}$ and an instance $\mathcal{I}$ of* IIK *with n items and $T \geq 2$ times, produces a $(1-\varepsilon)$-approximation to the optimum solution of $\mathcal{I}$ in time $O(T^{h(\varepsilon)} \cdot n f_{LP}(n))$. Here $f_{LP}(m)$ is the time required to solve a linear program with $O(m)$ variables and constraints, and $h : \mathbb{R}_{>0} \to \mathbb{R}_{\geq 1}$ is a function depending on ε only. In particular, there exists a PTAS for* IIK.

Theorem 1 dominates all previous results on IIK [1,2,3,5,6] and, due to the hardness results in [1], settles the complexity of the problem. Interestingly, it is based on designing a disjunctive formulation – a tool mostly common among integer programmers and practitioners – and then rounding the solution to its linear relaxation with a greedy-like algorithm. We see Theorem 1 as an important step towards understanding the complexity landscape of knapsack problems over time.

Extensions. Following Theorem 1, one could ask for a PTAS for the general *Incremental Knapsack* (IK) problem. This is the modification of IIK (also introduced in [1]) where the objective function is $p_\Delta(x) := \sum_{t\in[T]} \Delta_t \cdot p^T x_t$, where $\Delta_t \in \mathbb{Z}_{>0}$ for $t \in [T]$ can be seen as time-dependent discounts. We show here some partial results.

Corollary 1. *There exists a PTAS-preserving reduction from* IK *to* IIK*, assuming $\Delta_t \leq \Delta_{t+1}$ for $t \in [T-1]$. Hence, under the hypothesis above,* IK *has a PTAS.*

Of independent interest is the fact that there is a PTAS for the modified version of IIK when each item can be taken multiple times. Unlike Corollary 1, this is not based on a reduction between problems, but on a modification of our algorithm.

Corollary 2. *There is a PTAS for the generalization of* IIK *where we are allowed to take object i at most d_i times, for each $i \in [n]$, and the vector $d = (d_1, \ldots, d_n) \in \mathbb{Z}^n_{>0}$ is part of the input.*

References

1. Daniel Bienstock, Jay Sethuraman, and Chun Ye. Approximation algorithms for the incremental knapsack problem via disjunctive programming. *CoRR*, abs/1311.4563, 2013.
2. Federico Della Croce, Ulrich Pferschy, and Rosario Scatamacchia. Approximation results for the incremental knapsack problem. *Proceedings of IWOCA '17*, 2017.
3. Jeffrey Hartline. Incremental optimization. *PhD thesis, Cornell University*, 2008.
4. Jay Sethuraman and Chun Ye. Personal communication, 2016.
5. Alexa Sharp. Incremental algorithms: Solving problems in a changing world. *PhD thesis, Cornell University*, 2007.
6. Chun Ye. On the trade-offs between modeling power and algorithmic complexity. *PhD thesis, Columbia University*, 2016.

WEA3 : Algorithms on Graphs

- Efficient Algorithm for Shortest Path in Inner Dualist of Hexagonal Graphs
  Faqir M Bhatti, Khawaja M Fahd.

- Graph Orientation with Splits
  Yuichi Asahiro, Jesper Jansson, Eiji Miyano, Hesam Nikpey, Hirotaka Ono.

- Even flying cops should think ahead
  Anders Martinsson, Florian Meier, Patrick Schnider, Angelika Steger.

- Improved Algorithms for k-Domination and Total k-Domination in Proper Interval Graphs
  Valeria Alejandra Leoni, Nina Chiarelli, Tatiana Romina Hartinger, María Inés Lopez Pujato, Martin Milanic.

On the Efficient Algorithm for Shortest Path in Inner Dualist of Hexagonal Graphs

Khawaja M Fahd and Faqir M Bhatti

Riphah Institute of Computing and Applied Sciences (RICAS), Riphah International University, Township Lahore, Pakistan
Email:{muhammad.fahd,fmbhatti}@riphah.edu.pk

Abstract. Shortest path problem is a well-known problem in the area of graph algorithms and theoretical computer science. It has applications in finding the optimal path in internet to download files and other optimization problems. .

Hexagonal graphs are also very important because of their many applications in mathematics and computer science. A hexagonal block is a hexagon formed by creating a cycle with 6 vertices and 6 edges. A hexagonal graph is a graph which is created by joining different hexagonal blocks. This hexagonal graph is a planar graph in two dimensions.

The inner dual planar hexagonal graphs are well studied in the literature[**?**]. In our presentation we replace each finite face of the hexagonal graph by a vertex. Two vertices are joined by an edge if two faces in hexagonal graph are separated by an edge. In this way we form inner dual graph. See Fig. 1 below.
These standard techniques of representing graphs like adjacency matrix and adjacency lists do not preserve the information about the orientation of any edges in the inner dual. We are interested in keeping track of the angle of each edge with the x-axis. There are several methods to save this information and one of them is the He-Matrix[**?**] which is given by two Chinese scientists. (See [**?**])

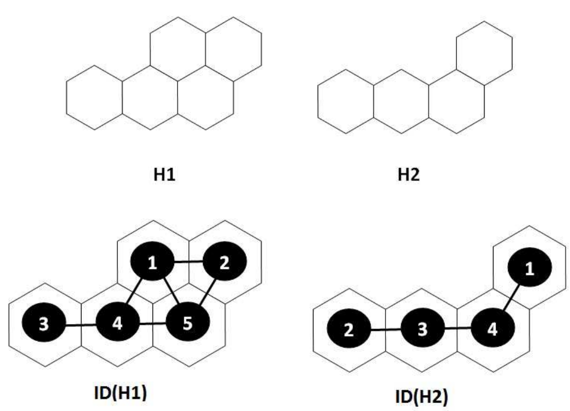


Fig. 1. Hexagonal Graphs H1 and H2 and their corresponding Inner Dual (ID) of Hexagonal Graphs

He-Matrix is an extension of adjacency matrix, where an entry of 0 means that the corresponding vertices are not connected directly by an edge. If there is an edge between two vertices then the entry in the matrix can be 1, 2 or 3 if the edge lies at an angle of 0°,60° or120° degrees with the x-axis respectively. He-matrix representation is only applicable to inner dual graphs of hexagonal systems because of its geometry.

If the inner dualist graph is rotated and reflected then there can be at most 6 non-isomorphic adjacency matrices of the graph. Each of them will correspond to a different orientation of the graph. These rotations and reflections will change the weight of each edge, resulting in the change in shortest path. The length of the shortest path in the new graph can increase or decrease depending on the orientation.

Finding shortest path between two given points is a well-known problem. Also there are known algorithms to solve this problem. Dijkstra's algorithm [**?**] and Bellman Ford's algorithm[**?**] are two of the fundamental algorithms used for solving the shortest path problem. None of the known algorithms for solving this problem have linear running time. In our presentation we give a linear time algorithm to find shortest path in the inner dual graphs of hexagonal graph for the first time using He-matrix.

Moreover, In this presentation we discuss different results regarding the shortest path of the inner dual graph in any given orientation. We present a formula that calculates the total weight of the shortest path between any two given points.

We compare shortest paths in different orientations which is in the form of an algorithm. This algorithm finds the shortest path between the given points which is of least weight among different shortest paths in all possible orientations. This algorithm also identifies the orientation that contains the minimum of all shortest paths. Our algorithm runs in deterministic linear time.

We also discuss the number of shortest paths present in the inner dual graph. A combinatorial formula for the number of paths is also given if all intermediate hexagons are present. Finally, we give a linear time algorithm to calculate the number of shortest paths in case where some hexagons are missing from the graph under consideration.

Graph Orientation with Splits (Short Paper)

Yuichi Asahiro[1], Jesper Jansson[2], Eiji Miyano[3], Hesam Nikpey[4], and Hirotaka Ono[5]

[1] Department of Information Science, Kyushu Sangyo University, Higashi-ku, Fukuoka 813-8503, Japan. E-mail: asahiro@is.kyusan-u.ac.jp

[2] Department of Computing, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong. E-mail: jesper.jansson@polyu.edu.hk

[3] Department of Systems Design and Informatics, Kyushu Institute of Technology, Iizuka, Fukuoka 820-8502, Japan. E-mail: miyano@ces.kyutech.ac.jp

[4] Department of Computer Engineering, Sharif University of Technology, Tehran, Iran. E-mail: hnikpey@ce.sharif.edu

[5] Graduate School of Informatics, Nagoya University, Japan. E-mail: ono@i.nagoya-u.ac.jp

An *orientation* of an undirected graph is an assignment of a direction to each of its edges. The computational complexity of constructing graph orientations that optimize various criteria has been studied, e.g., in [1–4, 6–9], and positive as well as negative results are known for many variants of these problems. An example is the *Minimum Maximum Outdegree Problem* (MMO) [3–6, 9], which takes as input an undirected, edge-weighted graph and asks for an orientation that minimizes the resulting maximum weighted outdegree taken over all vertices in the oriented graph. In this paper, we introduce a new variant of MMO called the *p-Split Minimum Maximum Outdegree Problem* (p-Split-MMO), where p is a specified non-negative integer. Here, one is allowed to perform a sequence of *p split operations* on the vertices before orienting the edges. When thinking of MMO as a load balancing problem, the split operation can be interpreted as a way to alleviate the burden on the existing machines by adding an extra machine.

Formally, let $G = (V, E, w)$ be an undirected, edge-weighted graph with vertex set V, edge set E, and edge weights defined by the function $w : E \to \mathbb{Z}^+$. An *orientation Λ of G* is an assignment of a direction to every edge $\{u, v\} \in E$, i.e., $\Lambda(\{u, v\})$ is either (u, v) or (v, u). For any orientation Λ of G, the *weighted outdegree* of a vertex u is $d^+_\Lambda(u) = \sum_{\substack{\{u,v\}\in E:\\ \Lambda(\{u,v\})=(u,v)}} w(\{u, v\})$ and the *cost* of Λ is $c(\Lambda) = \max_{u\in V}\{d^+_\Lambda(u)\}$. MMO is defined as follows:

> The Minimum Maximum Outdegree Problem (MMO):
>
> Given an undirected, edge-weighted graph $G = (V, E, w)$, where V, E, and w denote the set of vertices of G, the set of edges of G, and an edge-weight function $w : E \to \mathbb{Z}^+$, output an orientation Λ of G with minimum cost.

Next, for any $v \in V$, the set of vertices in V that are neighbors of v is denoted by $\Gamma[v]$ and the set of edges incident to v is denoted by $E[v]$. A *split operation* on a vertex v_i in G is an operation that transforms: (i) the vertex set of G to $(V \setminus v_i) \cup \{v_{i,1}, v_{i,2}\}$, where $v_{i,1}$ and $v_{i,2}$ are two new vertices; and (ii) the edge set of G to $(E \setminus E[v_i]) \cup \{\{v_{i,1}, s\} : s \in S\} \cup \{\{v_{i,2}, s'\} : s' \in \Gamma[v_i] \setminus S\}$ for some subset $S \subseteq \Gamma[v_i]$. For any non-negative integer p, a *p-split* on G is a sequence of p split operations successively applied to G. In a p-split, a new vertex resulting from a split operation may in turn be the target of a later split operation. The following new problem generalizes MMO:

The p-Split Minimum Maximum Outdegree Problem (p-Split-MMO):

Given an undirected, edge-weighted graph $G = (V, E, w)$, where V, E, and w denote the set of vertices of G, the set of edges of G, and an edge-weight function $w : E \to \mathbb{Z}^+$, output a graph G' and an orientation Λ' of G' such that: (i) G' is obtained by a p-split on G; (ii) Λ' has minimum cost among all orientations of all graphs obtainable by a p-split on G.

We have analyzed the computational complexity of p-Split-MMO and the results are summarized in the table below. (Please see the full version of this paper for details.) The number of vertices in G is denoted by n. Note that the edge weights are included in the input so it is possible to further classify the NP-hardness results as either weakly NP-hard or strongly NP-hard. The special case where $w(e) = 1$ for all $e \in E$ is referred to as the *unweighted* case.

	Unweighted case	General case
Constant p	$O((n + p)^p \cdot poly(n))$ time	Weakly NP-hard
Unbounded p	NP-hard	Strongly NP-hard

References

1. Y. Asahiro, J. Jansson, E. Miyano, and H. Ono. Graph orientations optimizing the number of light or heavy vertices. *Journal of Graph Algorithms and Applications*, 19(1):441–465, 2015.
2. Y. Asahiro, J. Jansson, E. Miyano, and H. Ono. Degree-constrained graph orientation: Maximum satisfaction and minimum violation. *Theory of Computing Systems*, 58(1):60–93, 2016.
3. Y. Asahiro, J. Jansson, E. Miyano, H. Ono, and K. Zenmyo. Approximation algorithms for the graph orientation minimizing the maximum weighted outdegree. *Journal of Combinatorial Optimization*, 22(1):78–96, 2011.
4. G. Borradaile, J. Iglesias, T. Migler, A. Ochoa, G. Wilfong, and L. Zhang. Egalitarian graph orientations. *Journal of Graph Algorithms and Applications*, 21(4):687–708, 2017.
5. G. S. Brodal and R. Fagerberg. Dynamic representations of sparse graphs. In *Proceedings of WADS 1999*, volume 1663 of *LNCS*, pages 342–351. Springer, 1999.
6. M. Chrobak and D. Eppstein. Planar orientations with low out-degree and compaction of adjacency matrices. *Theoretical Computer Science*, 86(2):243–266, 1991.
7. R. W. Deming. Acyclic orientations of a graph and chromatic and independence numbers. *Journal of Combinatorial Theory, Series B*, 26:101–110, 1979.
8. K. Khoshkhah. On finding orientations with the fewest number of vertices with small out-degree. *Discrete Applied Mathematics*, 194:163–166, 2015.
9. V. Venkateswaran. Minimizing maximum indegree. *Discrete Applied Mathematics*, 143(1–3):374–378, 2004.

Even flying cops should think ahead

Anders Martinsson, Florian Meier, Patrick Schnider, and Angelika Steger

Department of Computer Science
ETH Zürich, Switzerland
{maanders,meierflo,patrick.schnider,steger}@inf.ethz.ch

Abstract. We study the entanglement game, which is a version of cops and robbers, on sparse graphs. While the minimum degree of a graph G is a lower bound for the number of cops needed to catch a robber in G, we show that the required number of cops can be much larger, even for graphs with small maximum degree. In particular, we show that there are 3-regular graphs where a linear number of cops are needed.

Keywords: Cops and robbers, entanglement game, probabilistic method

In this paper we consider the *entanglement game*, introduced by Berwanger and Grädel [1] that is the following version of the cops and robbers game on a (directed or undirected) graph G. First, the robber chooses a starting position and the k cops are outside the graph. In every turn, the cops can either stay where they are, or they can fly one of them to the current position of the robber. Regardless of whether the cops stayed or one of them flew to the location of the robber, the robber then has to move to a neighbor of his current position that is not occupied by a cop. If there is no such neighbor, the cops win. The robber wins if he can run away from the cops indefinitely. The *entanglement number* of a graph G, denoted by $\text{ent}(G)$, is the minimal integer k such that k cops can catch a robber on G. In order to get accustomed to the rules of the game, it is a nice exercise to show that the entanglement number of an (undirected) tree is at most 2.

The main property that distinguishes the entanglement game from other variants of cops and robbers is the restriction that the cops are only allowed to fly to the current position of the robber. This prevents the cops from cutting off escape routes or forcing the robber to move into a certain direction. As we will show in this paper, it is this restriction that enables the robber to run away from many cops.

In a similar way to how the classical game of cops and robbers can be used to describe the treewidth of a graph, the entanglement number is a measure of how strongly the cycles of the graph are intertwined, see [2]. Just like many problems can be solved efficiently on graphs of bounded treewidth, Berwanger and Grädel [1] have shown that parity games of bounded entanglement can be solved in polynomial time.

As the cops do not have to adhere to the edges of the graph G in their movement, adding more edges to G can only help the robber. In fact, it can be seen easily that on the complete graph K_n with $n \geq 2$ vertices, n 1 cops are needed to catch the robber. Furthermore, observe that the minimum degree of the graph G is a lower bound on the entanglement number, as otherwise the robber will always find a free neighbor to move to. These observations seem to suggest that, on sparse graphs, the cops should have an advantage and therefore few cops would suffice to catch the robber. Indeed, on 2-regular graphs, it is easily checked that three cops can always catch the robber.

Motivated by this, we study the entanglement game on several classes of sparse graphs. We show that for sparse Erdős-Rényi random graphs, with high probability linearly many cops are needed.

Theorem 1. *For every $0 < \alpha < 1$ there exists a constant $C = C(\alpha) > 0$ such that for any $p \geq C/n$, αn cops do not suffice to catch the robber on $G_{n,p}$ with high probability. The same result holds for directed random graphs.*

We then apply similar ideas to show our main result.

Theorem 2. *There exists an $\alpha > 0$ such that with high probability αn cops do not suffice to catch the robber on the graph $G = M_1 \cup M_2 \cup M_3$, where M_1, M_2, M_3 are independent uniformly chosen random perfect matchings.*

Further, we give an upper bound on the entanglement number on 3-regular graphs.

Theorem 3. *For any 3-regular graph on n vertices, $\lfloor \frac{n}{4} \rfloor + 4$ cops suffice.*

Finally, we consider the entanglement game for graphs that are given by a more specific union of three perfect matchings, in fact, that are the union of a Hamilton cycle and a perfect matching. For graphs given by a Hamilton cycle and a perfect matching connecting every vertex to its diagonally opposite vertex, it may seem that the diagonal "escape" edges are quite nice for the robber. This, however, is not so: we show that for these graphs six cops are always sufficient. However, we also show, that if we replace this specific perfect matching by a random one, then with high probability a linear number of cops is needed.

We conclude that in contrast to the intuition that sparse graphs are advantageous for the cops, they are often not able to use the sparsity to their advantage. This shows that the freedom of the cops of being able to *fly* to any vertex is not helpful when they are only allowed to fly to the current position of the robber. In other words, they should not only *follow* the robber, but they should think ahead of where the robber might want to go, as the title of our paper indicates.

References

1. Dietmar Berwanger and Erich Grädel. *Entanglement – A Measure for the Complexity of Directed Graphs with Applications to Logic and Games*, pages 209–223. Springer Berlin Heidelberg, Berlin, Heidelberg, 2005.
2. Dietmar Berwanger, Erich Grädel, Łukasz Kaiser, and Roman Rabinovich. Entanglement and the complexity of directed graphs. *Theoretical Computer Science*, 463:2 – 25, 2012. Special Issue on Theory and Applications of Graph Searching Problems.

Improved Algorithms for k-Domination and Total k-Domination in Proper Interval Graphs*

Nina Chiarelli[1,2], Tatiana Romina Hartinger[1,2], Valeria Alejandra Leoni[3,4], Maria Inés Lopez Pujato[3,5], and Martin Milanič[1,2]

[1] University of Primorska, FAMNIT, Glagoljaška 8, SI-6000 Koper, Slovenia
nina.chiarelli@famnit.upr.si

[2] University of Primorska, IAM, Muzejski trg 2, SI-6000 Koper, Slovenia
tatiana.hartinger@iam.upr.si, martin.milanic@upr.si

[3] FCEIA, Universidad Nacional de Rosario, Rosario, Santa Fe, Argentina
valeoni@fceia.unr.edu.ar, mineslpk@hotmail.com

[4] CONICET, Argentina

[5] ANPCyT, Argentina

Among the many variants of the domination problems [5, 6], we consider in this work a family of generalizations of the classical domination and total domination problems known as k-domination and total k-domination. Given a positive integer k and a graph G, a *k-dominating set* in G is a set $S \subseteq V(G)$ such that every vertex $v \in V(G) \setminus S$ has at least k neighbors in S [4], and a *total k-dominating set* in G is a set $S \subseteq V(G)$ such that every vertex $v \in V(G)$ has at least k neighbors in S [9]. The k-domination and the total k-domination problems aim to find the minimum size of a k-dominating, resp. total k-dominating set, in a given graph. The k-domination and total k-domination problems are known to be NP-hard [7, 12] and also hard to approximate [3]. They are NP-hard not only for general graphs but also in the class of chordal graphs. More specifically, the problems are NP-hard in the class of split graphs [10, 12] and, in the case of total k-domination, also in the class of undirected path graphs [11]. We consider k-domination and total k-domination in another subclass of chordal graphs, the class of proper interval graphs. A graph G is an *interval graph* if it has an *interval model*, that is, a family $\mathcal{I}$ of intervals on the real line and a one-to-one correspondence between the vertices of G and the intervals of $\mathcal{I}$ such that two vertices are joined by an edge in G if and only if the corresponding intervals intersect. A *proper interval graph* is an interval graph that has an interval model in which no interval contains another one [13].

Recent results due to Kang et al. [8], building on previous works by Bui-Xuan et al. [2] and Belmonte and Vatshelle [1], imply that for each fixed integer $k \geq 1$, the k-domination and total k-domination problems are solvable in time $\mathcal{O}(n^{6k+4})$ in the class of interval graphs where n is the order of the input graph. In this work, we significantly improve the above result for the case of proper

* This work is supported in part by the Slovenian Research Agency (I0-0035, research program P1-0285 and research projects N1-0032, J1-6720, and J1-7051). The work for this paper was partly done in the framework of a bilateral project between Argentina and Slovenia, financed by the Slovenian Research Agency (BI-AR/15–17–009) and MINCYT-MHEST (SLO/14/09).

interval graphs. We show that for each positive integer k, the k-domination and total k-domination problems are solvable in time $\mathcal{O}(n^{3k})$ in the class of n-vertex proper interval graphs. Except for $k = 1$, this improves on the best known running times. Our approach is based on a reduction showing that for each positive integer k, the total k-domination problem on a given proper interval graph G can be reduced to a shortest path computation in a derived edge-weighted directed acyclic graph. A similar reduction works for k-domination. The reductions immediately result in algorithms with running time $\mathcal{O}(n^{4k+1})$. We show that with a suitable implementation the running time can be improved to $\mathcal{O}(n^{3k})$. The algorithms can be easily adapted to the weighted case, at no expense in the running time.

References

1. R. Belmonte and M. Vatshelle. Graph classes with structured neighborhoods and algorithmic applications. *Theoret. Comput. Sci.*, 511:54–65, 2013.
2. B.-M. Bui-Xuan, J. A. Telle, and M. Vatshelle. Fast dynamic programming for locally checkable vertex subset and vertex partitioning problems. *Theoret. Comput. Sci.*, 511:66–76, 2013.
3. F. Cicalese, M. Milanič, and U. Vaccaro. On the approximability and exact algorithms for vector domination and related problems in graphs. *Discrete Appl. Math.*, 161(6):750–767, 2013.
4. J. F. Fink and M. S. Jacobson. n-domination in graphs. In *Graph theory with applications to algorithms and computer science (Kalamazoo, Mich., 1984)*, Wiley-Intersci. Publ., pages 283–300. Wiley, New York, 1985.
5. T. W. Haynes, S. T. Hedetniemi, and P. J. Slater, editors. *Domination in graphs. Advanced topics*, volume 209 of *Monographs and Textbooks in Pure and Applied Mathematics*. Marcel Dekker, Inc., New York, 1998.
6. T. W. Haynes, S. T. Hedetniemi, and P. J. Slater. *Fundamentals of domination in graphs*, volume 208 of *Monographs and Textbooks in Pure and Applied Mathematics*. Marcel Dekker, Inc., New York, 1998.
7. M. S. Jacobson and K. Peters. Complexity questions for n-domination and related parameters. *Congr. Numer.*, 68:7–22, 1989. Eighteenth Manitoba Conference on Numerical Mathematics and Computing (Winnipeg, MB, 1988).
8. D. Y. Kang, O.-j. Kwon, T. J. F. Strømme, and J. A. Telle. A width parameter useful for chordal and co-comparability graphs. *Theoret. Comput. Sci.*, 704:1–17, 2017.
9. V. R. Kulli. On n-total domination number in graphs. In *Graph theory, combinatorics, algorithms, and applications (San Francisco, CA, 1989)*, pages 319–324. SIAM, Philadelphia, PA, 1991.
10. J. K. Lan and G. J. Chang. Algorithmic aspects of the k-domination problem in graphs. *Discrete Appl. Math.*, 161(10-11):1513–1520, 2013.
11. J. K. Lan and G. J. Chang. On the algorithmic complexity of k-tuple total domination. *Discrete Appl. Math.*, 174:81–91, 2014.
12. D. Pradhan. Algorithmic aspects of k-tuple total domination in graphs. *Inform. Process. Lett.*, 112(21):816–822, 2012.
13. F. S. Roberts. Indifference graphs. In *Proof Techniques in Graph Theory (Proc. Second Ann Arbor Graph Theory Conf., Ann Arbor, Mich., 1968)*, pages 139–146. Academic Press, New York, 1969.

WEB1 : Polyhedral Combinatorics

- The distance polytope for the vertex coloring problem
  Bruno Dias, Rosiane de Freitas, Nelson Maculan, Javier Marrenco.

- Computing the volume of the convex hull of the graph of a trilinear monomial using mixed volumes
  Emily Speakman, Gennadiy Averkov.

- A Polyhedral Study of the Connected Assignment Problem in Arrays
  Manoel Campelo, Joel Soares.

- Lovász-Schrijver PSD-operator on some graph classes defined by clique cutsets
  Annegret Wagler.

The distance polytope for vertex coloring

Bruno Dias, Rosiane de Freitas, Nelson Maculan, Javier Marenco

Instituto de Computação, Universidade Federal do Amazonas, Manaus, Brazil
PESC-COPPE, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil
Instituto de Ciencias, Universidad Nacional de General Sarmiento, Argentina

Abstract. In this work we consider the distance model for the classical vertex coloring problem, introduced by Delle Donne in 2009. This formulation involves decision variables representing the distance between the colors assigned to every pair of distinct vertices, thus not explicitly representing the colors assigned to each vertex. We show close relations between this formulation and the so-called orientation model for graph coloring. In particular, we prove that we can translate many facet-inducing inequalities for the orientation model polytope into facet-inducing inequalities for the distance model polytope, and viceversa.

Given a simple undirected graph $G = (V, E)$ and a set C of colors, the *vertex coloring problem* asks for a mapping $c : V \to C$ such that $c(i) \neq c(j)$ whenever $ij \in E$. Many integer programming formulations for this problem have been explored in the literature. In this work we are interested in the *distance formulation* [1] and its relations with the *orientation model* [2].

The orientation model involves an integer variable $z_i \in \{1, \ldots, |C|\}$ for each vertex $i \in V$ representing the color assigned to i, and an ordering variable y_{ij} for each edge $ij \in E$, $i < j$, in such a way that $y_{ij} = 1$ if and only if $z_i < z_j$. We refer the reader to [2] for details on this formulation. On the other hand, the *distance formulation* employs an integer variable x_{ij} for every $i, j \in V$, $i < j$, denoting the difference between the colors assigned to i and j (i.e., if i takes color $c(i)$ and j takes color $c(j)$, then $x_{ij} = c(i) - c(j)$), and, for every $ij \in E$, $i < j$, the variable y_{ij} is considered, with the same meaning as in the orientation model:

$$x_{ik} = x_{ij} + x_{jk} \qquad \forall i, j, k \in V,\ i < j < k \tag{1}$$
$$x_{ij} \geq 1 - |C| y_{ij} \qquad \forall (i, j) \in E,\ i < j \tag{2}$$
$$x_{ij} \leq -1 + |C|(1 - y_{ij}) \qquad \forall (i, j) \in E,\ i < j \tag{3}$$
$$x_i \in \{-|C| + 1, \ldots, |C| - 1\} \qquad \forall i \in V \tag{4}$$
$$y_{ij} \in \{0, 1\} \qquad \forall (i, j) \in E,\ i < j \tag{5}$$

The first set of constraints is composed by $O(|V|^3)$ equations, and may generate a large model in practice. The following result allows us to replace these contraints by $O(|V|^2)$ equations.

Theorem 1 **([1]).** *If $V = \{1, \ldots, n\}$, then constraints (1) are equivalent to*

$$x_{i,i+1} + x_{i+1,i+2} = x_{i,i+2} \qquad \forall i \in V, i \leq n-2 \tag{6}$$

$$x_{ij} + x_{i+1,j-1} = x_{i,j-1} + x_{i+1,j} \qquad \forall i,j \in V, i \leq n-3, i+3 \leq j \tag{7}$$

Let $PD(G,C)$ be the convex hull of the vectors (x,y) satisfying constraints (1)-(5). Theorem 1 allows us to calculate the dimension of this polytope.

Theorem 2. *If $|C| > \chi(G)+1$, then* $\dim(PD(G,C)) = |V| + |E| - 1$.

Call $PO(G,C)$ to the convex hull of feasible solutions $(z,y) \in \mathbb{Z}^{|V|+|E|}$ for the orientation model. The main results of this work show that certain facet-inducing inequalities for $PO(G,C)$ can be translated into $PD(G,C)$, and viceversa.

Theorem 3. *Let $\alpha z_i + \pi y \leq \alpha z_j + \pi_0$ be a valid (resp. facet-inducing) inequality for $PO(G,C)$, where $ij \in E$. Then, $\alpha x_{ij} + \pi y \leq \pi_0$ is valid (resp. facet-inducing if $|C| \geq \chi(G)+2$) for $PD(G,C)$.*

The construction in Theorem 3 can be generalized in a straightforward way to (facet-inducing) inequalities involving pairs of variables $\{(z_{i_k}, z_{j_k})\}_{k=1}^{p}$, where $i_k, j_k \in V$, $i_k \neq j_k$, for $k = 1, \ldots, p$, such that z_{i_k} and z_{j_k} appear with the same coefficient in both sides of the inequality. Many inequalities presented in [3] and [4] fit into this pattern, hence this result provides many facet-inducing inequalities for $PD(G,C)$. Finally, the following theorem provides a converse result, thus showing that all facets of $PD(G,C)$ come from facets of an orientation polytope.

Theorem 4. *Assume $C = \{1, \ldots, |C|\}$. Let $\gamma x + \pi y \leq \pi_0$ be a valid (resp. facet-inducing) inequality for $PD(G,C)$. Then, $\sum_{i \neq j} \gamma_{ij}(z_i - z_j) + \pi y \leq \pi_0$ is valid (resp. facet-inducing if $|C| \geq \chi(G)+2$) for $PD(G, C \cup \{|C|+1\})$.*

Theorem 3 and Theorem 4 imply that all facets of $PD(G,C)$ can be obtained from facets of $PO(G,C)$ and $PO(G, C \cup \{|C|+1\})$, although these polytopes are neither isomorphic nor combinatorially equivalent. This fact is not directly implied by the relation between the x- and the z-variables, but also relies on (simple) properties of vertex coloring. It would be interesting to study further pairs of polytopes with similar relations between their sets of variables, in order to gain more knowledge on these issues.

References

1. Delle Donne, D.: Un algoritmo branch&cut para un problema de asignación de frecuencias en redes de telefonía celular. Master's thesis, Departamento de Computación, FCEyN – Universidad de Buenos Aires (2009) (in Spanish).
2. Borndörfer, R., Eisenblätter, A., Grötschel, M., Martin, A.: The orientation model for frequency assignment problems. Technical report, ZIB Berlin (1998)
3. Marenco, J., Wagler, A.: Chromatic scheduling polytopes coming from the bandwidth allocation problem in point-to-multipoint radio access systems. Annals of Operations Research **150** (2007) 159–175
4. Marenco, J., Wagler, A.: Facets of chromatic scheduling polytopes based on covering cliques. Discrete Optimization **6** (2009) 64–78

Computing the volume of the convex hull of the graph of a trilinear monomial using mixed volumes

Gennadiy Averkov and Emily Speakman

Institute of Mathematical Optimization, Otto von Guericke Universität, Magdeburg

gennadiy.averkov@gmail.com emily.speakman@ovgu.de

1 Introduction

For difficult mixed integer non linear optimization (MINLO) problems, spatial branch-and-bound (sBB) is the algorithmic framework used to find globally optimal solutions (for example, see [1]). At a high level, sBB works by generating a convex relaxation of the problem over a given domain to obtain a bound, and then branching on the domain of a variable and re-convexifying to obtain better bounds. Therefore, the quality of the convexification obtained is very important in the success of the algorithm. Speakman and Lee ([4]) developed a method for comparing alternative convexifications of the graph of a trilinear monomial, $y = x_1x_2x_3$, over a nonnegative box $x_i \in [a_i, b_i],\ 0 \le a_i < b_i$. They calculated the 4-dimensional volume of various natural convexifications of the graph of $y = x_1x_2x_3$, and used this as a measure to compare the tightness of those convexifications. In particular, they computed the 4-dimensional volume of the convex hull. Define

$$\mathcal{P}_H^3 := \operatorname{conv}\left\{(y, x_1, x_2, x_3) \in \mathbb{R}^4 : y = x_1x_2x_3,\ \ x_i \in [a_i, b_i],\ i = 1, 2, 3\right\}.$$

The extreme points of $\mathcal{P}_H^3$ are the eight points that correspond to the $2^3 = 8$ choices of each x-variable at its upper or lower bound ([2]). Assuming we label the variables so that they are ordered by the ratios of their lower and upper bounds, i.e., $\frac{a_1}{b_1} \le \frac{a_2}{b_2} \le \frac{a_3}{b_3}$, the volume of $\mathcal{P}_H^3$ is given by the following:

Theorem 1 (see [4]).

$$\operatorname{Vol}(\mathcal{P}_H^3) = (b_1 - a_1)(b_2 - a_2)(b_3 - a_3) \times \left(b_1(5b_2b_3 - a_2b_3 - b_2a_3 - 3a_2a_3) + a_1(5a_2a_3 - b_2a_3 - a_2b_3 - 3b_2b_3)\right)/24.$$

2 Our contribution

In this work, we present an alternative method for the proof of Theorem 1. To do this, we note that the extreme points of P_H^3, lie in two parallel hyperplanes. Four points lie in the hyperplane $x_3 = a_3$ and four points lie in the hyperplane $x_3 = b_3$. In this way we can think of P_H^3 as the convex hull of two (3 dimensional) tetrahedra, let these be P and Q. The volume of P_H^3 can therefore be calculated via an integral as x_3 varies from a_3 to b_3.

$$\begin{aligned}\mathrm{Vol}(\mathcal{P}_H^3) &= \int_{a_3}^{b_3} \mathrm{Vol}\left(\frac{b_3-t}{b_3-a_3}P + \frac{t-a_3}{b_3-a_3}Q\right) dt \\ &= (b_3-a_3)^{-3}\int_{a_3}^{b_3} \mathrm{Vol}((b_3-a_3)P + (t-a_3)Q)dt \\ &= (b_3-a_3)^{-3}\int_{a_3}^{b_3} (b_3-t)^3 \,\mathrm{Vol}(P) + 3(b_3-t)^2(t-a_3)V(P,P,Q) \\ &\qquad + 3(b_3-t)(t-a_3)^2 V(P,Q,Q) + (t-a_3)^3 \,\mathrm{Vol}(Q)dt,\end{aligned}$$

where $V(P,P,Q)$ and $V(P,Q,Q)$ are so-called mixed volumes; for more information see [3]. Given that they are tetrahedra, the volumes of P and Q are easy to compute. The technical detail of the proof lies in using the appropriate formulae to compute $V(P,P,Q)$ and $V(P,Q,Q)$ and integrating. In doing this, we obtain the formulae of [4].

3 Possible extensions

A natural extension of this problem is to allow the variables to have mixed-sign domains. Our alternative proof method restricts the sections of the proof that depend on the sign of the bounds, and therefore, this method may make the extension to mixed-sign domains easier. Additionally, it would be interesting to be able to compute the volume of the convex hull of the graph of a general multilinear term (over a box). In the work of [4], the jump to the case of general n seemed to be computationally unrealistic. Now, with this alternative method, it remains unclear that the leap will be possible, but a natural approach for this extension is more obvious. In the case of general n, we can write down the volume as:

$$\mathrm{Vol}(\mathcal{P}_H^n) = \int_{a_n}^{b_n} \mathrm{Vol}\left(\frac{b_n-t}{b_n-a_n}P_{n-1} + \frac{t-a_n}{b_n-a_n}Q_{n-1}\right) dt,$$

where each of the polytopes P_{n-1} and Q_{n-1}, are no longer tetrahedra, but *are* closely related to P_H^{n-1} (the extreme points of both are the extreme points of P_H^{n-1} with a scaling applied to the first component). Being able to express the volume of $\mathcal{P}_H^n$ in terms of the volume of $\mathcal{P}_H^{n-1}$ is hopefully be the first step in obtaining more general results.

References

[1] Adjiman, C., Dallwig, S., Floudas, C., Neumaier, A.: A global optimization method, αBB, for general twice-differentiable constrained NLPs. Computers & Chem. Engin. **22**(9) (1998) 1137–1158

[2] Rikun, A.: A convex envelope formula for multilinear functions. Journal of Global Optimization **10** (1997) 425–437

[3] Schneider, R.: Convex bodies: the Brunn-Minkowski theory. 2nd edn. Volume 151 of Encyclopedia of Math. and its Apps. Cambridge University Press (2014)

[4] Speakman, E., Lee, J.: Quantifying double McCormick. Mathematics of Operations Research **42**(4) (2017) 1230–1253

A Polyhedral Study of the Connected Assignment Problem in Arrays (short version)

Manoel Campêlo and Joel Soares

Universidade Federal do Ceará, Campus do Pici, Bl. 910, 60440-554 Fortaleza, Brazil
{mcampelo,joelcruz}@lia.ufc.br

Abstract. We study the Connected Assignment Problem in Arrays. We present a large family of facet-defining inequalities. In particular, it comprises all facet-defining inequalities with righthand side 1. We prove that it gives a complete description of the polytope in some cases. We computationally evaluate its strength for tightening the linear relaxation.

Keywords: Connected Assignment Problem, Polyhedral Combinatorics.

1 Introduction

Motivated by Radio Resource Allocation in mobile networks [1, 2], we define the Connected Assignment Problem in Arrays (CAPA) as follows. Let us be given a set of symbols $I = \{1, 2, \ldots, M\}$, a set of indices $J = \{1, 2, \ldots, N\}$ of the entries of an array and a gain function $\rho : I \times J \to \mathbb{Q}_+$ such that $\rho(i, j)$ (or simply ρ_{ij}) describes the utility value of assigning i to j. A *connected assignment* is an assignment of exactly one symbol from I to each position indexed by J such that repeated symbols appear consecutively in the array. Precisely, it is a function $A : J \to I$ such that, for all $i \in I$, there is an interval $[a_i, b_i] \subseteq J$ (possibly empty) where $A(j) = i$ if, and only if, $j \in [a_i, b_i]$. The gain of an assignment A is $\rho(A) = \sum_{j \in J} \rho(A(j), j)$. Problem CAPA consists in finding a connected assignment of maximum gain.

2 The pattern-based polytope

Let us call *pattern* a nonempty interval of the set $J = \{1, 2, \ldots, N\}$ of positions. Let us denote P the set of all patterns. Note that $|P| = N(N + 1)/2$.

In CAPA, each symbol $i \in I$ is either not allocated or assigned to a pattern $p \in P$ thus returning a gain of $\rho_i(p) = \sum_{j \in p} \rho_{ij}$. Thus, using a binary variable $x_{ip} \in \{0, 1\}$ to indicate whether i is assigned or not to p, we get the model:

$$\max \sum_{i \in I} \sum_{p \in P} \rho_i(p) x_{ip} \tag{1}$$

$$\text{s.t.} \sum_{p \in P} x_{ip} \le 1, \forall i \in I, \qquad \sum_{i \in I} \sum_{p \in P : j \in p} x_{ip} \le 1, \forall j \in J, \tag{2}$$

$$x_{ip} \in \{0, 1\}, \forall i \in I, \forall p \in P. \tag{3}$$

Constraints (2) ensure that each symbol can be assigned according to at most one pattern and each position is occupied by at most one symbol.

We denote $\mathcal{P}(M,N)$ the polytope defined by the convex hull of the points satisfying (2)-(3). If M and N are not relevant in the context, we simply use $\mathcal{P}$.

Let $P_\cap(p) = \{p' \in P : p \cap p' \neq \emptyset\}$ and $P_\supset(p) = \{p' \in P : p' \supseteq p\}$. The results below hold for $\mathcal{P}$.

Theorem 1. *For $p \in P$ and $i \in I$, the inequality*

$$\sum_{i' \in I \setminus \{i\}} \sum_{p' \in P_\supset(p)} x_{i'p'} + \sum_{p' \in P_\cap(p)} x_{ip'} \leq 1 \qquad (4)$$

is valid for $\mathcal{P}$, and it is facet-defining if $M \geq 2$. For $M = 2$ or $N = 2$, $\mathcal{P}(M,N) = \{x \geq 0 : x$ satisfies (4) $\forall i \in I, \forall p \in P\}$.

We can show that the family of inequalities (4) comprises all facet-defining inequalities on integer coefficients with righthand side 1. On the other hand, this family can be generalized to include facet-defining inequalities with all possible righthand sides in the range $1 \ldots \min(M, N-1)$, as follows. For a nonempty subset $I' \subseteq I$ of symbols and p be a pattern, let us consider the inequality

$$\sum_{i' \in I'} \sum_{p' \in P_\cap(p)} \max(\delta_{p'}, 1) x_{i',p'} + \sum_{i' \in I \setminus I'} \sum_{p' \in P_\cap(p)} \max(\delta_{p'}, 0) x_{i',p'} \leq |I'|, \qquad (5)$$

where $\delta_{p'} = |I'| - |p \setminus p'|$, for all $p' \in P$. Observe that, if we restrict I' to be a singleton, we get inequality (4).

Theorem 2. *Let $I' \subseteq I$, $I \neq \emptyset$, and $p \in P$. Inequality (5) is valid for $\mathcal{P}$ if, and only if, $|I'| \leq |p|$ or $|p| = 1$. For $M \geq 3$, it defines a facet of $\mathcal{P}$ if, and only if, $|I'| \leq |p| - 1$ or $|p| = 1$.*

3 Computational Experiments

We carried out some computational experiments to evaluate the strenght of the derived facet-defining inequalities. We used 300 randomly generated instances with $N \in \{20, 40\}$ and $M \in \{kN/5 : k = 1, \ldots, 5\}$. The percentage gap of the linear relaxation is generally small but nonzero in all instances. Adding inequalities (4) reduces it in 18% in average. Moreover, the relaxed optimum solution becomes integer in 12.3% of the cases. Adding inequalities (5) with $|I'| = 2$ could further reduce the gap at much higher computational cost.

References

1. F. Lima, T. Maciel, and F. Cavalcanti. Radio resource allocation in SC-FDMA uplink with resource adjacency constraints. *J Commun Inf Syst*, 31:272–289, 2016.
2. I. Wong, O. Oteri, and W. McCoy. Optimal resource allocation in uplink SC-FDMA systems. *Trans. Wireless. Comm.*, 8(5):2161–2165, 2009.

Lovász-Schrijver PSD-operator on some graph classes defined by clique cutsets

Annegret Wagler
annegret.wagler@uca.fr

LIMOS (UMR 6158 CNRS), University Clermont Auvergne,
Clermont-Ferrand, France

In this work, we study the stable set polytope, some of its linear and semi-definite relaxations, and graph classes for which certain relaxations are tight.

The *stable set polytope* STAB(G) of a graph G is defined as the convex hull of the incidence vectors of all stable sets of G. Two canonical relaxations of STAB(G) are

$$\begin{aligned} \text{ESTAB}(G) &= \{\mathbf{x} \in [0,1]^V : x_i + x_j \ \le \ 1, ij \in E\} \\ \text{QSTAB}(G) &= \{\mathbf{x} \in [0,1]^V : x(Q) = \textstyle\sum_{i \in Q} x_i \ \le \ 1, \ Q \subseteq V \text{ clique}\}, \end{aligned}$$

where STAB(G) equals ESTAB(G) for bipartite graphs, and QSTAB(G) for perfect graphs only [3]. Lovász and Schrijver introduced in [5] the PSD-operator LS_+ which, applied to ESTAB(G), generates a positive semi-definite relaxation $LS_+(G)$ of STAB(G). Results of Lovász and Schrijver imply that

$$\text{STAB}(G) \subseteq LS_+(G) \subseteq \text{ASTAB}^*(G) \tag{1}$$

where ASTAB$^*(G)$ is a linear relaxation of STAB(G) given by *joined antiweb constraints*

$$\sum_{i \le k} \frac{1}{\alpha(A_i)} x(A_i) + x(Q) \le 1, \tag{2}$$

associated with the complete join of some antiwebs $A_1, \ldots, A_k$ and a clique Q, where an *antiweb* A_n^k is a graph with n nodes $0, \ldots, n-1$ and edges ij if and only if $k \le |i-j| \le n-k (\text{mod} n)$ and $i \ne j$.

Graphs G with STAB$(G) = LS_+(G)$ are called *LS_+-perfect*, and the following conjecture has been proposed in [1]:

Conjecture 1 (LS_+-Perfect Graph Conjecture). G is LS_+-perfect if and only if $LS_+(G) = \text{ASTAB}^*(G)$.

Note that graphs G with STAB$(G) = \text{ASTAB}^*(G)$ are called *joined a-perfect*. By (1), we have that all joined a-perfect graphs are LS_+-perfect. Subclasses of joined a-perfect graphs include, besides perfect graphs, t-perfect, h-perfect, a-perfect graphs as well as near-bipartite graphs. Moreover, we can easily see from the above remarks that the conjecture states that LS_+-perfect graphs coincide with joined a-perfect graphs. Conjecture 1 has been already verified for several graph classes, e.g., near-perfect graphs, webs, line graphs, and claw-free graphs.

Our aim is to verify Conjecture 1 for further graph classes and to identify further subclasses of joined a-perfect and LS_+-perfect graphs. For that, we study graph classes where *clique cutsets* play a role in a decomposition theorem: pseudothreshold graphs and graphs without certain Truemper configurations. We describe the facet-defining system of the stable set polytopes for each of those basic families and then apply the result of Chvátal [3] that the facets of STAB(G) belong to the union of the facets of the stable set polytopes of the blocks of the decomposition.

Chvátal and Hammer [4] characterize a graph $G = (V, E)$ as *pseudothreshold* if and only if there is a partition $V = S \cup Q \cup U$ such that

- S is stable, and there are no edges between S and U,
- Q is a clique, and there are all edges between Q and U,
- U does not contain a stable set of size 3.

We present the facet-defining inequalities of the stable set polytope STAB(G) for G pseudothreshold. As a consequence, we can verify the LS_+-Perfect Graph Conjecture for pseudothreshold graphs. Moreover, we define strongly pseudothreshold graphs as further subclass containing all pseudothreshold graphs G such that also the complement $\overline{G}$ is pseudothreshold and show that strongly pseudothreshold graphs are joined a-perfect.

Boncompagni et al. [2] define $\mathcal{G}_U$ as the class of all graphs which either are

- a *light clique* obtained from a clique by removing a (possibly empty) matching,
- a *fat universal wheel* obtained as complete join of a hole C_k with $k \geq 5$ and a (possibly empty) clique,

or have a clique cutset. Based on this result, we give a complete description of the stable set polytope for graphs in $\mathcal{G}_U$ and conclude that every graph in $\mathcal{G}_U$ is joined a-perfect

Finally, we discuss the relations of the studied graph classes, revealing that strongly pseudothreshold graphs form a subclass of $\mathcal{G}_U$ and that $\mathcal{G}_U$ is a new subclass of joined a-perfect graphs, being incomparable to all such classes known so far.

References

1. S. Bianchi, M. Escalante, G. Nasini, and L. Tunçel, "Lovász-Schrijver PSD-operator and a superclass of near-perfect graphs", *Electronic Notes in Discrete Mathematics*, vol. 44, pp. 339–344, 2013.
2. V. Boncompagni, I. Penev, K. Vuşković, "Clique-cutsets beyond chordal graphs", arXiv:1707.03252 [math.CO]
3. V. Chvátal, "On Certain Polytopes Associated with Graphs," *J. Combin. Theory (B)*, vol. 18, pp. 138–154, 1975.
4. V. Chvátal, P.L. Hammer, "Aggregation of inequalities in integer programming", *Annals of Discrete Math.*, vol. 1, pp. 145–162, 1977.
5. L. Lovász and A. Schrijver, "Cones of matrices and set-functions and 0-1 optimization," *SIAM J. on Optimization*, vol. 1, pp. 166–190, 1991.

WEB2 : Mixed Integer Programming I

- Team Formation on Social Networks
  Nihal Berktas, Hande Yaman.

- A New Mixed Integer Linear Program for the Graph Edit Distance Problem
  Mostafa Darwiche, Romain Raveaux, Donatello Conte, Vincent T'Kindt.

- Finding Minimum Stopping and Trapping Sets: An Integer Linear Programming Approach
  Alvaro Velasquez, K. Subramani, Steven Drager.

- A branch-and-bound procedure for the robust cyclic job shop problem
  Idir Hamaz, Laurent Houssin, Sonia Cafieri.

Team Formation on Social Networks

Nihal Berktaş and Hande Yaman

Bilkent University, Ankara, Turkey

Abstract. People work in teams to complete a job or project that requires a number of skills. The success of a team depends on the technical capabilities of people as well as the quality of the communication among them. We study the team formation problem in which communication is taken into consideration by assuming the potential members constitute a social network. The aim is to construct a capable team with minimum communication cost. We formulate the problem as an integer program and develop a branch and bound algorithm which are tested using different social networks.

Keywords: team formation, integer programming, branch and bound

1 Introduction

The complexity of products and services in today's world requires many skills, knowledge and experience from different fields while the pace of consumption demands agility in the production and development phases. To be able to meet these requirements, we see people working in teams in various organizations. The quality of work done depends on the technical capabilities of people as well as the quality of the communication among them. We define *Team Formation Problem* (TFP) as finding a group of people who posses a required set of skills and can function as a team. To measure the functionality of the team in terms of communication, we assume that the candidates constitute a social network from which the communication quality can be obtained.

In a social network, nodes correspond to candidates and edges among them represent the relations. The weight of an edge can be interpreted as the effort required for those people to communicate effectively as team members. In other words, it is the cost of communication between the people connected by that edge. The existence of such a network is an acceptable assumption because either such network really exists, e.g. LinkedIn (www.linkedin.com) and Stack Overflow (www.stackoverflow.com), or people are part of an actual organization and their relations can be modeled via a social network.

The first study of TFP on social networks was by Lappas et al. [1]. The authors define two different communication cost functions: first as the diameter of the subgraph induced by the team members and second as the cost of the minimum spanning tree on this subgraph. Algorithms are developed for both versions of the problems. Following this work, the problem is studied with different cost definitions and settings, all of which is compared by Wang et al. [2].

This comparative study concluded that it is a more robust approach to define the communication cost as the sum of distances between each team member since the other cost functions are very sensitive the changes in the team. Minimizing sum of distances amounts to maximizing the average familiarity of the team which has a positive effect on the performance as discussed in [3] therefore we use this communication cost function in our study.

2 Solution Methods

We formulated TFP first as a quadratic set covering problem and then the objective function is linearized by defining a new set of binary variables. The models are implemented in Java using Cplex 12.7 and run on a a 64-bit machine with Intel Xeon E5-2630 v2 processor at 2.60 GHz and 96 GB of RAM. Our preliminary analysis show that with both of the quadratic and mixed integer programming formulations the optimal solution can be obtained within few minutes for the instances from IMDB network which has 1021 nodes. We construct a collaboration network with 12855 nodes using DBLP database and for some instances from this larger network the state-of-the-art solver fails to build the mixed integer model. To be able to reach the optimal solutions in reasonable time for the large instances, we develop a branch and bound algorithm. First a new set of constraints are obtained from the set covering constraint applying the general procedure proposed in [4]. We add these new constraints to the integer program and then we relax some of the constraints which connect the binary variables so that the problem decomposes. In our branch and bound algorithm we solve $n+1$ set covering problems in the root node where n is the number nodes in the network. In other nodes of the tree, at most 3 set covering problems are required to be solved.

3 Conclusion

To the best of our knowledge, this is the first study where exact solution techniques are developed for the team formation problem in the presence of a social network. A novel branch and bound scheme which can be generalized to other combinatorial optimization problems is developed and shown to be efficient.

References

1. Lappas, T., Liu, K., Terzi, E.: Finding a team of experts in social networks. In Proceedings of the 15th ACM SIGKDD 467–476 (2009)
2. Wang, X., Zhao, Z., Ng, W.: Ustf: A unified system of team formation. IEEE Transactions on Big Data, 2(1), 70–84 (2016)
3. Huckman, R. S., Staats, B. R., Upton, D. M.: Team familiarity, role experience, and performance: Evidence from Indian software services. Management science, 55(1), 85–100 (2009)
4. Adams, W. P. and Sherali, H. D.: A tight linearization and an algorithm for zero-one quadratic programming problems. Management Science, 32(10), 1274–1290 (1986)

A New Mixed Integer Linear Program for the Graph Edit Distance Problem

Mostafa Darwiche[1,2], Romain Raveaux[1], Donatello Conte[1], and Vincent T'kindt[2]

[1] Laboratoire d'Informatique Fondamentale et Appliquée (LIFAT, EA6300), Université François Rabelais Tours, France
[2] Laboratoire d'Informatique Fondamentale et Appliquée (LIFAT, EA6300), ERL-CNRS 6305, Université François Rabelais Tours, France
{mostafa.darwiche,romain.raveaux,donatello.conte,tkindt}@univ-tours.fr

Keywords: Graph Edit Distance, Graph Matching, Mixed Integer Linear Program.

1 Introduction

When talking about structural representation of objects and patterns, one could consider graph-based representation, which has been proven in the past decades to be efficient and convenient in many fields. A graph consists of two sets of vertices and edges, where vertices depict the main components of the objects and edges draw the relationships between them. Also, a group of numerical or nominal values can be assigned to vertices and edges in order to provide more information and characteristics. Such values are referred to as attributes/labels. Throughout the years, the attention towards using graphs to model objects have grown in many fields such as *Pattern Recognition* and *Chemionformatics* [4]. The problematic then occurs when having two graphs, how to compare and measure the (dis)similarities between them? Such question has intrigued many researchers who have come up with different classes of problems that are all *Graph Matching problems.* The Graph Edit Distance (GED) problem, which is one of them, provides a dissmilarity measure between two graphs and belongs to *Error-tolerant graph matching* class of problems in particular. The GED problem defines a set of edit operations, which are substitution, insertion and deletion of a vertex or edge where each operation has an associated cost. Solving the problem consists in finding the set of edit operations that transform one graph into another while minimizing the total cost. Let $G = (V, E, \mu, \xi)$ and $G' = (V', E', \mu', \xi')$ be two graphs, with μ and ζ the functions to assign attributes for vertices/edges. The optimal solution of the GED problem is the set of operations $\lambda(G, G') = \{o_1, ..., o_k\}$ with o_i an elementary vertex/edge edit operation and k the number of operations with the minimum cost. This problem has been proved to be NP-hard [5] and numerous heuristics can be found in the literature to solve it. However, only two *mixed integer linear programs* (MILP) exist in the literature [1, 2]. The intent of this work is to propose a new MILP formulation.

2 New MILP formulation for the GED problem

The proposed MILP formulation is inspired from the formulation presented in [2], referred to as (F2). It defines two sets of binary variables: variables $x_{i,k}$ represent the substitution of two vertices $u_i \in V$ and $v_k \in V'$, variables $y_{ij,kl}$ represent the substitution of two edges $e_{ij} \in E$ and $f_{kl} \in E'$. The number of variables in total is $(|V| \times |V'|) + (|E| \times |E'| \times 2)$ for undirected graphs, where y variables are doubled, in comparison with (F2), by considering $y_{ij,kl}$ and $y_{ij,lk}$ for every two edges e_{ij} and f_{kl}. Another variation from (F2) is the constraint that preserves the topology of the graphs, where a new constraint is introduced that only depends on the number of vertices in the new formulation. The number of constraints is then $|V'| + |V| + (|V| \times |V'|)$, against $|V| + |V'| + (|V'| \times |E|)$ constraints in (F2). Two assumptions are made here: the new formulation has a number of constraints independent from the number of edges of the graphs, which should logically lead to a better formulation than (F2) especially in the case of dense (highly connected) graphs. The second assumption is that even having more variables and reducing the number of constraints, the new formulation will perform better than (F2). These assumptions can only be validated through experiments. So far, the new formulation is tested against (F2) on CMU-House graph database [3] of medium graph sizes. Over 660 instances, the new formulation was able to solve 333 instances to optimality against 25 instances by (F2). Both formulations were solved by CPLEX 12.6.0 with 900 seconds as time limit. This preliminary result is promising and shows that the two assumptions hold for this graph database. More graph databases will be considered to evaluate both formulations with different graph sizes and structures in order to confirm the assumptions and the effectiveness of the proposed formulation. The obtained results will be presented at the conference.

References

1. Lerouge, J., Abu-Aisheh, Z., Raveaux, R., Héroux, P., Adam, S.: Exact graph edit distance computation using a binary linear program. In: Joint IAPR International Workshops on Statistical Techniques in Pattern Recognition (SPR) and Structural and Syntactic Pattern Recognition (SSPR). pp. 485–495. Springer (2016)
2. Lerouge, J., Abu-Aisheh, Z., Raveaux, R., Héroux, P., Adam, S.: New binary linear programming formulation to compute the graph edit distance. Pattern Recognition 72, 254–265 (2017), https://doi.org/10.1016/j.patcog.2017.07.029
3. Moreno-García, C.F., Cortés, X., Serratosa, F.: A graph repository for learning error-tolerant graph matching. In: Joint IAPR International Workshops on Statistical Techniques in Pattern Recognition (SPR) and Structural and Syntactic Pattern Recognition (SSPR). pp. 519–529. Springer (2016)
4. Sanfeliu, A., Alquézar, R., Andrade, J., Climent, J., Serratosa, F., Vergés, J.: Graph-based representations and techniques for image processing and image analysis. Pattern recognition 35(3), 639–650 (2002)
5. Zeng, Z., Tung, A.K., Wang, J., Feng, J., Zhou, L.: Comparing stars: On approximating graph edit distance. Proceedings of the VLDB Endowment 2(1), 25–36 (2009)

Finding Minimum Stopping and Trapping Sets: An Integer Linear Programming Approach

Alvaro Velasquez[1], K. Subramani[2], Steven L. Drager[3]

Department of Computer Science, University of Central Florida, Orlando, FL[1]
LCSEE, West Virginia University, Morgantown, WV[2]
Information Directorate, Air Force Research Laboratory, Rome, NY[3]
velasquez@cs.ucf.edu, ksmani@csee.wvu.edu, steven.drager@us.af.mil

In this paper, we discuss the problems of finding minimum stopping sets and trapping sets in Tanner graphs, using integer linear programming. These problems are important for establishing reliable communication across noisy channels. Indeed, stopping sets and trapping sets correspond to combinatorial structures in information-theoretic codes which lead to errors in decoding once a message is received. In this paper, we make two contributions. First, we propose integer linear programming solutions for finding stopping sets and several trapping set variants. We improve on the results of [1], which demonstrate that the minimum stopping set in the $(4896, 2474)$ Margulis code [2] is of size 24 and that no stopping sets of size 25 and 26 exist. Our results establish that the next largest stopping sets in said code are of sizes 36 and 48 (See Figures 10 and 11 in https://www.cs.ucf.edu/~velasquez/StoppingSets/). As a point of reference, the number of points in the search spaces for finding stopping sets of sizes 26, 36, and 48 are $\binom{4896}{26} \approx 2 \times 10^{69}$, $\binom{4896}{36} \approx 1.61 \times 10^{91}$, and $\binom{4896}{48} \approx 8.28 \times 10^{115}$, respectively. The second contribution we make pertains to the previously unknown complexities of two trapping set variants. We prove that these variants are **NP-hard**, thereby rounding out the complexity results in the literature.

Stopping sets and trapping sets are defined by simple combinatorial structures in the graph representation of the underlying code. Due to space constraints, we will focus on stopping sets in this brief. However, trapping sets are similarly defined. An (n, k) code is one whose codewords are bit-vectors of length n and whose dimension is k. The dimension k of the code specifies the number of linearly independent codewords that form the basis $\underline{c}_1, \ldots, \underline{c}_k \in \{0, 1\}^n$ for the code. That is, any codeword can be expressed as a linear combination of these basis vectors. Any given codeword of length n contains k original bits of information and $n - k$ redundant check bits that are used to detect and correct errors that arise during message transmission across a noisy channel.

Given an (n, k) code, its corresponding parity-check matrix $H \in \{0, 1\}^{(n-k)\times n}$ defines the linear relations among codeword variables. Each column in H corresponds to a bit in the codeword and each row corresponds to a redundant check bit. Given a codeword $\underline{c} = \underline{c}_1, \ldots, \underline{c}_n$, the entry H_{ij} is 1 if $\underline{c}_j$ is involved in a check operation, which is used to detect errors after transmission. For any such matrix H, let $G = (V \cup C, E)$ denote its representation as a bipartite graph, where $V = \{v_1, \ldots, v_n\}$ and $C = \{c_1, \ldots, c_m\}$ are the sets of variable and check nodes, and $E = \{(v_i, c_j)|H_{ji} = 1\}$ defines the adjacency set. This is known as the Tanner graph of a code. We can now define the stopping and trapping set problems.

Definition 1 (Stopping set). *Given a Tanner graph $G = (V \cup C, E)$, a stopping set $S \subseteq V$ is a set of variable nodes such that all neighbors of nodes in S are connected to S at least twice.*

As an example, suppose we are given the parity-check matrix H below with Tanner graph $G = (V \cup C, E)$. We can determine the minimum stopping set $S = \{v_7, v_9\}$ as pictured in Figure 1. A comparison of our approach against methods in the literature can be seen in Table 1.

$$H = \begin{pmatrix} 0&0&1&1&0&0&1&1&1&0 \\ 1&0&0&1&0&1&0&0&0&1 \\ 1&0&1&0&1&0&0&1&0&1 \\ 0&1&0&0&1&1&1&0&1&0 \\ 0&1&0&0&0&0&1&1&1&1 \end{pmatrix}$$

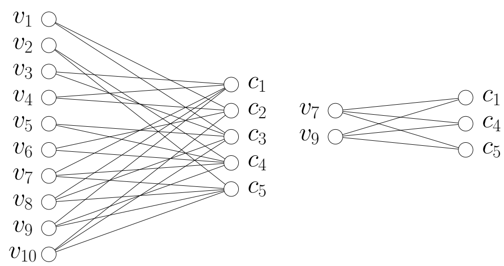


Fig. 1. (*left*) Parity-check matrix H for some code. (*center*) Tanner graph $G = (V \cup C, E)$ of H. (*right*) Subgraph G_S induced by the minimum stopping set $S = \{v_7, v_9\}$ in G.

Minimum Stopping Set Size

Code	[3]	[4]	[1]	Us
$(504, 252)$ Mackay	16 (N/A)	16 (600 hours)	16 (N/A)	16 (37 seconds)
$(504, 252)$ PEG	N/A (N/A)	19 (25 hours)	19 (N/A)	19 (365 seconds)
$(1008, 504)$ Mackay	28 (N/A)	26 (3085 hours)	N/A (N/A)	26 (18.73 hours)
$(4896, 2474)$ Margulis	24 (N/A)	24 (162 hours)	24 (N/A)	24 (267 seconds)

Table 1. The size of minimum stopping sets in 4 popular codes are presented based on the results of various methods in the literature. The numbers in parentheses denote the execution time to find said sets. See figures 6 through 9 in https://www.cs.ucf.edu/~velasquez/StoppingSets/ for a visualization of the sets found by our approach.

References

1. Eirik Rosnes and Øyvind Ytrehus. An efficient algorithm to find all small-size stopping sets of low-density parity-check matrices. *IEEE Transactions on Information Theory*, 55(9):4167–4178, 2009.
2. J Rosenthal and PO Vontobel. Construction of ldpc codes based on ramanujan graphs and ideas from margulis. In *Proc. 38th Annual Allerton Conf. on Communication, Computing and Control*, pages 248–257.
3. Xiao-Yu Hu and Evangelos Eleftheriou. A probabilistic subspace approach to the minimal stopping set problem. In *Turbo Codes&Related Topics; 6th International ITG-Conference on Source and Channel Coding (TURBOCODING), 2006 4th International Symposium on*, pages 1–6. VDE, 2006.
4. Masanori Hirotomo, Yoshiho Konishi, and Masakatu Morii. On the probabilistic computation algorithm for the minimum-size stopping sets of ldpc codes. In *Information Theory, 2008. ISIT 2008. IEEE International Symposium on*, pages 295–299. IEEE, 2008.

A branch-and-bound procedure for the robust cyclic job shop problem

Idir Hamaz[1], Laurent Houssin[1] and Sonia Cafieri[2]

[1] LAAS-CNRS, Universite de Toulouse, CNRS, UPS, Toulouse, France
{ihamaz, lhoussin}@laas.fr
[2] ENAC, Universite de Toulouse, F-31055 Toulouse, France
sonia.cafieri@enac.fr

This paper deals with the cyclic job shop problem where the task durations are uncertain and belong to a polyhedral uncertainty set. We formulate the cyclic job shop problem as a two-stage robust optimization model. The cycle time and the execution order of tasks executed on the same machines correspond to the *here-and-now* decisions and have to be decided before the realization of the uncertainty. The starting times of tasks corresponding to the *wait-and-see* decisions are delayed and can be adjusted after the uncertain parameters are known. In the last decades, different solution approaches have been developed for two-stage robust optimization problems. Among them, the use of affine policies, column generation algorithms, row and row-and-column generation algorithms. In this paper, we propose a Branch-and-Bound algorithm to tackle the robust cyclic job shop problem with the worst case cycle time minimization. The algorithm uses, at each node of the search tree, a robust version of the Howard algorithm to derive a lower bound on the optimal cycle time. We also develop a heuristic method that permits to compute an initial upper bound for the cycle time. Finally, encouraging preliminary results of numerical experiments performed on randomly generated instances are presented.

WEB3 : Models and Algorithms for Network Optimization problems

- A p-Median Based Exact Method for the Large-Scale Optimal Diversity Management Problem
  Adriano Masone, Antonio Sforza, Claudio Sterle, Igor Vasilyev, Anton Ushakov.

- On the Optimal Cost of the Transportation Problem with Interval Right-Hand-Sides
  Raffaele Cerulli, Ciriaco D'Ambrosio, Monica Gentili.

- An exact algorithm for the split-demand one-commodity pickup-and-delivery travelling salesman problem
  Hipolito Hernandez Perez, Juan José Salazar González.

- A MILP Formulation for a Multi-Tiered Vehicle Routing Problem with Global Cross-Docking
  Anthony Smith, Paolo Toth, Jan H. van Vuuren.

A p-Median Based Exact Method for the Large-Scale Optimal Diversity Management Problem

A. Masone[1], A. Sforza[1], C. Sterle[1], A. Ushakov[2] and I. Vasilyev[2]

[1]Department of Electrical Engineering and Information Technology - Optimization and Problem Solving Laboratory, University "Federico II" of Naples, 80125 Naples, Italy
[2]Matrosov Institute for System Dynamics and Control Theory, Siberian Branch of the Russian Academy of Sciences, Lermontov str. 134, Irkutsk 664033, Russia

{adriano.masone, antonio.sforza, claudio.sterle}@unina.it
{aushakov, vil}@icc.ru

Abstract. The p-median problem (PMP) is the well known network optimization problem of discrete location theory. In many real applications PMP is defined on large scale networks, for which ad-hoc exact and/or heuristic methods have to be developed. A very interesting industrial application is constituted by the optimal diversity management problem (ODMP) which arises when a company producing a good and/or a service (which can be customized with options) needs to satisfy many client demands with various subset of options, but only a limited number of option configurations can be produced. Exploiting a suitable network representation, ODMP can be formulated as a PMP on a large-scale disconnected network. In this paper we revise and improve a decomposition approach where a lot of smaller PMPs related to the network components can be solved instead of the initial large-scale problem. Proposed approach drastically reduces number and dimension of these subproblems, solving them to optimality by a MIP solver, and combining their solutions to find the optimal solution of the original problem, formulated as a multiple choice knapsack problem. The computational tests show that our approach is able to find optimal solutions of known and new test instances, considerably outperforming state-of-the-art approaches to the large-scale PMP on disconnected networks.

Keywords: p-median, optimal diversity management, decomposition approach.

1 Problem definition and proposed method

The optimal diversity management problem (ODMP) is a well-known optimization problem arising in many application fields, i.e. every time a company produces a good and/or a service which can be provided with options. In this case the product can be personalized by the customer that can choose different option combinations (configurations) depending on her/his needs or preferences. In this context, satisfying all the possible demands with exactly the required options would impose the company to produce in advance all the possible configurations at the assembly lines. Moreover, the production operations could start only after a demand is received, so providing huge delays in satisfying the request. To overcome these drawbacks, the company usually produces a limited number of opportunely chosen configurations to cover all the possible ones. In this way a demand, if no available configuration covers all the required options, can be satisfied by a compatible configuration, i.e. by a configuration containing all the required options plus some others not demanded by the customer. This implies that a client could receive some not demanded options, so generating an over-cost for the company.

The ODMP consists in choosing a subset of configurations to cover all the customer demands, minimizing the total over-cost. It was introduced in [1] for a car industry application. Indeed, a car could be equipped with a large number of options, each of them requiring the installation of related electrical wiring. The total number of options can be up to forty and possible electrical wiring configurations can be many thousands, but obviously only a limited number of them can be made available at the assembly line. For this reason, a not available wiring configuration has to be replaced by another one containing all the needed wirings and some others not required.

ODMP can be formulated in terms of p-median problem (PMP) on a disconnected network [1]. This representation allows a natural decomposition of the initial large-scale problem into several smaller p-median sub-problems (sub-PMPs). The solutions of the sub-problems can be combined by solving a multiple choice knapsack problem (MCKP) to find an optimal solution of the original problem.

In this work we develop an improved version of the solution method presented in [2]. The improvements affect both the solution quality and the computation time. They include a reduction of the number of sub-PMPs to be solved, a more efficient versions of Lagrangian relaxation-based and heuristic techniques for solving the sub-PMPs, a very fast dynamic programming algorithm to optimally solve the MCKP, and a parallel implementation of particular components of the method. The proposed approach has been tested on a wide range of problem instances, known from the literature and randomly generated. The obtained results show that our approach is able to find optimal solutions to large-scale problem instances, considerably outperforming state-of-the-art approaches to large-scale PMPs on disconnected networks [3].

Research perspectives are devoted to apply the proposed method in the application fields where ODMP can play a relevant role, in all the systems where it is necessary to resolve effectively the trade-off between diversity and redundancy of the decision problem solutions.

References

1. Briant, O., Naddef, D.: The Optimal Diversity Management Problem. Operations Research, 52(4), 515–526, (2004).
2. Avella, P., Boccia, M., Martino, C.D., Oliviero, G., Sforza, A., Vasilyev, I.: A decomposition approach for a very large scale optimal diversity management problem. 4OR, 3(1), 23–37 (2005).
3. Agra, A., Cerdeira, J.O., Requejo, C.: A decomposition approach for the p-median problem on disconnected graphs. Computers & Operations Research, 86, 79-85 (2017).

On the Optimal Cost of the Transportation Problem with Interval Right-Hand-Sides

R. Cerulli, C. D'Ambrosio, M. Gentili

Abstract The Interval Trasportation Problem (ITP) is a variant of the well known transportation problem where the coefficients of the problem range in an interval (i.e., are interval numbers). In this paper we focus on the special case of ITP when only right-hand-sides are interval numbers. We focus in determining the best and worst values of the optimal cost of the ITP among all the feasible realizations of the right-hand-side parameters. While finding the best optimum is an easy task [1], to the best of our knowledge, a formal proof of the computational complexity of finding the worst optimum is still missing. In this paper we prove some general properties of the best and worst optimum values, and we propose a new heuristic approach that outperforms the existing approaches on a set of benchmark instances.

Key words: Transportation Problem, Uncertainty, Interval Optimization

References

1. S Chanas, M Delgado, JL Verdegay, and MA Vila. Interval and fuzzy extensions of classical transportation problems. *Transportation Planning and Technology*, 17(2):203–218, 1993.

R. Cerulli, C. D'Ambrosio
Department of Mathematics, University of Salerno, Giovanni Paolo II, 132, 84084, Fisciano (SA), Italy, e-mail: {raffaele, cdambrosio}@unisa.it

M. Gentili
Industrial Engineering Department, University of Louisville, 132 Eastern Parkway Louisville, KY 40292, e-mail: monica.gentili@louisville.edu,
Department of Mathematics, University of Salerno, Giovanni Paolo II, 132, 84084, Fisciano (SA), Italy, e-mail: mgentili@unisa.it

An exact algorithm for the split-demand one-commodity pickup-and-delivery travelling salesman problem*

Hipólito Hernández-Pérez and Juan José Salazar-González

DMEIO, Facultad de Ciencias, Universidad de La Laguna,
38200 La Laguna, Tenerife, Spain
hhperez@ull.edu.es , jjsalaza@ull.es

The *Split-Demand One-Commodity Pickup-and-Delivery Travelling Salesman Problem* (SD1PDTSP) is defined as follows. Let us consider a finite set of locations. Each location is related to a customer, with a known positive or negative demand of a commodity. For example, the commodity can be bicycles of identical type, the locations can represent bike stations in a city, and the demand can be the difference between the number of bicycles at the beginning of a day and at the end of the previous day in each station. We assume that the sum of all demands is equal to zero. Customers with negative demands correspond to pickup locations, and customers with positive demands correspond to delivery locations. The travel distances (or costs) between the locations are assumed to be known. There is *one* vehicle with a given capacity that must visit each location *at least* once through a route to move the commodity between the customers as they require. Each visit may partially satisfy the demand of a customer, and all the visits to that customer must end up with exactly its complete demand. The SD1PDTSP consists of finding a minimum-cost route for the vehicle such that it satisfies the demand of all customers without violating the vehicle capacity. Although a customer may be visited several times, a maximum number of allowed visits is assumed on each customer. The vehicle is not required to leave any location with an a-priori known load (neither empty nor full). Thus, if a location is considered the starting (ending) point of the route, the initial (final) load of the vehicle in the SD1PDTSP is a decision that must be determined within the optimization problem. Although our results can be adapted to the variant with a fixed initial load of the vehicle in a location, we do not consider it in this paper. Since several visits to a location are allowed, the vehicle could deliver some units of the commodity in a location and collect them later in another visit. Similarly, the vehicle can collect some units of the commodity in a location and deliver them later in another visit. The SD1PDTSP allows these solutions and therefore it can be seen as an inventory-routing problem where each customer has an a-priori stock of the commodity, requires to have an a-posteriori stock, has a capacity, and the demand is the difference between the a-priori and a-posteriori stocks. In other words, a customer in the SD1PDTSP may be used to *temporarily* deliver or collect units of commodity. This charac-

* This work has been partially supported by MTM2015-63680-R (MINECO/FEDER) and by Fundación CajaCanarias.

teristic is called *preemption* and it may provide routes with smaller costs respect to the non-preemption variant.

Let V_i be an ordered set of m nodes representing potential visits to location i. The set $V = \cup_{i \in I} V_i$ is the node set of a directed graph $G = (V, A)$ where A is the arc set connecting nodes associated with different locations. For a given subset S of nodes, we write $\delta^+(S) = \{(v, w) \in A : v \in S, w \notin S\}$ and $\delta^-(S) = \{(v, w) \in A : v \notin S, w \in S\}$. Given an arc $a = (v, w)$ we also denote the cost c_a from v to w as the travel cost c_{ij} from the location i associated with v to the location j associated with w. We consider the following mathematical variables. For each arc $a \in A$, a binary variable x_a assumes value 1 if and only if the route includes a, and a continuous variable f_a is the load of the vehicle when traversing a. For each node $v \in V$, a binary variable y_v assumes value 1 if and only if the route includes v, and a continuous variable g_v determines the number $|g_v|$ of units delivered (if $g_v > 0$) or collected (if $g_v < 0$) when performing the visit v. Then, the SD1PDTSP can be formulated as $\min \sum_{a \in A} c_a x_a$ subject to:

$$\sum_{a \in \delta^+(v)} x_a = \sum_{a \in \delta^-(v)} x_a = y_v \quad \text{for all } v \in V$$

$$\sum_{a \in \delta^+(S)} x_a \geq y_v + y_w - 1 \quad \text{for all } S \subseteq V \ , \ v \in S \ , \ w \in V \setminus S$$

$$\sum_{a \in \delta^+(S) \setminus \delta^+(1_1)} x_a \geq y_{i_{l+1}} \quad \text{for all } i \in I \ , \ l = 1, \ldots, m-1 \ (i_l \neq 1_1), \quad S \subseteq V : \ 1_1, i_l \in S \ , \ i_{l+1} \in V \setminus S$$

$$\sum_{a \in \delta^-(v)} f_a - \sum_{a \in \delta^+(v)} f_a = g_v \quad \text{for all } v \in V$$

$$0 \leq f_a \leq Q x_a \quad \text{for all } a \in A$$

$$y_{i_1} = 1 \ , \ \sum_{1 \leq l \leq m} g_{i_l} = d_i \quad \text{for all } i \in I$$

$$0 \leq p_i + \sum_{1 \leq k \leq l} g_{i_k} \leq q_i \quad \text{for all } i \in I \ , \ l = 1, \ldots, m-1$$

$$-q_i y_{i_l} \leq g_{i_l} \leq q_i y_{i_l} \quad \text{for all } i \in I \ , \ l = 2, \ldots, m$$

$$y_v, x_a \in \{0, 1\} \quad \text{for all } v \in V \ , \ a \in A.$$

Based on this model, our talk at ISCO2018 analyzes a new branch-and-cut approach to solve the SD1PDTSP that outperforms the ones proposed in [1,2].

References

1. Erdogan, G., Battarra, M., Wolfler Calvo, R., 2015. An exact algorithm for the static rebalancing problem arising in bicycle sharing systems. European Journal of Operational Research 245, 667–679.
2. Salazar-González, J. J., Santos-Hernández, B., 2015. The split-demand one-commodity pickup-and-delivery travelling salesman problem. Transportation Research Part B: Methodological 75, 58–73.

A MILP Formulation for a Multi-Tiered Vehicle Routing Problem with Global Cross-Docking

Anthony Smith (Department of Industrial Engineering, Stellenbosch University, South Africa)

Paolo Toth * (DEI, University of Bologna, Italy)

Jan H. van Vuuren (Department of Industrial Engineering, Stellenbosch University, South Africa)

In most *Vehicle Routing Problems* (VRPs), a characterization of the visited customers or facilities in terms of different commodity demand needs is not considered. In this talk we tackle a variant of the VRP with time-windows that arises in a real-life application related to the collection and delivery of pathological specimens in the transportation network of a pathology healthcare service provider. There are different types of specimens that have to be collected from a set of hospitals and clinics, and processed in potentially different ways at a set of laboratories within a transportation network. The variation of the pathological specimen types may be due to the nature of the specimens themselves, such as their purpose and processing requirements, as well as to the need of maintaining standards associated with a specimen, or may even be due to the intended destinations of the specimens. We segregate the available specimen processing facilities according to their respective processing and storage capabilities into a set of *tiers*. This tier allocation is nested in the sense that a facility of tier i can process any type of specimen that can be processed at a facility of tier j if $j < i$, but there exist certain commodity types which can be processed at a facility of tier i that cannot be processed at any facility of a lower tier. Facilities of the lowest tier represent customers (i.e. hospitals and clinics) at which the specimens originate and have to be collected --- these facilities have no specimen processing or storage capabilities --- their only role is that of introducing new specimens into the system. Facilities of higher tiers (i.e. laboratories) may or may not introduce new specimens into the system, but their distinguishing feature is that they all offer specimen processing capabilities or intermediate specimen storage capabilities. All facilities, excluding the facilities of the lowest tier, are assumed to offer the same storage capabilities.

We allow for handover of specimens at facilities in the sense that a specimen requiring processing at a facility of a specific tier may be transported by one vehicle to a facility of a lower tier than the required one, and then be collected later by some other vehicle which transports it to a facility of the required tier. We refer to this type of specimen handover, which may occur at a facility of any tier (save the lowest and the highest ones), as *global cross-docking*.

Another novel feature of the considered VRP variant is that we allow demand for specimen collection to spill-over into a subsequent planning period. We essentially assume that the time continuum may be partitioned into planning periods of fixed length. One planning period is considered at a time, and if the demand for a specimen collection occurs at a facility after the last vehicle has departed from that facility, then this specimen is simply collected from the facility during the following planning period.

It is also assumed that the possible deterioration of the quality of a specimen over time (*specimen expiration time*) is given by the time required for the specimen to be collected from a facility of the lowest tier and transported to the first facility that has the appropriate processing, storage or consolidation capabilities (i.e. the specimen deterioration occurs only as a result of being in transit prior to the first facility of a tier not of the lowest tier). It is therefore assumed that once a specimen has been delivered to a facility (of tier greater than the lowest one), the specimen is either processed there or stored in such a manner that its expiration window remains unaffected during storage (i.e. in a vacuum or at a low temperature) or future transportation (i.e. repackaged in such a manner so as to retain the specimen's integrity).

The specimen collection and processing system, with global cross-docking and demand spill-over to subsequent planning periods, described above is formulated, through a *Mixed Integer Linear Programming* (MILP) model, as a tri-objective VRP minimizing: i) the total time required to transport the specimens, ii) the difference between the longest and the shortest travel times associated with the vehicles (i.e., the balancing of the driver workload), and iii) the number of vehicles required to implement the specimen collection routing schedule. The MILP model builds on a combination of several models proposed in the literature for well-known variants of the VRP, but exhibits various novel features to take into account the characteristics previously outlined. The proposed MILP model is validated by implementing it through the MILP Solver CPLEX, and by applying it to a small hypothetical problem instance.

Although conceived within the context of pathological specimen collection, the considered problem also has several alternative applications, such as that of a national postal service and other organizations that incorporate consolidation centers within their distribution network.

WEC1 : Polyhedral Approaches II

- The Benders Algoritm for the b-CMS Dual Problem
  Eleazar Madriz, Yuri Tavares.

- Primal-dual approach to the multi-activity tour scheduling problem
  Stefania Pan, Mahuna Akplogan, Lucas Létocart, Nora Touati, Roberto Wolfler Calvo, Louis-Martin Rousseau.

- Beyond Vehicle Routing: a general purpose branch-cut-and-price code for applications where pricing is a resource constrained shortest path
  Artur Pessoa, Ruslan Sadykov, Eduardo Uchoa, Francois Vanderbeck.

The Benders Algoritm for the b-CMS Dual Problem.

E. Madriz, and Y. Tavares

CETEC-UFRB,
Cruz da Almas, BA, Brazil
eleazar@ufrb.edu.br, yuri.passo@ufrb.edu.br

Abstract. In this work, we present the Benders algorithm for the dual of the a b-Complementary Multisemigroup problem.

1 Introduction

Let $(A, \widehat{+})$ be a multisemigroup b-complementary and the b-Complementary Multisemigroup problem is

$$P : min\{ct : b \in \widehat{\sum_{g\in A}} t(g)g, t \in \mathbb{Z}_+^A\}$$

where $c \in \mathbb{R}^A$ [2].

Let (L, E) is a base for $C(A)$. In this paper we considered the following problems:$P_p : max\{ct : \ \Gamma t = \rho_r, \ \Pi t \geq \pi_r, \ t \in \mathbb{R}_+^r\}$; $P_d : max\{v\rho_r + w\pi_r : v\Gamma + w\Pi \leq c, \ v \in \mathbb{R}^l, w \in \mathbb{R}_+^e\}$. where $t = (t_1, , , , t_r)$; $\rho_r = (\rho_i(g_r) : i \in \{1, .., l\})$; $\pi_r = (\pi_j(g_r) : j \in \{1, .., e\})$; $\Gamma = [\rho_{ij}]_{l\times r}$, where $\rho_{ij} = \rho_i(g_j)$ for all $i \in \{1, .., l\}$ and $j \in \{1, .., r\}$; $\Pi = [\pi_{kj}]_{e\times r}$, where $\pi_{kj} = \pi_k(g_j)$ for all $k \in \{1, .., e\}$ and $j \in \{1, .., r\}$, for $r = |A_+|$, $l = |L|$ and $e = |E|$

In [2] we proved that he dual problem of P_p is the problem P_d. In the section 2, we present an algorithm for the dual problem of the a b-Complementary Multisemigroup used Bendres Decompostion [1].

2 The Benders Algorihtm of the P_d problem

For $C = \{(t_0, t) \in \mathbb{R} \times \mathbb{R}^r \mid \Pi t \geq \pi_r t_0, t \geq 0, t_0 \geq 0\}$. And $H(C)$ will denote the intersection of all subsets:

$$H(C) = \bigcap_{(t_0,t)\in C} \{(x_0, v) \in \mathbb{R} \times \mathbb{R}^l \mid x_0 t_0 + t(\Gamma^T v) - t_0(\rho_r v) \leq tc\}.$$

References

1. J. F. Benders. Partitioning procedures for solving mixed-variables programming problems. *Numerische Mathematik*, 4:238–252, 1962.
2. M. E. Duality for a b-complementary multisemigroup master problem. *Discrete Otimization*, 22:363–371, 2016.

```
n = 0;
Select a finite set Q^n ⊂ C ;
if H(Q^n) = ∅ then
    Algorithm terminates;
else
    if there are t_0 > 0 and t ∈ ℝ^r_+ such that (t_0, t) ∈ Q^n then
        solveMaxProblem = True;
    else
        put x_0 = +∞, take (x_0, v^n) ∈ H(Q^n) ;
        solveMaxProblem = False;
    end
end
terminated = False;
repeat
    if solveMaxProblem then
        Solve the problem max{x_0 | (x_0, v) ∈ H(Q^n)}                    (1)
        if the problem (1) is not feasible then
            Algorithm terminates;
        end
        Take (x_0^n, v^n) the optimal solution of the problem (1);
    end
    Solve the problem min{(c − Γ^T v^n)t | Πt ≥ π_r, t ≥ 0}                (2)
    solveMaxProblem = True;
    if the problem (2) is not feasible then
        terminated = True;
    else
        if the problem (2) has a finite optimal solution t^n then
            if (c − Γ^T v^n)t^n = x_0^n − ρ_r v^n then
                Solve the dual problem of the problem (2);
                Take (v^n, w^n) the optimal solution of problem P_d;
                terminated = True;
            else
                if (c − Γ^T v^n)t^n < x_0^n − ρ_r v^n then
                    Q^{n+1} = Q^n ∪ {(1, t^n)};
                    n = n + 1;
                end
            end
        else
            Select a vertex t^n ∈ P and extreme direction d^n ∈ C_0 such that
              t^n + λd^n → ∞, where λ → +∞;
            if (c − Γ^T v^n)t^n ≥ x_0^n − ρ_r v^n then
                Q^{n+1} = Q^n ∪ {(0, d^n)};
                n = n + 1;
            else
                Q^{n+1} = Q^n ∪ {(1, t^n), (0, d^n)};
                n = n + 1;
            end
        end
    end
until terminated;
```

Primal-dual approach to the multi-activity tour scheduling problem

Stefania Pan[1,3], Mahuna Akplogan[3], Lucas Létocart[1], Louis-Martin Rousseau[2], Nora Touati[3], and Roberto Wolfler Calvo[1]

[1] LIPN, UMR 7030 CNRS, Université Paris 13,
99 avenue Jean-Baptiste Clément 93430, Villetaneuse, France
{pan,letocart,wolfler}@lipn.fr,
[2] CIRRELT - Polytechnique Montréal, CP6079 Succ Centre Ville,
Montréal, Canada, H3C3A8,
louis-martin.rousseau@cirrelt.net,
[3] Horizontal Software, 9 Rue de l'Isly, 75008 Paris, France
{makplogan,ntouati}@horizontalsoftware.com

Abstract. In this work, we investigate the tour scheduling problem with a multi-activity context, a challenging problem that often arises in personnel scheduling. We propose a primal-dual approach, which makes use of column generation to get a lower bound, and large neighbourhood search to get upper bounds and good primal integer solutions. The two methods are embedded in the overall approach by working on and exchanging sets of columns.

Keywords: multi-activity tour scheduling, column generation, large neighborhood search

1 Introduction

Personnel scheduling problems consists of constructing feasible shift schedules to be assigned to staff, in order to satisfy workload requirements. These problems arise in several organizations such as airline and railways companies, hospitals, restaurants, retail stores and call centres. Due to economic considerations, personnel scheduling represents an intense and challenging research field [1]. Three main categories of problems can be distinguished in personnel scheduling: shift scheduling, days-off scheduling and tour scheduling. In this work, we deal with a problem in the latter category. We aim at specifying the time periods of the day and the days of the week in which employees have to work. Moreover, more than one work activity has to be scheduled, making the problem fall into multi-activity tour scheduling category. In these problems, we need not only to define the working days and the working periods, but also to specify the allocation of work activities.

2 Primal-dual approach

We solve the multi-activity tour scheduling problem by combining column generation and large neighborhood search into a primal-dual approach. The first solves the linear relaxation providing lower bound and fractional solutions, while the second starts from integer rounded solutions and aims at improving them, providing upper bounds on the integer problem.

Column generation (CG). CG is a classical technique to solve linear programs with a large number of variables. Recently, CG approaches have been widely used to solve multi-activity tour scheduling problems [1]. This method considers on one hand a master problem that takes into account workload requirements, minimizing the total cost given by under and over coverage. On the other hand, new feasible schedules are built solving the subproblems, where the legal constraints, such as consecutive working hours, daily working hours, breaks and skills, are considered. Each subproblem is decomposed and solved in different phases. First (*phase 1*), work activities are combined to build feasible timeslots, which are in turn used to build feasible daily shifts (*phase 2*). Finally (*phase 3*), daily shifts are combined to build feasible weekly schedules. Rules defining timeslots and daily shifts are taken into account both by means of automata, and by solving resource constrained shortest path problems on the extended graphs.

Large Neighborhood Search (LNS). The LNS algorithm was first introduced by [3], and it iteratively destroys part of the current solution and repairs it in the hope of finding a better solution. Similarly to [2], destroying here means choosing an employee and removing his schedule, while repairing means assigning a new schedule to the selected employee in order to improve the global solution.

The overall primal-dual approach. The proposed primal-dual approach iteratively calls CG and LNS at each iteration. It starts by solving the linear relaxation using CG. When optimality is achieved, a lower bound and a fractional solution are available. The latter is rounded to provide a feasible integer solution which is improved by means of LNS. As soon as a local optimum is found, the best solution is given as initial solution to CG, which again solves the linear relaxation and stops when the value of the master problem is close to optimality. Preliminary computational results on instances built with input from quick service restaurants, show that the primal-dual method proposed allows an average gap of 5% between the lower bound given by the linear relaxation and the upper bound of the best integer solution obtained.

References

1. Jorne Van den Bergh, Jeroen Beliën, Philippe De Bruecker, Erik Demeulemeester, and Liesje De Boeck. Personnel scheduling: A literature review. *European Journal of Operational Research*, 226(3):367–385, 2013.
2. Claude-Guy Quimper, and Louis-Martin Rousseau. A large neighbourhood search approach to the multi-activity shift scheduling problem. *Journal of Heuristics*, 16 (3):373–392, 2010.
3. Paul Shaw. Using constraint programming and local search methods to solve vehicle routing problems. In *Proceedings of the Fourth International Conference on Principles and Practice of Constraint Programming CP 1998*, 417–431, 1998.

[3] This work has been funded by BPIfrance in the frame of the French cooperative project ADAMme, Projet Investissement Avenir (FSN: AAP CISN2).

Beyond Vehicle Routing: a general purpose branch-cut-and-price code for applications where pricing is a resource constrained shortest path

Artur Pessoa[1], Ruslan Sadykov[2,3], Eduardo Uchoa[1], and François Vanderbeck[3,2]

[1] Universidade Federal Fluminense, LOGIS, Niteroi, Brazil
[2] INRIA Bordeaux, team RealOpt, Talence, France
[3] Univ. of Bordeaux, Institute of Mathematics, Talence, France UMR 5251

Column generation algorithms where the pricing is solved as a resource constrained shortest path problem have been used in a variety of applications, as surveyed in [5]. Pioneering work on a generic solver using column generation based on a resource constrained shortest path subproblem was the GenCol software [13]. Our aim is to develop such a platform that includes both generic modeling tools and an highly efficient branch-cut-and-price. Our solver relies on generalizing the most advanced techniques that were recently developed for classical variants of the vehicle routing problem. It considers several resource constraints simultaneously, even allowing for continuous resources (as opposed to the discrete assumptions made by traditional dynamic programming approaches), sometimes even allowing zero or negative resource consumptions. The pricing is done by a bi-directional labeling algorithm, implemented over the so-called bucket graph (as proposed in [11]). Besides the good performance of the pricing oracle, the overall efficiency of the branch-cut-and-price relies on advanced features such as a procedure for fixing arc variables by reduced costs [4,8]; an algorithm for gradually enforcing total or partial elementarity of subproblem solution paths [10]; an self-adjusting dual price smoothing stabilization for improving the convergence of the column generation [7]; a heuristic local search separation procedure for limited-memory rank-1 Chvatal–Gomory cuts [6]; a labeling dynamic programming algorithm for enumerating elementary subproblem solution paths [1]; a multi-phase pseudo-costs based strong branching procedure [6]; and the generic diving heuristic for improving the initial primal bound of [12].

In this presentation we will focus on the scope of applications that are amenable to our branch-cut-and-price solver. The goal is to convey the ease of access to an efficient solver for the many combinatorial optimization problems that can be decomposed into resource constrained shortest path subproblems, once linking constraints have been dualized in a Lagrangian way. Beyond the case of vehicle routing problems (VRP) for which the solver was originally developed, we will focus on problems where the VRP-like structure is not evident, including machine scheduling, packing, resource allocation, and network design problems [8,2,9,3]. After showing how such problems reduce to our approach, we evaluate how it performs in practice when compared to the best existing approaches.

References

1. R. Baldacci, N. Christofides, and A. Mingozzi. An exact algorithm for the vehicle routing problem based on the set partitioning formulation with additional cuts. *Mathematical Programming*, 115(2):351–385, 2008.
2. G. Belov and G. Scheithauer. A branch-and-cut-and-price algorithm for one-dimensional stock cutting and two-dimensional two-stage cutting. *European Journal of Operational Research*, 171(1):85 – 106, 2006.
3. B. Gendron and M. Larose. Branch-and-price-and-cut for large-scale multicommodity capacitated fixed-charge network design. *EURO Journal on Computational Optimization*, 2(1):55–75, Jun 2014.
4. T. Ibaraki and Y. Nakamura. A dynamic programming method for single machine scheduling. *European Journal of Operational Research*, 76(1):72 – 82, 1994.
5. S. Irnich and G. Desaulniers. Shortest path problems with resource constraints. *Column generation*, pages 33–65, 2005.
6. D. Pecin, A. Pessoa, M. Poggi, and E. Uchoa. Improved branch-cut-and-price for capacitated vehicle routing. *Mathematical Programming Computation*, 9(1):61–100, 2017.
7. A. Pessoa, R. Sadykov, E. Uchoa, and F. Vanderbeck. Automation and combination of linear-programming based stabilization techniques in column generation. *INFORMS Journal on Computing*, (Forthcoming), 2017.

8. A. Pessoa, E. Uchoa, M. P. de Aragão, and R. Rodrigues. Exact algorithm over an arc-time-indexed formulation for parallel machine scheduling problems. *Mathematical Programming Computation*, 2:259–290, 2010.
9. M. Posta, J. A. Ferland, and P. Michelon. An exact method with variable fixing for solving the generalized assignment problem. *Computational Optimization and Applications*, 52:629–644, 2012.
10. R. Roberti and A. Mingozzi. Dynamic ng-path relaxation for the delivery man problem. *Transportation Science*, 48(3):413–424, 2014.
11. R. Sadykov, E. Uchoa, and A. Pessoa. A bucket graph based labeling algorithm with application to vehicle routing. Technical Report L-2017-7, Cadernos do LOGIS-UFF, Niterói, Brazil, October 2017.
12. R. Sadykov, F. Vanderbeck, A. Pessoa, I. Tahiri, and E. Uchoa. Primal heuristics for branch-and-price: the assets of diving methods. *INFORMS Journal on Computing*, (Forthcoming), 2018.
13. B. Sansó, M. Desrochers, J. Desrosiers, Y. Dumas, and F. Soumis. Modeling and solving routing and scheduling problems: GENCOL user's guide. Research report, GERAD, Montreal, Canada, 1990.

WEC2 : Approximation Algorithms II

- Approximation algorithm for scheduling a chain of tasks on hybrid platform with energy constraint1
  Massinissa Ait Aba, Lilia Zaourar, Alix Munier.

- Polytope Membership in High Dimension
  Evangelos Anagnostopoulos, Ioannis Emiris, Vissarion Fisikopoulos.

- 2CSPs all are approximable within some constant differential factor
  Sophie Toulouse, Jean-François Culus.

- A tight linear time 13/12-approximation algorithm for the P2 — — Cmax problem
  Federico Della Croce, Rosario Scatamacchia, Vincent T'Kindt.

Approximation algorithm for scheduling a chain of tasks on hybrid platform with energy constraint

Massinissa Ait aba[1]✉, Lilia Zaourar[1] and Alix Munier[2]

[1] CEA, LIST, Computing and Design Environment Laboratory
91191 GIF SUR YVETTE CEDEX, FRANCE.
[2] LIP6-UPMC, 4 place Jussieu, 75005 Paris, France

Keywords: Hybrid platform, Linear chain of tasks, Makespan, Energy, Approximation algorithm

1 Introduction

Heterogeneous computing platforms offer significant computational power while preserving the energy consumption for running High Performance Parallel (HPC) applications [1]. Thus, they represent nowadays interesting means of calculations. In particular, hybrid platforms using GPU accelerators in addition to classical computing units such as CPU are widely used architectures in HPC.

In order to reap the benefits of heterogeneous platforms, efficient and automatic strategies to manage computing resources is increasingly important for running applications. These new hybrid architectures have given rise to new scheduling problems for allocating and sequencing calculations on different resources by optimizing one or more criteria.

This paper addresses task scheduling of a chain of task application onto a hybrid platform composed of two types of resources (CPU et GPU) which is a special case of [2]. Compared to the work presented in [3,4] we introduce energy constraint and communication delays between the resources as well as between the tasks that are linked by precedence constraints. We present a detailed mathematical formulation as well as some preliminary approximation results.

2 Problem definition

This study considers a heterogeneous platform composed of one GPU and one CPU. An application A of n tasks is represented by an oriented chain of tasks Graph $G_{Ch}(V, E)$, each vertex $v \in V$ represents a task t_i. Each arc $e = \{t_i, t_{i+1}\} \in E$ represents a precedence constraint between two successive tasks t_i and t_{i+1}, $i = \overline{1..n-1}$. The communication cost between t_i and t_{i+1} is denoted by $ct_{i,i+1}$.

A task can be executed by CPU or GPU. Executing the task t_i on a CPU (resp. GPU) generates an execution time equal to w_{i0} (resp. w_{i1}) and an energy equal to e_{i0} (resp. e_{i1}). We denote by E the allowed quantity of energy consumed during the execution. E represents in our case an energy bound that should not be exceeded during the execution. A communication delay is also considered if two tasks are executed on two different resource types, [5] provides the exact formula to evaluate $ct_{i,i+1}$ which takes into consideration latencies and available bandwidth between processors. A task t_i can be executed only after the complete execution of its predecessor t_{i+1}. We do not allow duplication of tasks and preemption. We denote by C_i the completion time of the task t_i and C_{max} the completion time of the application A (makespan). The aim is to find the minimum makespan scheduling respecting the energy bound E.

3 Preliminary results

We have proposed a two-phase algorithm to solve the problem of scheduling chain of tasks with energy constraint. In the first phase, we start by solving an assignment problem to find which resource (CPU or GPU) will execute each task. We propose mathematic model ($P1$) for solving the assignment problem while the energy and precedence constraints are satisfied. The solution obtained by this model represents a lower bound for the final makespan. Then we solve the relaxation ($P1^{'}$) of the model ($P1$). In order to obtain a feasible assignment for the tasks, we rounded up the fractional solution of program ($P1^{'}$). In the second phase, we use the assignment of tasks to get a feasible schedule. This algorithm guarantees a ratio of 2 compared to the optimal solution. Tests on large instances close to reality demonstrated the efficiency of our method and shows the limits of solving the problem with a solver such as *CPLEX* [6].

As part of the future, we will use the energy constraint by varying the value E to find a Pareto set which minimizes both the total execution time (makespan) and the energy consumption. An extension to more general classes of graphs is also planned to handle real applications.

References

1. Bosilca George, Bouteiller Aurelien, Danalis Anthony, Faverge Mathieu, Herault Thomas, Dongarra Jack J. *PaRSEC: Exploiting Heterogeneity to Enhance Scalability.* Computing in Science & Engineering, 2013.
2. Massinissa Ait Aba, Lilia Zaourar, Alix Munier Kordon. *Approximation Algorithm for scheduling dependent tasks in heterogeneous systems.* Fifteenth International Workshop HeteroPar 2017, Santiago de Compostela, Spain.
3. Amaris Marcos, Lucarelli Giorgio, Mommesin Clment, Denis Trystram. *Generic algorithms for scheduling applications on hybrid multi-core machines.* European Conference on Parallel Processing. Springer, Cham, 2017.
4. Kedad Sidhoum Safia, Monna Florence, Trystram Denis. *Scheduling tasks with precedence constraints on hybrid multi-core machines.* IEEE International Parallel and Distributed Processing Symposium Workshop (IPDPSW), 2015.
5. Zaourar Lilia, Ait Aba Massinissa, Briand David, Jean-Marc Philippe. *Modeling of applications and hardware to explore task mapping and scheduling strategies on a heterogeneous micro-server system.* IEEE IPDPSW, 2017.
6. IBM ILOG CPLEX V12.5 User's Manual for CPLEX, http://www.ibm.com

Polytope Membership in High Dimension

Evangelos Anagnostopoulos[1], Ioannis Z. Emiris[1] and Vissarion Fisikopoulos[2]

[1] National and Kapodistrian University of Athens, Athens, Greece,
{aneva,emiris}@di.uoa.gr
[2] Oracle Corp., Greece, vissarion.fysikopoulos@oracle.com

Abstract. We study the fundamental problem of polytope membership aiming at large convex polytopes, i.e. in high dimension and with many facets, given as an intersection of halfspaces. Standard data-structures as well as brute force methods cannot scale, due to the curse of dimensionality. We design an efficient algorithm, by reduction to the approximate Nearest Neighbor (ANN) problem based on the construction of a Voronoi diagram with the polytope being one bounded cell. We thus trade exactness for efficiency so as to obtain complexity bounds polynomial in the dimension, by exploiting recent progress in the complexity of ANN search. We employ this algorithm to present a novel practical boundary oracle based on a Newton-like iterative intersection procedure. We implement our algorithms and compare with brute-force approaches to show that they scale very well as the dimension and number of facets grow larger.

Keywords. Geometric optimization, convex polytope, membership oracle, approximation algorithms, general dimension, nearest-neighbor search, implementation.

Introduction. Let us consider a convex polytope P in H-representation, i.e. as the intersection of a finite set of linear inequalities: $P = \{x \in \mathbb{R}^d \mid Ax \leq b,\ A \in \mathbb{R}^{n\times d},\ b \in \mathbb{R}^n\}$. The main focus of our work is to create efficient data structures for the problems of polytope membership and boundary when the dimension d is high. The membership problem is to preprocess P such that, given a query point q, we can efficiently decide whether q lies inside or outside P. The polytope boundary problem is to preprocess P such that, given a query ray emanating from inside the polytope, we can efficiently compute the point where it intersects the boundary of the polytope. We will allow ourselves (i.e. our data structures) to answer correctly within some approximation error ϵ and with some success probability, in order to gain some margin for efficient high-dimensional solutions.

The motivation stems from an algorithmic point of view, where improving the complexity of oracles, in particular membership, implies improvements to algorithms used to solve combinatorial optimization problems such as the ellipsoid, interior point or randomized methods. Another important example of application is volume approximation which has also an established connection to combinatorial optimization. For example, the volume of order polytopes gives the number of linear extentions of the associated partial order set.

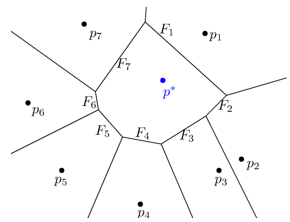


Fig. 1. A conceptual presentation of the alternative view of the polytope as a Voronoi cell instead of an intersection of hyperplanes.

Recent advances in approximate nearest neighbor (ANN) search guide us to using ANN for answering these crucial geometric predicates in convexity theory.

Our contribution. We describe a simple constructive reduction from the polytope membership problem to point location (Figure 1), then show under which conditions this reduction holds for the respective approximate versions of the problems. This gives us the flexibility to exploit advances in the research of ANN, but for the approximate polytope membership problem. Thus, we use a high dimensional solution to ANN[4] in order to offer a practical approximate polytope membership oracle in high dimension with complexity bounds polynomial in the dimension d and sublinear in the number of inequalities n. We remark that any (high-)dimensional ANN solution can be utilized and we can inherit its complexity and its properties.

We further present an application of the designed membership data structures in order to obtain novel experimental (approximate) solutions for the polytope boundary problem. Our approach exploits the nature of the aforementioned reduction in order to iteratively close in on the boundary point, in a Newton's method-like manner.

We implement and experimentally examine our algorithms and illustrate that they scale well as the dimension and the number of facets grow larger, compared to the trivial approach of checking every hyperplane, which is a plausible solution in the exact setting, especially in the high-dimensional case. Specifically we demonstrate results up to $d = 1000$ and $n = 10^6$ and show that that we have a $10\times$ speed-up.

References

1. Dudley, R.: Metric entropy of some classes of sets with differentiable boundaries. J. Approximation Theory (1974)
2. Bentley, J., Preparata, F., Faust, M.: Approximation algorithms for convex hulls. Commun. ACM (1982)
3. Arya, S., da Fonseca, G.D., Mount, D.: Optimal approximate polytope membership. In: Proc. ACM-SIAM Symposium on Discrete Algorithms. (2017)
4. Andoni, A., Razenshteyn, I.: Optimal data-dependent hashing for approximate near neighbors. In: Proc. ACM STOC. (2015)

2 CSPs all are approximable within some constant differential factor

Jean-François Culus[1] and Sophie Toulouse[2]

[1] CEREGMIA, Université des Antilles et de la Guyane, France
[2] LIPN (UMR CNRS 7030), Institut Galilée, Université Paris 13, France

1. Addressed issue. Given an integer $q \geq 2$, a *Constraint Satisfaction Problem* (CSP in short) over $\mathbb{Z}_q$ considers a set $\{x_1, \ldots, x_n\}$ of $\mathbb{Z}_q$-valued variables and a set $\{C_1, \ldots, C_m\}$ of constraints, where a constraint C_i consists of the application of a (non constant) predicate $P_i : \mathbb{Z}_q^{k_i} \to \{0, 1\}$ to a tuple $x_{J_i} = (x_{i_1}, \ldots, x_{i_{k_i}})$ of variables. In the most general case, a weight $w_i \in \mathbb{Q}$ is associated with each constraint C_i and functions P_i may be $\mathbb{Q}$-valued. The goal then is to assign values to the variables so as to optimize objective function $\sum_{i=1}^{m} w_i P_i(x_{J_i})$ over $\mathbb{Z}_q^n$. Given a universal constant integer k, we denote by $\mathsf{k\,CSP-q}$ the restriction of the corresponding unconstrained optimization problem to the case where each constraint C_i acts on at most k of the variables.

This work concerns the *differential approximability* [3] of $\mathsf{k\,CSP-q}$. Given an instance I of $\mathsf{k\,CSP-q}$, we denote by $v(I, .)$ its objective function, by $\mathrm{opt}(I)$ and $\mathrm{wor}(I)$ respectively the optimum and the worst solution values on I. A solution x of I is ρ-differential approximate for some $\rho \in]0, 1]$ if it performs a differential ratio $(v(I, x) - \mathrm{wor}(I))/(\mathrm{opt}(I) - \mathrm{wor}(I))$ at least ρ. $\mathsf{k\,CSP-q}$ is approximable within differential factor ρ if it is possible to compute within polynomial time on each of its instances a solution with differential ratio at least ρ.

A natural question as regards differential approximability of CSPs is: what are the greatest integers $q, k \geq 2$ for which $\mathsf{k\,CSP-q}$ is approximable within a constant factor? $\mathsf{2\,CSP-2}$ and $\mathsf{3\,CSP-2}$ are approximable within factor respectively $2 - \pi/2 > 0.429$ [4] and $1 - \pi/4$ [2]. When $q \geq 3$ or $k \geq 4$, the question whether $\mathsf{k\,CSP-q}$ is approximable within any constant differential factor is open.

A common way to exhibit approximability bounds for a given optimization problem consists in reducing to or from another optimization problem for which approximability bounds are known. We here analyse a specific reduction from $\mathsf{k\,CSP-q}$ to $\mathsf{k\,CSP-p}$ given three constant integers q, p, k with $q > p \geq k \geq 2$.

2. Our approach. Let $\mathcal{P}_p(\mathbb{Z}_q)$ refer to the set of the p-cardinality subsets of $\mathbb{Z}_q$. As observed in [1], given an instance I of $\mathsf{k\,CSP-q}$, restriction $I(S)$ of I to any $S \in \mathcal{P}_p(\mathbb{Z}_q)^n$ can be assimilated to an instance of $\mathsf{k\,CSP-p}$. A natural way to derive from a hypothetical approximation algorithm for $\mathsf{k\,CSP-p}$ approximate solutions on I therefore consists in solving subinstances $I(S), S \in \mathcal{P}_p(\mathbb{Z}_q)^n$. This is precisely what we do, but restricting to a constant number of subsets S. Namely, we restrict to the $\binom{q}{p} = O(q^p)$ subsets of the form T^n where $T \in \mathcal{P}_p(\mathbb{Z}_q)$.

Our goal is to express a sum of solution values on subinstances $I(T^n)$ as a sum of solution values on I that include $\mathrm{opt}(I)$. To do so, we model multisubsets of $\{T^n \,|\, T \in \mathcal{P}_p(\mathbb{Z}_q)\}$ and $\mathbb{Z}_q^n$ by two arrays Ψ and Φ with q columns on $\mathbb{Z}_q$. Let x^* be an optimal solution of I. We identify a row $(u_0, \ldots, u_{q-1})$ of these arrays with the vector x of $\mathbb{Z}_q^n$ whose coordinates x_j satisfy the following rule: if $x_j^* = c$, then $x_j = u_c$. We then introduce the following family of combinatorial designs:

Definition 1. *Let q, p, k be three integers with $q \geq p \geq k \geq 2$. Then given any two integers $R \geq 1, R^* \in \{1, \ldots, R\}$, $\Gamma(R, R^*, q, p, k)$ refers to the (possibly empty) set of pairs (Ψ, Φ) of arrays with q columns over $\mathbb{Z}_q$ that satisfy that:*

1. *the components of each row of Ψ take at most p distinct values;*
2. *$(0, 1, \ldots, q-1)$ occurs R^* times as a row in Φ;*
3. *for all sequences $J = (c_1, \ldots, c_k)$ of column indices, subarrays $(\Psi^{c_1}, \ldots, \Psi^{c_k})$ and $(\Phi^{c_1}, \ldots, \Phi^{c_k})$ coincide up to the ordering of their rows.*

Furthermore, $\gamma(q, p, k)$ refers to the greatest real γ for which there exist two integers R, R^* such that $R^*/R = \gamma$ and $\Gamma(R, R^*, q, p, k) \neq \emptyset$.

Requirements *1.*, *2.* and *3.* respectively ensure that array Ψ models vectors of $\{T^n \,|\, T \in \mathcal{P}_p(\mathbb{Z}_q)\}$, x^* occurs at least R^* times in the solution family modelled by Φ, and the sum of solution values over the solution family modelled by Ψ equals the sum of solution values over the solution family modelled by Φ.

3. Obtained results. We summarize below the results we obtain:

Lemma 2. *For all integers q, p, k with $q > p \geq k \geq 2$, on any instance I of* k CSP−q*, solutions that perform the best objective value among those whose coordinates take at most p distinct values are $\gamma(q, p, k)$-differential approximate.*

Theorem 3. *For all constant integers q, p, k with $q > p \geq k \geq 2$, if* k CSP−p *is approximable within some differential factor ρ, then* k CSP−q *is approximable within differential factor $\gamma(q, p, k) \times \rho$.*

Theorem 4. *For all integers q, k with $q > k \geq 2$ and all $p \in \mathbb{N}$, we have:*

$$\gamma(q+p, k+p, k) \geq 2/\sum_{r=0}^{k} \binom{q}{r}\binom{q-1-r}{k-r} \geq 1/(q-k/2)^k$$

Corollary 5 (using [4]). *For all constant integer $q \geq 2$,* 2 CSP−q *is approximable within differential factor $(2-\pi/2)/(q-1)^2$.*

Lemma 2 (and then Theorem 3) follows from Definition 1. Theorem 4 relies on a recursive construction for families $\Gamma(R, 1, q, k, k)$ of combinatorial designs. Corollary 5 is then straightforward. Notice that Theorems 3 and 4 additionally reduce the question whether k CSP−q is approximable within some constant differential factor to the consideration of integers q, k with $k \geq q \geq 2$.

References

[1] Charikar, M., Makarychev, K. Makarychev, Y.: Near-optimal algorithms for maximum constraint satisfaction problems. ACM Trans. on Alg., 5, pp. 1–14 (2009)

[2] Culus, J.-F., Toulouse, S., Roupin, F.: Differential approximation of SNP optimization problems. in ISCO 2012 (short paper) (2012)

[3] Demange, M. and Paschos, V. T.: On an approximation measure founded on the links between optimization and polynomial approximation theory. TCS, 158, pp. 117–141 (1996)

[4] Nesterov, Y.: Semidefinite relaxation and nonconvex quadratic optimization. Optimization Methods and Software, 9, pp. 141–160 (1998)

A tight linear time $\frac{13}{12}$-approximation algorithm for the $P2||C_{\max}$ problem

Federico Della Croce[1*], Rosario Scatamacchia[1], and Vincent T'kindt[2]

[1] DIGEP, Politecnico di Torino, Torino, Italy
[2] Université de Tours, Laboratoire d'Informatique Fondamental et Appliqué (EA 6300), ERL CNRS 6305, Tours, France

Keywords: Two Identical Parallel Machines Scheduling, Makespan, LPT rule, Linear Programming, Approximation

We consider the $P2||C_{max}$ scheduling problem, where the goal is to schedule n jobs on 2 identical parallel machines M_1, M_2 to minimize the makespan. $P2||C_{max}$ is NP-hard in the ordinary sense. A pioneering approximation algorithm for the problem is the Longest Processing Time (LPT) rule proposed in [5] for the more general $P||C_{max}$ problem. It requires to sort the jobs in non-ascending order of their processing times p_j $(j = 1, \dots, n)$ and then to assign one job at a time to the machine whose load is smallest so far. This assignment of jobs to machines is also known as List Scheduling (LS). Several properties have been established for LPT in the last decades [1, 2, 4, 5].
We mention other popular approximation algorithms which exploit connections of $P||C_{max}$ with bin packing: *Multifit* [3], *Combine* [8] and *Listfit* [6]. In [9], a fully polynomial time approximation scheme (FPTAS) was devised for $P2||C_{max}$ (and for the more general $Pm||C_{max}$ if the number of machines is fixed) which solves the problem with accuracy $1 + \epsilon$ in time $O((n/\epsilon)^{1/\epsilon^2})$. Such algorithms provide better worst case performance than LPT but at the cost of higher running times. The current best algorithm running with low polynomial complexity and providing constant approximation ratio has ratio $\frac{12}{11}$ and is due to [7].
We consider the jobs sorted by non-ascending p_j $(p_j \geq p_{j+1},\ \ j = 1, \dots, n-1)$. We denote by *critical* the job (the machine) that provides the makespan of the given schedule. Let C^* be the optimal makespan. The following proposition and related corollary (proof omitted, will be presented at the Conference) hold.

Proposition 1 *Consider a schedule obtained by assigning jobs* $1, ..., 2k$ *to the machines according to some policy and then apply LS to the remaining jobs* $2k+1, ..., n$. *If the critical job* j *is such that* $j \geq 2k+1$, *then* $\rho \leq 1 + \frac{1}{2(k+1)}$.

Corollary 1 *Given a problem* P_1 *with* n *jobs, consider the subproblem* P_{red} *with the first* $2k$ *jobs only. If problem* P_{red} *is solved by an algorithm with approximation ratio* $1 + \frac{1}{2(k+1)}$, *then the same approximation ratio holds for* P_1 *by applying LS to the remaining subset of jobs.*

By exploiting, Corollary 1, we can both improve the time complexity of the fptas in [9] and the approximation ratio of [7] by means of the following propositions (proof omitted, will be presented at the Conference).

* federico.dellacroce@polito.it

Proposition 2 *Given a $P2||C_{\max}$ instance with n jobs, an approximation ratio $(1+\epsilon)$ can be reached with complexity $O(\frac{1}{\epsilon^3}+n)$ for all $\epsilon > 1/n$.*

Consider the following approximation algorithm.

Algorithm 1

1: **Input:** $P2||C_{max}$ instance I with n jobs and 2 machines, parameter k.
2: Select jobs $1...k$ inducing a k-job instance I': apply LPT to I' and get schedule S'.
3: Search for the best swap $SW^1_{i,j}$ (if any) between any job i on machine M_1 and any job j on machine M_2 that improves the makespan of S'.
4: Search for the best swap $SW^2_{i,j,k}$ (if any) between any job i on machine M_1 and any pair of jobs j,k on machine M_2 that improves the makespan of S'.
5: Search for the best swap $SW^3_{i,j,k}$ (if any) between any job i on machine M_2 and any pair of jobs j,k on machine M_1 that improves the makespan of S'.
6: Apply the best swap (among $SW^1_{i,j}$, $SW^2_{i,j,k}$, $SW^3_{i,j,k}$) to S' reaching schedule $S^{''}$.
7: Given $S^{''}$, apply LS to the remaining $n-k$ jobs and return the final schedule S^*.

In practice, Algorithm 1 applies first LPT to the reduced instance I' composed by the largest k jobs yelding subschedule S'. Then, a single step of local search between pairs or triples of jobs is applied to I' yelding subschedule $S^{''}$. Finally, starting from $S^{''}$, List Scheduling is applied to the remaining $(n-k)$ jobs.

Proposition 3 *Algorithm* 1 *applied with* $k=10$ *reaches a tight* $\frac{13}{12}$ *approximation ratio.*

Remark 1 *We employ linear programming to show that Algorithm* 1 *applied to the reduced instance* I' *cannot have approximation ratio superior to* $13/12$ *(a tight example in this case occurs with* 8 *jobs with processing times* $7,5,2,2,2,2,2,2$*).*

References

1. Blocher, J.D., Sevastyanov, D.: A note on the coffman–sethi bound for lpt scheduling. *Journal of Scheduling* 18, 325–327 (2015)
2. Chen, B.: A note on LPT scheduling. *Operation Research Letters* 14, 139–142 (1993)
3. Coffman(Jr.), E.G., Garey, M.R., Johnson, D.S.: An application of bin-packing to multiprocessor scheduling. *SIAM Journal on Computing* 7, 1–17 (1978)
4. Coffman(Jr.), E.G., Sethi, R.: A generalized bound on LPT sequencing. *RAIRO Supplement* 10, 17–25 (1976)
5. Graham, R.L.: Bounds on multiprocessors timing anomalies. *SIAM J. Applied Math.* 17, 416–429 (1969)
6. Gupta, J.N.D., Ruiz-Torres, A.J.: A listfit heuristic for minimizing makespan on identical parallel machines. *Production Planning & Control* 12(1), 28–36 (2001)
7. He, Y., Kellereri, H., Koto, V.: Linear compound algorithms for the partitioning problems. *Naval Research Logistics* 47, 593–601 (2000)
8. Lee, C.Y., Massey, J.D.: Multiprocessor scheduling: combining lpt and multifit. *Discrete Applied Mathematics* 20(3), 233–242 (1988)
9. Sahni, S.: Algorithms for scheduling independent tasks. *Journal of the ACM* 23, 116–127 (1976)

WEC3 : Metaheuristics

- A Heuristic for Maximising Energy Efficiency in an OFDMA System Subject to QoS Constraints
  Adam Letchford, Qiang Ni, Zhaoyu Zhong.

- Multi-start local search procedure for the maximum fire risk insured capital problem
  Maria Isabel Gomes, Lourdes Belchior Afonso, Nelson Chibeles-Martins, Joana Fradinho.

- Selective and dynamic distribution - Petrol Secondary Distribution
  Adnane Elhilali Alaoui, Youssef Benadada.

- Descent with mutations applied to the linear order problem
  Olivier Hudry.

A Heuristic for Maximising Energy Efficiency in an OFDMA System Subject to QoS Constraints

Adam N. Letchford[1], Qiang Ni[2], and Zhaoyu Zhong[1]

[1] Department of Management Science, Lancaster University,
Lancaster LA1 4YX, United Kingdom
{A.N.Letchford,z.zhong1}@lancaster.ac.uk
[2] School of Computing and Communications, Lancaster University,
Lancaster LA1 4WA, United Kingdom
q.ni@lancaster.ac.uk

Orthogonal Frequency-Division Multiple Access (OFDMA) is a popular multiplexing scheme in modern mobile wireless communications systems [1]. In OFDMA, we have a set I of *subcarriers* and a set J of *users*. Each subcarrier can be assigned to at most one user, but a user may be assigned to more than one subcarrier. For each $i \in I$, we are given a *bandwidth* B_i (in megahertz), and a *noise power* N_i (in watts). If we allocate p_i watts of power to subcarrier i, the data rate of that subcarrier (in megabits per second) is

$$f_i(p_i) = B_i \log_2 (1 + p_i/N_i) .$$

Several different optimisation problems have been defined in connection with OFDMA systems (e.g., [2–8]). In [3], we considered the following specific problem. In addition to the above data, we are given a total power limit P (in watts), a *system power* σ and, for each $j \in J$, a *demand* d_j (in megabits per second). The task is to simultaneously allocate power to subcarriers, and subcarriers to users, so that energy efficiency is maximised, the total power limit is not exceed, and the demand of each user is satisfied. The user demand constraints are one specific way of ensuring *quality of service* (QoS).

This problem was formulated as a *mixed 0-1 fractional program* in [3]. For all $i \in I$ and $j \in J$, let the binary variable x_{ij} indicate whether user j is assigned to subcarrier i, let the non-negative variable p_{ij} represent the amount of power supplied to subcarrier i to serve user j, and let r_{ij} denote the associated data rate. The formulation was:

$$\max \quad \frac{\sum_{i\in I}\sum_{j\in J} r_{ij}}{\sigma + \sum_{i\in I}\sum_{j\in J} p_{ij}} \tag{1}$$

$$\begin{aligned}
\textstyle\sum_{i\in I}\sum_{j\in J} p_{ij} &\le P - \sigma && \\
\textstyle\sum_{j\in J} x_{ij} &\le 1 && (\forall i \in I) \\
\textstyle\sum_{i\in I} r_{ij} &\ge d_j && (\forall j \in J) \\
r_{ij} &\le f_i(p_{ij}) && (\forall j \in J) \qquad (2) \\
p_{ij} &\le (P-\sigma) x_{ij} && (\forall i \in I,\, j \in J) \\
x_{ij} &\in \{0,1\} && (\forall i \in I,\, j \in J) \\
p_{ij}, r_{ij} &\in \mathbb{R}_+ && (\forall i \in I,\, j \in J).
\end{aligned}$$

Unfortunately, the fractional objective function (1) and the nonlinear constraints (2) make this problem especially challenging. An exact algorithm was developed in [3], but it is rather slow, taking several minutes in some cases. This makes it of little use in a highly dynamic environment, when users may arrive and depart every few seconds.

This led us to devise a fast heuristic for the problem. An overview of the heuristic is as follows.

1. Let D be the sum of the user demands.
2. Solve a relaxation of the problem, in which, instead of $|J|$ users, we have just one user whose demand is $D + \epsilon$, where ϵ is a small positive parameter. (The relaxation can be formulated as a convex program with only $|I|$ variables.)
3. Let $p^* \in \mathbb{R}_+^{|I|}$ be the optimal power allocation for the above relaxation.
4. "Freeze" the power allocation p^* and attempt to find a feasible solution to the real problem that uses that allocation. (This can be done by solving a 0-1 linear program with $|I||J|$ variables, or with a constructive heuristic.)
5. If a feasible solution is found, try to improve its energy efficiency. (This can be done by solving another convex program with only $|I|$ variables.)
6. If no feasible solution is found, repeat steps to 2 to 5 several times, using binary search to find a promising value for ϵ.

This heuristic turns out to be remarkably effective, being able to solve realistic instances (following the IEEE 802.16 standard) to within 1% of optimality within a couple of seconds. This is fast enough to be useful in practice.

References

1. Fazel, K., Kaiser, S.: Multi-Carrier and Spread Spectrum Systems (2nd edn). Chichester: Wiley (2008)
2. Letchford, A.N., Ni, Q., Zhong, Z.:An exact algorithm for a resource allocation problem in mobile wireless communications. Comput. Optim. Appl. 68, 193–208 (2017)
3. Letchford, A.N., Ni, Q., Zhong, Z.: Bi-perspective functions for mixed-integer fractional programs with indicator variables. Working paper, Department of Management Science, Lancaster University (2017)
4. Liu, Y.F., Dai, Y.H.: On the complexity of joint subcarrier and power allocation for multi-user OFDMA systems. IEEE Trans. Signal Process. 62, 583–596 (2014)
5. Ting, T, Chien, S.F., Yang, X.S., Lee, S.:Analysis of quality-of-service aware orthogonal frequency division multiple access system considering energy efficiency. IET Commun. 8, 1947–1954 (2014)
6. Xiao, X., Tao, X., Lu, J.: QoS-aware energy-efficient radio resource scheduling in multi-user OFDMA systems. IEEE Commun. Lett. 17, 75–78 (2013)
7. Xiong, C., Li, G., Zhang, S., Chen, Y., Xu, S. Energy-and spectral-efficiency tradeoff in downlink OFDMA networks. IEEE Trans. Wirel. Commun. 10, 3874–3886 (2011)
8. Zarakovitis C, C.C., Ni, Q.: Maximising energy efficiency in multi-user multi-carrier broadband wireless systems: convex relaxation and global optimisation techniques. IEEE Trans. Veh. Technol. 65, 5275–5286 (2016)

Multi-start local search procedure for the maximum fire risk insured capital problem

M. I. Gomes[1], L. B. Afonso[1], N. Chibeles-Martins[1], J. M. Fradinho[2]

[1] Centre of Mathematics and Applications, Faculty of Science and Technology, Nova University of Lisbon, Portugal mirg@fct.unl.pt

[2] Department of Mathematics, Faculty of Science and Technology, Nova University of Lisbon, Portugal

Abstract. A recently European Commission regulation requires insurance companies to determine the maximum value of insured fire risk policies of all buildings that are partly or fully located within circle of a radius of 200 meters. This work proposes a multi-start local search meta-heuristics developed to solve the real case of an insurance company having more than 400 thousand insured buildings in mainland Portugal. A random sample of the data set was used and the meta-heuristics solutions were compared with the optimal solution of a MILP model based on the Maximal Covering Location Problem. The results show the proposed approach to be very efficient and effective in solving the problem.

Keywords: Meta-heuristics, Local Search, Solvency II, Continuous Location Problem

Recently the European Union (EU) has published a new legistative programme - Solvency II - aiming at the harmonization of insurance industry across the European market and defining a policyholders protection framework that is risk-sensitive. Among other aspects Solvency II comprises risk-based capital requirements that need to be allocate in order to ensure the financial stability of insurance companies with assets and liabilities valued on a market consistent basis (the Solvency Capital Requirement, SRC) This work focus on the man-made catastrophe risk which comprises extreme events directly accountable to men. Specifically it addresses the capital requirement for fire risk assuming 100% damage on the total sum of the capital insured for each building located partly or fully within a 200 meters radius [1]. Until now and to best of the authors knowledge, the choice of 200 meters as the radius for the concentration was based on statistics and expert judgement.

The problem can be stated as "find the centre coordinates of a circle with a fixed radius that maximizes the coverage of total fire risk assured". This can be viewed as a particular instance of the Maximal Covering Location Problem (MCLP) with fixed radius [2]. Under the assumption that demand point is either covered or not by the facility, it has been proven that a discrete and finite set contains an optimal coverage solution [3].

This work has been motivated by the real case of an insurance company that, having more than 400 thousand buildings in Portugal covered by a fire insurance policy, needs to determine the maximum accumulated risk within a circle with 200 meters radius. The number of points makes impossible to use a 01 linear programming model since only "super" computers might be able to cope with such an amount of data. An algorithm has to be designed so that the insurance company could use it at least once a year. Therefore, we have developed a meta-heuristic - the Fire Risk Meta-heuristic (FRM) - inspired by the pattern search method proposed by Custódio and Vicente [4] that can be run in an ordinary desktop computer.

The FRM is a multi-start local search procedure where intensification and exploration strategies have been defined. In a nutshell, this procedure can be stated as: given an initial coordinate point (randomly selected) for the circle centre, determine the total fire risk within a k meters radius; generate and evaluate four neighbourhood points by increasing/decreasing each coordinate by a Δ value; make the best neighbourhood point as the new center. The step size Δ_i varies according to the quality of the neighbour solutions. Two stopping rules have been defined: one for the local search procedure and one for the multi-start algorithm. For the local search procedure one assumes as stopping criteria a small value for Δ_i, Δ_{min}. The meta-heuristic stopping criteria has two components: minimum number of iterations, i_{min}, and maximum fire risk of a single building, $Best$. The i_{min} is set empirically so that an adequate exploration of the search space is performed. The second component assures that the optimal circle must have a total fire risk (the objective function value) greater than the largest fire risk associated to a single building.

Given the volume of data for a national study and being this work a first step towards the development of an optimization approach, we confined the study to Lisbon area. The Insurance Company provided a data set which encompasses the chosen geographical area and has 46 843 buildings (points). Each points is defined by the two geographical coordinates (longitude and latitude) and the fire capital insured. The maximal covering location problem (MCLP), as proposed in Farahani et al. [5], was applied and proved to be unable to solve real instance since 46 thousand points leads to out of memory issues. A random sample of tractable size (1000 points) was then use to validate the FRM. The meta-heuristic parameters were set to $\Delta_0 = 200 \cdot 2^7 = 25\,600$ m, $\Delta_{min} = 50$ m and $i_{min} = 50\,000$ iterations and 40 runs were performed. The starting point coordinates were randomly chosen while assuring an adequate covering the search space. The results showed that only two different values for fire risk were found, with the best one coming out in 39 of the 40 runs. All 39 solutions have different centre coordinates, which is a consequence of the assumption that the centre coordinates could be any point in the chosen area. This highest fire risk value improves in 15% the optimal value found by the MCLP model ($33\,323, 56$ *vs.* $28\,857, 46$). The MCLP solution was also found by the meta-heuristic at run 19 and is the lowest reported value.

The meta-heuristic has been applied to a data set of more than 46 thousand points [6] and is being currently used by the insurance company. As future work, two main directions will be persued. Firstly, the computational experiments will be extended and the meta-heuristic solution quality will be acces with regard to the 01 integer formulation proposed by Mehrez, in 1983, [7]. Secondly, the meta-heuristic will be apploed to the 400 thousand points of mainland Portugal.

References

[1] European Insurance and Occupational Pensions Authority (25 July 2014). *The underlying assumptions in the standard formula for the Solvency Capital Requirement calculation.*

[2] Plastria F. (2002). Continuous covering location problems. In *Drezner Z, Hamacher HW, editors. Facility location: applications and theory.* Berlin: Springer, pages 37–79.

[3] Church RL, ReVelle CS (1974) The maximal covering location problem. *Pap Reg Sci Assoc* 32(1): 101–118.

[4] Custódio, A. L., and Vicente, L. N. (2007). Using sampling and simplex derivatives in pattern search methods. *SIAM Journal on Optimization*, 18(2), 537–555.

[5] Farahani, R.Z., Asgari, N., Heidari, N., Hosseininia, M., Goh, M., (2012). Covering problems in facility location: A review. *Comput. Ind. Eng.* 62: 368–407.

[6] Afonso, L.B. , Fradinho, J.M. Chibeles-Martins, N., Gomes, M.I. (2018) Metaheuristic method to evaluate the fire risk sub-module in Solvency II. (submitted)

[7] Mehrez, A., (1983). A note on the linear integer formulation of the maximal covering location problem with facility placement on the entire plane. *J. Reg. Sci.* 23: 553–555.

Selective and dynamic distribution - Petrol Secondary Distribution

Adnane ELHILALI ALAOUI[1], Pr. Youssef BENADADA[2]

[1,2]Smart Systems Laboratory, Rabat IT Center, ENSIAS, Mohammed V University in Rabat
Madinat Al Irfane, BP 713, Agdal Rabat, MOROCCO
elhilali.alaoui@gmail.com, youssef.benadada@um5.ac.ma

Abstract.

Oil products are essential to meet society's everyday needs. The downstream oil sector plays a pivotal role in powering countries economic engine by supplying the transportation fuels and the crucial feedstocks for other industry sectors and processes which greatly contribute to the nation's economic growth. Thus, petroleum products such as gasoline, kerosene, diesel and heating fuel are nowadays essential strategic products, they must be permanently available to satisfy the growing and immediate needs of professionals and particulars consumers. However, in the face of market liberalization, unpredictable oil prices fluctuations and the risk of stocks depreciation, distribution companies are forced to reduce their stocks levels as much as possible, some even tend to work just in time. This new approach to managing oil distribution requires some specific adaptation of the decision support systems used by dispatchers to prepare delivery programs. This problem of fuel distribution and replenishment of service stations has aroused the interest of several researchers and has been the subject of many research and publication works in recent decades, generally geared towards the optimization of the vehicles routes in order to minimize the global cost. The purpose of this work is the study of a problem of selective and dynamic transport, it is the case where the available resources are not sufficient to satisfy all the needs of the customers and where urgent requests may arise during the execution of the initial program, this is the most common case in the business environment. Our goal is to determine the optimal set of customers to serve, the vehicles to use and the routes to be done, respecting all the constraints of safety, capacity and time, in order to improve the quality of service of the company, to avoid or minimize stockouts at customers, and maximize their satisfaction.

Keywords: VRP, TOP, Petrol Secondary Distribution

1 Problematic

Our problem is to define a distribution policy that is:

- ***Selective:*** we have to select from among all the customers who have expressed a need, a limited set of customers to serve, and adjust their orders, taking into account logistic capacities, available stocks and some others criteria

- ***Dynamic***: that is, designing an adaptable and modifiable delivery program, which should be able, to take into consideration new urgent orders received during its execution

1.1 Constraints

The set of constraints to consider:

- Homogeneous fleet of m compartmented tank trucks,
- Each vehicle contains a set of compartments of different volume,
- Each compartment must contain a single product,
- Each vehicle in service must start its tour and finish it at the logistics center,
- All trucks that visit a customer or depot must leave it,
- The truck must respect a succession of customers,
- Each truck must pass through one and only one loading depot,
- The sum of the volumes delivered on the same tour must not exceed the total capacity of the vehicle,
- For security reasons an open compartment must be totally emptied at the same customer,
- A customer can be delivered partially, quantity delivered can be less than the requested quantity.

1.2 Objective

As objective we intend to seek a compromise between maximizing the company's profile and maximizing the quality of service customers given all the parameters and variables defined above,

2 Conclusion

Our problem is NP-hard, we will adopt the Ant Colony Optimization algorithm (ACO), which have proved to be very effective and efficient in problems of high complexity (NP-hard) in combinatorial optimization and then compare the experimental results with those of the literature. We will present a dynamic resolution that will be able to take into account real-time data changes and provide optimal results in a short adequate time. The numerical results will be published in a future article.

3 References

1. Avella, P., Boccia M, Sforza A, 2004. Solving a Fuel Delivery Problem by Heuristic and Exact Aproaches. European Journal of Operational Research 2004, 152:170-9.
2. Ben Abdelaziz, F., Roucairol, C., Bacha, C., 2002. Deliveries of liquid fuels to SNDP gas stations using vehicles with multiple compartments. In:Systems Man and Cybernetics IEEE International Conference,

Descent with mutations applied to the linear ordering problem

Olivier Hudry

LTCI - Telecom ParisTech, 46, rue Barrault, 75634 Paris, France
{olivier.hudry@telecom-paristech.fr}
http://www.infres.enst.fr/~hudry

Abstract. We study here the application of the "descent with mutations" metaheuristic to the linear ordering problem. We compare this local search metaheuristic with another very efficient metaheuristic, obtained by the hybridization of a classic simulated annealing with some ingredients coming from the noising methods. The computational experiments on the linear ordering problem show that the descent with mutations provides results which are comparable to the ones given by this improved simulated annealing, or even better, while the descent with mutations is much easier to design and to tune, since there is no parameter to tune (except the CPU time that the user wants to spend to solve his or her problem).

Keywords: Combinatorial optimization, metaheuristics, simulated annealing, noising methods, linear ordering problem, median order, Slater's problem, Condorcet-Kemeny's problem

1 Introduction

We deal here with a metaheuristic called "descent with mutations" (DWM). This method looks like the usual descent, but with random elementary transformations which are performed, from time to time, in a blind way, in the sense that they are accepted whatever their effects on the function f to optimize (such an elementary transformation performed without respect to its effect on f will be called a *mutation* in the sequel). The density of performed mutations decreases during the process, so that the method at its end is the same as a classic descent.

We study the application of DWM to two problems arising from the field of the aggregation and the approximation of binary relations: the approximation of a tournament (i.e. an oriented complete graph) by a linear order (i.e. a transitive tournament) at minimum distance (this problem is also known as *Slater's problem*) and the aggregation of linear orders into a median linear order (this problem is sometimes called *Condorcet-Kemeny's problem*). Both can be represented by the *linear ordering problem* (LOP in the following). In LOP, we are given a weighted tournament $T = (V, A)$ and we look for a subset of A with a minimum weight such that reversing the elements of A in T makes T become transitive, i.e. become a linear order. These problems are NP-hard.

2 Principle of DWM

As the other metaheuristics, DWM is not designed to be applicable to only one combinatorial problem, but to many of them. Such a problem can be stated as follows:

$$\text{Minimize } f(s) \text{ for } s \in S,$$

where S is assumed to be a finite set and f is a function defined on S; the elements s of S will be called *solutions*.

As many other metaheuristics, DWM is based on *elementary transformations*. A *transformation* is any operation changing a solution into another solution. A transformation will be considered as *elementary* (or *local*) if, when applied to a solution s, it changes one feature of s without modifying its global structure much. For instance, if s is a binary string, a possible elementary transformation would be to change one bit of s into its complement.

DWM can be described as follows:

- Repeat:
  - with a certain probability, apply an arbitrary elementary transformation (irrespective improvement or worsening: this is a mutation)
  - otherwise, apply an elementary transformation which brings an improvement
- until a given condition is fulfilled.

During the process, we apply less and less mutations. Observe that removing the first instruction of the loop (i.e. the possibility to perform a mutation), we recover the basic scheme of a descent.

3 Application of DWM to LOP

The application of DWM to LOP requires the definition of an elementary transformation. Let O be the current linear order to which we want to apply the elementary transformation. This one consists in considering another linear order O' obtained from O by moving a vertex v from its current place in O to another place: if v is ranked i in O, its rank will be $j \neq i$ in O', and the remaining of O' is the same as in O.

The experimental results of DWM applied to LOP are compared with those provided by repeated descents and by a method based on simulated annealing. They show that DWM provide very good results, with about the same quality (usually even better) as the ones obtained by an improved version of simulated annealing, which proved to be already very efficient, within the same CPU time. But, beyond this efficiency, the main advantage of DWM is that this metaheuristic is much easier to design and to apply since there is no parameter to tune (except the CPU time, which in its turn defines the number of iterations performed by the method)!

THA1 : Graph Structures and Polyhedra

- A polyhedral view to vector domination and limited packing
  Jose Neto.

- Parameterized Algorithms for Module Map Problems
  Frank Sommer, Christian Komusiewicz.

- The minimum rooted-cycle cover problem
  Denis Cornaz, Youcef Magnouche.

A polyhedral view to vector domination and limited packing

José Neto

Samovar, CNRS, Telecom SudParis, Université Paris-Saclay,
9 rue Charles Fourier, 91011 Evry, France
Jose.Neto@telecom-sudparis.eu

Given an undirected simple graph $G = (V, E)$ and integer values $f_v, v \in V$, a node subset $D \subseteq V$ is called an *f-tuple dominating set* if, for each node $v \in V$, its closed neighborhood intersects D in at least f_v nodes. The natural complementary notion to f-tuple domination is that of f-limited packing: a node subset $D \subseteq V$ is called an *f-limited packing* if the closed neighborhood of each node $v \in V$, intersects D in at most f_v nodes. The concepts of domination and limited packing naturally arise in location problems for the strategic placement of facilities in a network.

The *minimum weight f-tuple dominating set problem* in G can be stated as follows: Given a vector of node weights $w \in \mathbb{R}_+^{|V|}$ and $f \in \widehat{\mathcal{F}}_G$ with $\widehat{\mathcal{F}}_G = \{f \in \mathbb{Z}_+^{|V|} : 0 \leq f_v \leq d_v + 1, \forall v \in V\}$, where d_v stands for the degree of node v, find a minimum weight f-tuple dominating set of G, i.e. find a node subset $D \subseteq V$ such that D is an f-tuple dominating set and the weight of D: $\sum_{v \in D} w_v$, is minimum. This problem may be formulated as the following integer program.

$$(IP) \begin{cases} \min \sum_{v \in V} w_v x_v \\ s.t. \\ \qquad \sum_{u \in N[v]} x_u \geq f_v, \forall v \in V, \\ \qquad x \in \{0,1\}^{|V|}, \end{cases}$$

where $N[v]$ represents the closed neighborhood of node v. We investigate the polyhedral structure of the convex hull of the feasible region of the formulation (IP) , i.e. the polytope that is defined as the convex hull of the incidence vectors in $\mathbb{R}^{|V|}$ of the f-tuple dominating sets in G.

Some specific families of facet-defining inequalities are presented. In particular, we provide a complete formulation for the case of trees. Two corollaries of this result are a proof of a conjecture present in the literature on the formulation of the 2-tuple dominating set polytope (i.e. the f-tuple dominating set polytope for the case when $f_v = 2, \forall v \in V$) of trees, and a new family of (generally exponentially many) inequalities which are valid for the f-tuple dominating set polytope of any graph and that can be separated in linear time. We also prove that, for the case of trees, the integrality gap of the linear relaxation of the formulation (IP) is upper bounded by $\frac{4}{3}$, and that this bound is asymptotically tight.

Parameterized Algorithms for Module Map Problems

Frank Sommer and Christian Komusiewicz

Fachbereich für Mathematik und Informatik, Philipps-Universität Marburg, Marburg, Germany
{fsommer,komusiewicz}@informatik.uni-marburg.de

1 Problem Definition

We investigate the MODULE MAP problem which has applications in computational biology [1, 3] and is defined on edge-colored graphs $G = (V, E_b, E_r)$ with a set E_b of *blue edges* and a set E_r of *red edges*. We say that G fulfills the *cluster property* if the graph $G_b := (V, E_b)$ is a cluster graph, that is, a disjoint union of cliques called *clusters*. Further, G fulfills the *link property* if for distinct clusters A and B of G_b, the graph $G_r[A \cup B]$, with $G_r := (V, E_r)$, is either edgeless or complete bipartite with partite sets A and B. In our computational problem, we aim to establish both properties.

Definition 1. *A graph $G = (V, E_b, E_r)$ is a* module graph *if G satisfies the cluster property and the link property.*

A module graph is shown in Fig. 1. Our aim is to find a module graph which can be obtained from the input graph G by as few edge transformations as possible. This leads to the following problem, here formulated in its decision version.

MODULE MAP
Input: A graph $G = (V, E_b, E_r)$ and a non-negative integer k.
Question: Can we transform the graph $G = (V, E_b, E_r)$ into a module graph by deleting or adding at most k red and blue edges?

In practice, it is useful to consider edge-weighted versions of the problem, where the input includes a weight function $g : V^2 \to \mathbb{N}^+$ on vertex pairs. The higher the weight, the more confidence we have in the observed edge type. This gives the following problem:

WEIGHTED MODULE MAP
Input: A graph $G = (V, E_b, E_r)$ with edge weights $g : V^2 \to \mathbb{N}^+$ and a non-negative integer k.
Question: Can we transform the graph G into a module graph by edge transformations of cost at most k?

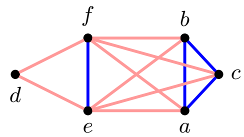


Fig. 1: A module graph with the clusters $\{a, b, c\}$, $\{d\}$, and $\{e, f\}$.

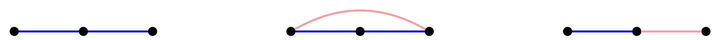


Fig. 2: The forbidden induced subgraphs for module graphs. From left to right: a blue P_3, consisting of two (dark) blue edges, a two-colored K_3, consisting of two blue and one (light) red edge and a two-colored P_3, consisting of one blue and one red edge.

2 Our Results

We first provide a characterization of module graphs in terms of forbidden induced subgraphs, see Fig. 2.

Theorem 1. *A graph G is a module graph if and only if G has no blue P_3, no two-colored K_3, and no two-colored P_3 as induced subgraph.*

Using this characterization, we obtain a simple linear-time algorithm for every fixed value of k.

Proposition 1. Module Map *can be solved in* $\mathcal{O}(3^k \cdot (|V| + |E|))$ *time.*

To obtain a better running time dependency on k, we adapt a branching algorithm for Cluster Editing [2] which is a special case of Module Map.

Theorem 2. Weighted Module Map *can be solved in* $\mathcal{O}(2^k \cdot |V|^3)$ *time.*

Finally, we show that in polynomial time we can obtain an equivalent instance whose number of vertices is bounded by a polynomial function in k.

Theorem 3. Weighted Module Map *admits a problem kernel of* $\mathcal{O}(k^2)$ *vertices which can be found in* $\mathcal{O}(|V|^3 + k \cdot |V|^2)$ *time.*

References

[1] D. Amar and R. Shamir. Constructing module maps for integrated analysis of heterogeneous biological networks. *Nucleic Acids Research*, 42(7):4208–4219, 2014.

[2] S. Böcker, S. Briesemeister, Q. B. A. Bui, and A. Truß. Going weighted: Parameterized algorithms for cluster editing. *Theoretical Computer Science*, 410(52):5467–5480, 2009.

[3] I. Ulitsky, T. Shlomi, M. Kupiec, and R. Shamir. From E-maps to module maps: dissecting quantitative genetic interactions using physical interactions. *Molecular Systems Biology*, 4(1):209, 2008.

The minimum rooted-cycle cover problem

D. Cornaz and Y. Magnouche*

March 15, 2018

Abstract

Given an undirected rooted graph, a cycle containing the root vertex is called a rooted cycle. We study the combinatorial duality between vertex-covers of rooted-cycles, which generalize classical vertex-covers, and packing of disjoint rooted cycles, where two rooted cycles are vertex-disjoint if their only common vertex is the root node. We use Menger's theorem to provide a characterization of all rooted graphs such that the maximum number of vertex-disjoint rooted cycles equals the minimum size of a subset of non-root vertices intersecting all rooted cycles, for all subgraphs.

Keywords: Kőnig's theorem, Menger's theorem.

1 Introduction

Throughout $G = (V, E)$ is a simple undirected graph. The *minimum vertex-cover problem* is to find a *vertex-cover* (that is, a set $T \subseteq V$ so that every edge of G has at least one vertex in T) minimizing $|T|$. This is a very well studied NP-hard problem, equivalent to finding a maximum *stable set* (equivalently, the complement of a vertex-cover, or a *clique* in the complementary graph) [6]. In this paper, we introduce the *minimum rooted-cycle cover problem* which contains the vertex-cover problem, and which is, given a root vertex r of G, to remove a minimum size subset of $V \setminus \{r\}$ so that r is contained in no cycle anymore. The minimum vertex-cover problem is the particular case where r is adjacent with all other vertices.

If we are given a set of terminal vertices of G, with at least two vertices, the *minimum multi-terminal vertex-cut problem* is to remove a minimum number of vertices, so that no path connects two terminal vertices anymore, see [1, 2]. The weighted version of the minimum rooted-cycle cover problem contains the minimum multi-terminal vertex-cut problem which is the particular case where the neighborhood $N(r)$ of r is the set of terminal vertices with infinite weight. In turn, if we replace r by $|N(r)|$ terminal vertices $t_1, \ldots, t_k$ where $N(r) = \{v_1, \ldots, v_k\}$ and link t_i to v_i, then we obtain an instance of the minimum multi-terminal vertex-cut problem the solution of which is a solution for the original instance of the minimum rooted cycle cover problem.

*LAMSADE, University Paris-Dauphine, Place du Maréchal de Lattre de Tassigny, 75016 Paris, France. {denis.cornaz,youcef.magnouche}@dauphine.fr

Our main motivation to introduce the minimum rooted cycle cover problem is that it allows us to give short proofs of some min-max theorems, such results being fundamental in combinatorial optimization and linear programming [5]. Jost and Naves gave such results for the minimum multi-terminal vertex-cut problem in an unpublished manuscript [2] (actually we found independently this result).

The paper is organized as follows. In Section 2, we recall two classical theorems and give formal definitions. In Section 3, we give a characterization of all rooted graphs (G, r) so that the minimum number of non-root vertices intersecting all rooted cycles equals the maximum number of rooted cycles having only the root as common node, for all partial subgraphs. In Section 4, we revisit a result by Jost and Naves [2] in terms of rooted cycles. This is a structural characterization in terms of excluded minors of pseudo-bipartite rooted graphs, that is, rooted graphs satisfying the min-max equality for all rooted minors.

References

[1] N. Garg, V.V. Vazirani, M. Yannakakis, Multiway cuts in vertex weighted graphs. *J. of Algorithms* 50 (2004) 49–61.

[2] V. Jost, G. Naves, The graphs with the max-Mader-flow-min-multiway-cut property. *CoRR abs/1101.2061* (2011).

[3] D. Kőnig, Graphok és alkalmazásuk a determinánsok és a halmazok elméletére [Hungarian], *Mathematikai és Természettudományi Értesitő* 34 (1916) 104–119.

[4] K. Menger, Zur allgemeinen Kurventheorie, *Fundamenta Mathematicae* 10 (1927) 96–115.

[5] A. Schrijver, Theory of linear and integer programming, *John Wiley & Sons Ltd* England (1986).

[6] A. Schrijver, Combinatorial Optimization: Polyhedra and Efficiency, *Springer-Verlag Berlin Heidelberg* (2003).

THA2 : Scheduling

- MIP Formulations for Just-in-Time Scheduling with Common Due-Date
  Anne-Elisabeth Falq, Pierre Fouilhoux, Safia Kedad-Sidhoum.

- A generalized Tool Switching Scheduling Problem with Unlimited Buffers
  Khadija Hadj Salem, Yann Kieffer, Luc Libralesso.

- On scheduling a conference program
  Philippe Depouilly, Laurent Facq, Adeline Fonseca, Pierre Pesneau, Issam Tahiri, Francois Vanderbeck.

MIP Formulations for Just-in-Time Scheduling with Common Due-Date

Anne-Elisabeth Falq, Pierre Fouilhoux, Safia Kedad-Sidhoum

Sorbonne Université, CNRS UMR 76006, LIP6, 4 Place Jussieu, 75005 Paris, France
{anne-elisabeth.falq,pierre.fouilhoux,safia.kedad-sidhoum}@lip6.fr

We consider a set of tasks J that have to be processed non-preemptively on a single machine around a common due-date d. The processing time of a task $j \in J$ is denoted by p_j. A schedule will be given by the task completion times $(C_j)_{j\in J}$. A task $j \in J$ is early (resp. tardy) if $C_j \leq d$ (resp. $C_j > d$). The earliness (resp. tardiness) of any task $j \in J$ is given by $\max(0, d–C_j)$ (resp. $\max(0, C_j–d)$). Given unitary earliness penalties $(\alpha_j)_{j\in J}$ (resp. tardiness penalties $(\beta_j)_{j\in J}$), the problem aims at finding a schedule that minimizes the sum of earliness-tardiness penalties.

This problem falls into common due-date just-in-time scheduling literature [1, 2]. To the best of our knowledge, no polyhedral approach has been considered for this class of problems. In this work, we propose MIP formulations based on some polyhedral properties.

When $d \geq \sum p_j$, the problem is said unrestrictive. In this case, using the so-called V-shaped dominance property [1], we ensure that there exists an optimal solution such that early tasks are scheduled by increasing α_j/p_j ratio while tardy tasks are scheduled by decreasing β_j/p_j ratio. We can also restrict to schedules without idle time and with a task ending exactly at d (called an on-time task). This unrestrictive problem is NP-hard even if $\alpha_j = \beta_j$ for all $j \in J$ [1].

In the general case, a small due-date can lead to the occurrence of a straddling task in every optimal schedule. The problem is still NP-hard, even if $\alpha_j = \beta_j = 1$ for all $j \in J$ [2].

1 Polyhedral Approach for the Unrestrictive Case

Queyranne [3] consider the classical scheduling problem where the objective is to minimize the weighted sum of completion times. His goal is to describe the convex hull of vectors $(C_j)_{j\in J}$ encoding valid schedules using linear inequalities and then, in the resulting polyhedron, all extreme points describe solutions. A natural way to express the non-overlapping of tasks is to use the disjunctive constraint : $\forall (i, j) \in J^2$, $i \neq j$, $C_j \geq C_i + p_j$ or $C_i \geq C_j + p_i$. However Queyranne succeeds in avoiding the introduction of binary variables and provides an exponential number of inequalities that can be separated in polynomial time. Note that some points within the interior of the polyhedron encode non-valid schedules.

In a similar way, our goal is to provide a polyhedron whose extreme points encode dominant schedule for our just-in-time problem. However, due to the earliness-tardiness partition of tasks, most of Queyranne's arguments are no longer valid.

To encode the earliness-tardiness partition, we introduce binary variables $(\delta_j)_{j\in J}$. They allow to replace C_j variables by $(e_j)_{j\in J}$ and $(t_j)_{j\in J}$ which respectively represents earliness and tardiness of tasks, thus inducing a linear objective function. Thanks to some dominance properties, the solution space can be restricted to non-idle time schedules with one on-time task. Adding some linearization variables, we provide a familiy of linear inequalities, derived from Queyranne's ones, that express the non-overlapping of tasks before d and another family for the tasks after d.

Here again, the resulting formulation is not a classical MIP formulation, since some points within the interior of the polyhedron, even if they are integer, encode non-valid schedules. The number of inequalities is exponential, but we reduce the associate separation problem to the min-cut problem.

It is important to remark that, knowing the earliness-tardiness partition given by $(\delta_j)_{j\in J}$, an optimal solution can be easily derived by sorting the tasks on either side of d according to their ratios. Once an ordering has been set for either side, we provide a linear function based on variables δ_j (without e_j and t_j) corresponding to the cost of a non-idle and V-shaped schedule. This leads to a compact formulation. Our first experimental results with Cplex show that this second formulation outperforms the first one.

2 Polyhedral Approach for the General Case

To extend the previous formulation to the general case, we should be able to find a dominant solution set for which $(\delta_j)_{j\in J}$ is a sufficient encoding. Unfortunately it can happen that straddling tasks occur in all optimal schedules. Hence, even if we know which are the early tasks of an optimal schedule, we cannot say which task is the first among the tardy ones.

Alternatively, we extend the first formulation. We add a continuous variable a corresponding to the earliness of the last early task. We replace e_j by e'_j (resp. t_j by t'_j) which express the earliness (resp. tardiness) of task j regarding $d-a$ (which becomes the reference point). In the same way as for the unrestrictive case, we provide inequalities (similar to Queyranne's ones) to express the non-overlapping of tasks on each side of $d-a$.

We devise a Branch-and-Cut algorithm in which the non-overlapping inequalities are separated and we compare our three formulations.

References

1. Hall, N.G., Posner, M.E.: Earliness-tardiness sceduling problems, 1: Weighted deviation of completion times about a common due date. Operations Research **Vol 39** (Sep-Oct 1991) 836–846
2. Hoogeveen, J., van de Velde, S.: Scheduling around a small common due date. European Journal of Operational Research **Vol 55** (1991) 237–242
3. Queyranne, M.: Structure of a simple scheduling polyhedron. Mathematical Programming **Vol 58** (1993) 263–285

A generalized Tool Switching Scheduling Problem with Unlimited Buffers

Khadija HADJ SALEM[1], Yann KIEFFER[2], and Luc LIBRALESSO[1]

[1] GSCOP Laboratory, 46 avenue Félix Viallet, 38031 GRENOBLE, France
[2] LCIS Laboratory, 50 rue Barthélémy de Laffemas, 26902 VALENCE, France
`khadija.hadj-salem@grenoble-inp.fr` ✉

This paper considers the `Minimum Completion Time of 3-PSDDP` (MCTP), a non-standard scheduling optimization problem arising from the design of efficient embedded vision systems [3]. It can be also seen as a generalized Tool Switching Problem (ToSP), which is a $\mathcal{NP}$-hard combinatorial optimization problem arising from computer and manufacturing systems [4].

We are given a set of output tiles $\mathcal{Y} = \{1, \ldots, Y\}$ (also called tasks) to be computed sequentially without preemption on one machine, and a set of input tiles $\mathcal{X} = \{1, \ldots, X\}$ (also called prerequisites). Each task y requires a subset of prerequisites $\mathcal{R}_y$, where $\mathcal{R}_y \subseteq \mathcal{X}$, to be loaded on a second machine with an unlimited (infinite number) buffers capacity Z. Also, the duration of a prefetch step α, and that of a computation step β, have to be given as inputs.

The MCTP consists of simultaneously determining the processing schedule of tasks (a permutation of the tasks) and a corresponding schedule of prefetches (a permutation of the prerequisites) in order to minimize the total completion time C_{max}.

Despite the existence of several trivial variants which can be solved in polynomial time, we have proved that MCTP is $\mathcal{NP}$-Hard, by giving a polynomial reduction from the k-`weak visit` problem described in [1].

An example of an instance of the MCTP is shown in Fig. 1. Each column in the incidence matrix refers to a particular task in the considered set $\mathcal{Y} = \{a, b, c, d, e\}$ and each row refers to a particular prerequisite in the considered set $\mathcal{X} = \{1, 2, 3, 4\}$.

$$
\begin{array}{c}
\\ 1 \\ 2 \\ 3 \\ 4
\end{array}
\begin{array}{c}
\begin{matrix} a & b & c & d & e \end{matrix} \\
\begin{pmatrix}
1 & 1 & 0 & 1 & 0 \\
0 & 1 & 1 & 1 & 1 \\
0 & 0 & 1 & 1 & 0 \\
1 & 0 & 1 & 0 & 0
\end{pmatrix}
\end{array}
$$

Fig. 1. Incidence matrix

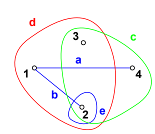


Fig. 2. Hyper-graph diagram

2	1	3	4		
	e	b	d	a	c

Fig. 3. An optimal solution

As shown in Fig. 2, it is then easy to see that this can be represented as a hyper-graph, where the set of vertices $V = \mathcal{X}$ and the set of hyper-edges $E = \mathcal{Y}$.

An optimal solution ϕ is given in Fig. 3, in which ϕ requires $N = 4$ prefetches and $C_{max} = 6$ units of time.

We developed two approaches for solving the MCTP. The first one is a MILP model, in which some variables are binary and the others are continuous. In this model, the objective function minimizes the total completion time C_{max}, while respecting the following constraints.

$$\sum_{y\in\mathcal{Y}} c_{yj} = 1 \qquad \forall\ j \in \mathcal{Y} \tag{1}$$

$$\sum_{j\in\mathcal{Y}} c_{yj} = 1 \qquad \forall\ y \in \mathcal{Y} \tag{2}$$

$$\sum_{x\in\mathcal{X}} d_{xi} = 1 \qquad \forall\ i \in \mathcal{X} \tag{3}$$

$$\sum_{i\in\mathcal{X}} d_{xi} = 1 \qquad \forall\ x \in \mathcal{X} \tag{4}$$

$$u_j - t_i \geq \alpha - \Lambda * (3 - r_{xy} - c_{yj} - d_{xi}) \qquad \forall\ y, j \in \mathcal{Y},\ x, i \in \mathcal{X} \tag{5}$$

$$t_{i-1} + \alpha \leq t_i \qquad \forall\ i \in \mathcal{X}\backslash\{1\} \tag{6}$$

$$u_{j-1} + \beta \leq u_j \qquad \forall\ j \in \mathcal{Y}\backslash\{1\} \tag{7}$$

$$C_{max} \geq \max_{j\in\mathcal{Y}} u_j + \beta \tag{8}$$

$$c_{yj}, d_{xi} \in \{0, 1\} \qquad \forall\ y, j \in \mathcal{Y},\ x, i \in \mathcal{X} \tag{9}$$

$$u_j, t_i \geq 1 \qquad \forall\ j \in \mathcal{Y},\ i \in \mathcal{X} \tag{10}$$

Computational experiments, using a data-sets possessing different characteristics available at `http://www.unet.edu.ve//~jedgar/ToSP/ToSP.htm`, indicate that instances involving up to 20 tasks and 20 prerequisites can be solved optimally using the MILP approach (using Gurobi optimizer version 7.5.1).

For solving large-sized problems, we have develop a Simulated Annealing (SA), a meta-heuristic that has been widely used in optimization and present in most of the textbooks [2]. SA algorithms can be seen as iterated local searches, where moves decreasing the solution quality are allowed with some probability. This mechanism helps a escaping for local optima.

We compared the numerical results obtained by the SA, the MILP using Local Solver 7.5 and the heuristics described by [3]. The results show that the SA performance is promising.

References

[1] C. Arbib, M. Flammini, and E. Nardelli. How to survive while visiting a graph. *Discrete applied mathematics*, 99(1-3):279–293, 2000.

[2] M. Gendreau and J.Y. Potvin. *Handbook of metaheuristics*, volume 2. 2010.

[3] K. Hadj Salem, Y. Kieffer, and M. Mancini. *Meeting the Challenges of Optimized Memory Management in Embedded Vision Systems Using Operations Research*, pages 177–205. Springer International Publishing, Cham, 2018.

[4] C.S. Tang and E.V. Denardo. Models arising from a flexible manufacturing machine, part i: Minimization of the number of tool switches. *Operations Research*, 36(5):767–777, 1988.

On scheduling a conference program

Philippe Depouilly (U. of Bordeaux & CNRS), Laurent Facq (U. of Bordeaux & CNRS), Adeline Fonseca (Inria-Realopt), Pierre Pesneau (U. of Bordeaux & Inria-Realopt) Issam Tahiri (Inria-Realopt), and F. Vanderbeck (U. of Bordeaux & Inria-Realopt)

January 18, 2018

Abstract

Optimizing the schedule of the scientific program of a large conference such as the International Symposium in Mathematical Programming (ISMP) is quite challenging given the multitude of objectives, the lack of data, and the scale of the instance : there are about 520 sessions (with 3 or 4 talks in each) to schedule over a dozen time slots with 40 parallel tracks. Starting with sessions that have defined by the scientific committee, our scheduler outputs time slot and room assignment for each session. Its main goal is to spread the program evenly over the time horizon to maximize the offer that our public can attend. In this aim, the first objective is to minimize the number of parallel tracks in each thematic area. The second issue is to avoid to schedule in parallel sessions that are destined to a same public. Although the latter can not be measured precisely, we record referees' and attendees' inputs to define both hard and soft conflict constraints between sessions. The third measure of the quality of the program is the extend to which the scientific interest is evenly spread so as to avoid having all the hight profile talks into some time slots and none in other time slots. This goal is modeled as a min max of the interest measure for each time slot. All these goals are driving the optimization in the same direction of a well balance program. We develop a hierarchical optimization approach based on solving a sequence of mixed integer programs, that do scale up to our typical input size.

keywords: Planning, Coloring, Bin packing, Conflict Graph.

Aknowledgment: The contributions of Master Students have been essential to this project: Yacine Al-najjar, Mickael Gaury, Johan Leveque, and Nicolas Murcia

THA3 : Network Design

- Design of Multi-Echelon, Collaborative and Sustainable Supply Chain Networks
  Ouhader Hanan, El Kyal Malika.

- Polyhedral Investigation of The Proactive Countermeasure Selection Problem
  Ridha Mahjoub, Mohamed Yassine Naghmouchi, Nancy Perrot.

- Domain Creation in Heterogeneous Mobile Networks: Models and Algorithms
  Wesley Da Silva Coelho, Amal Benhamiche, Nancy Perrot.

Design of Multi-Echelon, Collaborative and Sustainable Supply Chain Networks

Hanan Ouhader[1] , Malika El kyal[1]

[1] National School of Applied Sciences Ibn Zohr University, Agadir
Morocco
ouhader@gmail.com

Abstract. In this paper, we study the impact of network structure in the economic and environmental benefits to assess the opportunities offered by the horizontal collaboration. We analyse three network configurations which differ in the location of satellites (intermediate depots) within the distribution area. The goal is to present to decision makers a preliminary mechanism to gain general insights into beneficial network structure for the coalition. we exploit a bi-objective mathematical model for a two echelon location routing problem (2E-LRP) to test if partners fit for the collaboration or not and if opportunity for each partner to make economic and environmental benefits exists. Extended known instances reflecting the real distribution in urban area are regenerated to evaluate several goods' delivery strategies. Shapley value method, belonging to the field of cooperative game theory, is used to allocate cost and CO2 emissions to partners of the coalition. This approach proposes a coalition formation mechanism allowing the decision makers to measure the sustainability performance of partners during the design phase of the network..

Keywords. Horizontal collaboration, network design, Sustainable urban road transport, Two-echelon Location Routing problem, Multi-objective optimization

1 Introduction

According to the Accenture and World Economic Forum Report 2016 , logistics and transportation activities contribute approximately to 13% of total greenhouse gas (GHG) emissions and 57% of the transport emissions came from road freight. Experts estimate that urban goods movements account for 20 % to 30% of total vehicle kilometers driven. Accordingly, the urban road transport sector can play a considerable role in reducing emissions. To ensure that environmental, social, and economic considerations are factored into decisions affecting urban transportation activity is the goal of sustainable urban transportation [1].

Several strategies with the aim of improving efficiency and sustainability from urban road transport have been suggested both in practice and in the academic literature. Logistics collaboration is gaining traction as a one of the key policies to assure this mission.

We talk about collaborative supply chain when two players (or more) of the "Supply Chain" seek to optimize together the logistics of the distribution circuit in which they are linked [2].

Logistics collaboration was studied in three main areas: Vertical, horizontal and lateral collaboration. The vertical collaboration occurs between members of the same chain value (industrial and distributor). Horizontal collaboration occurs between companies (may be competitors or not) that can provide goods or complementary services [3]. Nevertheless, less attention has given to research on horizontal logistics collaboration ([4], [5], [6] ,[7] ,[8]).

There are various areas of research and opportunities in the collaborative supply chain field. From the transportation management's point of view , the most recent literature review articles on horizontal collaboration among supply chain partners ([9] ,[6],[10]) and recent studies proposing quantitative models for establishing horizontal collaboration, reveal the majorities of papers on the subject are based on vehicle routing problem by proposing models only for the operational level of the supply chain and assuming that strategic facility location decisions have met in a prior step and cannot be modified. Also the optimization of horizontal collaborative supply chain was mainly based in single objective mathematical modeling approach dealing with economic concern and the integration of sustainability is accordingly in his infancy.

To overcome this drawback, we quantified in our previous work [11] the aggregated economic benefit of horizontal collaboration basing on a single-objective two echelons Location Routing problem (2E-LRP) model and we performed a posterior evaluation of the impact of collaboration in CO2 emissions based on travelled

Polyhedral Investigation of the Proactive Countermeasure Selection Problem

A. Ridha Mahjoub[1], M. Yassine Naghmouchi[1,2], and Nancy Perrot[2]

[1] Université Paris-Dauphine, PSL Research University, CNRS, LAMSADE, 75016, Paris, France,
{mahjoub,mohamed-yassine}@dauphine.fr
[2] Orange Labs, France, {firstname.lastname}@orange.com

The Proactive Countermeasure Selection Problem (PCSP) is defined as follows. Given 1) a directed graph $G = (V, A)$ called the *Risk Assessment Graph* [2], [1], where $V = S \cup T$, $S \cap T = \emptyset$, and with each arc $(i, j) \in A$ is associated a weight $w_{ij} \in \mathbb{R}_+$. 2) A set of available countermeasures $K = \{(t, k) : k \in K_t, t \in T\}$ such that K_t is the set of countermeasures associated with t. The placement of k on t has a positive cost $c_t^k \in \mathbb{R}_+$, and yields an increase of the weight of the ongoing arcs of t by a positive effect $\alpha_t^k \in \mathbb{R}_+^*$. 3) A positive vector $D = (d_s^t)_{s \in S, t \in T} \mathbb{R}_+$. The PCSP consists in selecting a set of countermeasures, at minimal cost, such that for each $s \in S$ and $t \in T$ the length of the $s - t$ shortest path is at least d_s^t. A bilevel model was introduced in [3] to solve the PCSP. The goal of this paper is to investigate the PCSP polytope in order to improve the algorithmic aspect.

ILP single-level reformulation and Associated Polytope A vector x induced by a solution of the set of feasible solutions of the PCSP, denoted by $S(G, K, D)$, satisfies the following constraints:

$$\sum_{u \in P, u \neq s} \sum_{k \in K_u} \alpha_u^k x_u^k \geq d_s^t - \sum_{ij \in P} w_{ij} \quad \forall s \in S, t \in T, P \in P_{s,t}, \tag{1}$$

$$0 \leq x_t^k \quad \forall (t, k) \in K, \tag{2}$$

$$x_t^k \leq 1 \quad \forall (t, k) \in K. \tag{3}$$

Inequalities (1) are called *security inequalities*. They ensure for each access point $s \in S$ for each asset-vulnerability node $t \in T$ that the length of the $s - t$ shortest path is at least d_s^t. Inequalities (2) and (3) are *the trivial inequalities*. The PCSP problem is equivalent to the following integer program $min\{c^T x | x \in \{0,1\}^{|K|}$: x satisfies $(1)-(3)\}$. We will denote by $PCSP(G, K, D)$, the polytope associated with the PCSP. A countermeasure $(t, k) \in K$ is said to be *essential* for $PCSP(G, K, D)$ if and only if the set $S(G, K \backslash \{(t, k)\}, D) = \emptyset$. We will denote by K^* the set of all essential countermeasures of $PCSP(G, K, D)$. We have the following result.

Theorem 1. *$dim(PCSP(G, K, D)) = |K| - |K^*|$, and PCSP is full dimensional if and only if $K^* = \emptyset$.*

Facial Aspect and Valid Inequalities

Theorem 2. *Let $(t, k) \in K$. 1) Inequality (2) defines a facet of $PCSP(G, K, D)$ if and only if $(t, k) \in K \backslash K^*$ and $(K \backslash \{(t, k)\})^* = \emptyset$. 2) Inequality (3) defines a facet of $PCSP(G, K, D)$ if and only if $(t, k) \in K \backslash K^*$.*

Theorem 3. *Let $s \in S$, $t \in T$ and $P \in P_{s,t}$. Inequality (1) defines a facet of $PCSP(G, K, D)$ if*

1) *For all $(u, l) \in K(P)$ $\alpha_u^l = \alpha$,*
2) *$\exists \rho \in \mathbb{R}_+$ such that $1 \leq \rho \leq |K(P)|$ and $\rho\alpha = d_s^t - V(P)$,*
3) *For all $(u, l) \in K \backslash \{K^* \cup K(P)\}$, for all $J \subset K(P)$ such that $|J| = |K(P)| - \rho$ we have $S(G, K \backslash \{J \cup \{(u, l)\}\}, D) \neq \emptyset$.*

Theorem 4. *Let $s \in S$, $t \in T$ and $P \in P_{s,t}$. Inequality (1) defines a facet of $PCSP(G, K, D)$ only if*

1) *$\exists (u, l) \in K(P)$ such that $\alpha_u^l \leq d_s^t - V(P)$,*
2) *$\exists (u, l) \in K^* \cap K(P)$ such that $\alpha_u^l \neq \frac{1}{|K(P)|}(d_s^t - V(P))$,*
3) *For all $J \subset K^* \cap K(P)$ $\sum\limits_{(u,l) \in T} \alpha_u^l \leq d_s^t - V(P)$.*

Theorem 5. *Let $(t_i, k_i) \in K \backslash K^*$, $i = 1, \ldots, n$, $1 \leq n \leq |K| - |K^*|$ such that $(t_{i+1}, k_{i+1}) \in (K \backslash \{(t_i, k_i)\})^*$ $i = 1, \ldots, n-1$. The following inequality is valid for $PCSP(G, K, D)$:*
$\sum\limits_{i=1}^{n} x_{t_i}^{k_i} \geq \lceil \frac{n-1}{2} \rceil$.

Theorem 6. *Let $(t_i, k_i) \in K \backslash K^*$, $i = 1, \ldots, n$, $1 \leq n \leq |K| - |K^*|$ such that $(t_{i+1}, k_{i+1}) \in (K \backslash \{(t_i, k_i)\})^*$ $i = 1, \ldots, n-1$. Inequality $\sum\limits_{i=1}^{n} x_{t_i}^{k_i} \geq \lceil \frac{n-1}{2} \rceil$ is facet defining if $\exists J \subset \bigcup\limits_{i=1}^{n} \{(t_i, k_i)\}$ such that $|J| = n - \lceil \frac{n-1}{2} \rceil$ and $S(G, K \backslash J, D) \neq \emptyset$, and for all $J \subset K$ such that $|J| \geq n - \lceil \frac{n-1}{2} \rceil$ we have $S(G, K \backslash J, D) \neq \emptyset$.*

Theorem 7. *Let $(t_i, k_i) \in K \backslash K^*$, $i = 1, \ldots, n$, $1 \leq n \leq |K| - |K^*|$ such that $(t_{i+1}, k_{i+1}) \in (K \backslash \{(t_i, k_i)\})^*$ $i = 1, \ldots, n-1$. Inequality $\sum\limits_{i=1}^{n} x_{t_i}^{k_i} \geq \lceil \frac{n-1}{2} \rceil$ is facet defining only if n is even, and $\exists J \subset \bigcup\limits_{i=1}^{n} \{(t_i, k_i)\}$ such that $|J| \geq n - \lceil \frac{n-1}{2} \rceil$ and $S(G, K \backslash J, D) \neq \emptyset$.*

References

1. Kheir, N., Mahjoub, A. R., Naghmouchi, M. Y., Perrot, N., Wary, J. P: Assessing the risk of complex ICT systems. Annals of Telecommunications, 1-15 (2017)
2. Naghmouchi, M. Y., Perrot, N., Kheir, N., Mahjoub, A. R., Wary, J. P: A New Risk Assessment Framework Using Graph Theory for Complex ICT Systems. In Proceedings of the 2016 International Workshop on Managing Insider Security Threats (pp. 97-100). ACM (2016)
3. Mahjoub, A. R., Naghmouchi, M. Y., Perrot, N: The proactive Countermeasure Selection Problem in Complex ICT Systems (2017)

Domain Creation in Heterogeneous Mobile Networks: Models and Algorithms

Wesley Da Silva Coelho, Amal Benhamiche, and Nancy Perrot

Orange Labs, France,
{amal.benhamiche,nancy.perrot}@orange.com,
Orange Gardens, 44 avenue de la République, 92326 Châtillon, France

Introduction

In future 5G networks, mobile User Equipment (UEs) will be able to host functions that give them new abilities such as sharing connectivity, capacity, or CPU resources with other UEs, regardless of the ongoing traditional communications. In particular, the arrival of 5G wireless technology along with the evolution of mobile users behavior and needs, make the current schema of communication (UE to Base Station) no longer optimal in terms of radio resource utilization. The Device-to-Device (D2D) communication mode is one of the new approaches presented as a promising alternative to traditional communication in cellular networks. A D2D communication is defined as a direct communication between two mobile or fix user devices, without traversing the Base Station (BS) [1]. This technology allows to reuse radio resources and to decrease end-to-end latency of local communications. Then, D2D would allow a set of UEs, interested in a given service (like video streaming, gaming, etc.) and geographically close to each other, to establish direct or multi-hop D2D communications, while ensuring the quality required by the service. In this context, a D2D service *domain* is defined as the set of UEs that are used to establish mobile communications (D2D or traditional) related to this specific service. The communication is either direct or spans multiple links (D2D or via the BS). In both cases, one or several radio resources should be allocated to every active link and the SINR (Signal-to-Interference-plus-Noise Ratio) level required by the service should be ensured.

Problem definition

We consider a set of devices in a cellular network and a traffic matrix containing a volume of data for a specific service to be sent, possibly using D2D communications (e.g. exchange of content among a subset of devices). Each pair of devices can be connected either by one or more D2D radio links (as many as the number of available radio resources), or by conventional cellular communication, through the BS. Each link is associated with a non-negative weight which is the SINR between both end devices. Each service requires a certain minimum level of quality (in terms of technical capabilities of links and devices) for the communication. The *Domain Creation Problem* (DCP) consists then in finding a partition of the device pairs into k subsets, and the radio resource assignment to the D2D links so that: (i) every pair (link, resource) is assigned to a unique domain (ii) the SINR of each pair is above the quality threshold required by the domain (iii) every demand is routed from its origin to its destination within a domain (iv)

all the types of device capacities (CPU, RAM, battery) are respected and (v) the total cost is minimum. This problem is a variant of the so-called *Routing and Wavelength Assignment* (RWA) problem [2], [3] which arises in Optical Networks. The specificities of the (DCP) are that flows are *unsplittable* (each flow has to be sent along a unique path using D2D or cellular links), several types of capacities on the devices are considered, and a radio resource assigned to an active link can be reallocated only to a distant link to avoid communication interference.

Our contribution

In this work, we formally define the (DCP) and we propose two ILP formulations to model it: an arc variable (compact) and a path variable (non-compact) formulations. The linear relaxation of the non compact formulation is solved with a column generation procedure. We further propose an efficient heuristic based on a decomposition of the problem in two problems, namely *routing problem* and *resource allocation problem*, that are solved successively. The routing problem is solved with a column generation procedure while the allocation problem is transformed into a vertex coloring problem that is solved heuristically by an improved greedy algorithm, and computing a dual bound by solving exactly the Max Clique Problem. Numerical experiments are made on instances generated thanks to realistic parameters of Orange mobile networks. Cplex 12 is used to solve the linear relaxations to optimality. Our experiments show that the heuristic approach performs well, even on large instances with up to 2100 devices and 1500 service requests on a 6 cells network. Table 1 shows the empirical quality and the efficiency of our heuristic compared to a branch-and-bound over the compact formulation. For all instances tested, the heuristic gives the optimal solution, in a very short time.

Instance	heuristic			exact formulation		
#users # links # requests	# active links	# resources	CPU (s)	# active links	# resources	CPU (s)
140_L334_D35	83	70	0.19	83	70	7816.55
U140_L346_D35	85	68	0.18	85	68	7822.87
U210_L518_D52	125	100	0.46	125	100	38292.83
U210_L501_D52	135	100	0.43	135	100	37808.41

Table 1. Comparing heuristic and non-compact formulation

This algorithm offers to the network architects a way to evaluate the efficiency of D2D technology compared to classical cellular communication, and allow to simulate scenarios of new and flexible service deployment in future 5G networks.

References

1. Asadi, A., Wang, Q., Mancuso, V.: A survey on device-to-device communication in cellular networks. IEEE Communications Surveys Tutorials 16(4), 1801–1819 (2014)
2. Jaumard, B., Meyer, C., Thiongane, B.: ILP Formulations for the Routing and Wavelength Assignment Problem: Symmetric Systems, pp. 637–677. Springer US (2006)
3. Ozdaglar, A.E., Bertsekas, D.P.: Routing and wavelength assignment in optical networks. IEEE/ACM Trans. Netw. 11(2), 259–272 (April 2003)

THB1 : Integer Programming Theory

- The distance between optimal IP solutions and optimal solutions of MIP relaxations
  Joseph Paat, Robert Weismantel, Stefan Weltge.

- Breaking symmetries for the UCP by intersecting the full orbitope with an hypercube face
  Pascale Bendotti, Pierre Fouilhoux, Cecile Rottner.

- Supervised learning with parameter tuning in Branch-and-bound strategies
  Mohamed Mustapha Kabbaj, Abdellatif El Afia.

- Characterization and Approximation of Strong General Dual Feasible Functions
  Matthias Koeppe, Jiawei Wang.

The distance between optimal IP solutions and optimal solutions of MIP relaxations

Joseph Paat, Robert Weismantel, and Stefan Weltge

Institute for Operations Research, ETH Zürich

Let $A \in \mathbb{Z}^{m\times n}$, $b \in \mathbb{Z}^m$, and $c \in \mathbb{R}^n$. For $\ell \in \{0, \ldots, n\}$, we consider the mixed-integer linear program with ℓ variables

$$\max\{c^\intercal x : \ Ax \leq b, \ x_i \in \mathbb{Z} \ \forall i \in \{1, \ldots, \ell\}\}. \qquad (\ell\text{-MIP})$$

Observe that (0-MIP) is a linear program, while (n-MIP) is an integer linear program.

A common technique for solving (ℓ-MIP) is to first solve a relaxed program (k-MIP) with fewer integer variables, i.e. $k < \ell$, and then use a rounding scheme to find a nearby optimal solution for (ℓ-MIP). This leads to our main question.

Question 1. Let $k, \ell \in \{0, \ldots, n\}$. Does there exist a $\beta \in \mathbb{R}_+$ such that, given *any* optimal solution w of (k-MIP), there exists some optimal solution z of (ℓ-MIP) with $\|k - \ell\|_\infty \leq \beta$?

We assume that (ℓ-MIP) has an optimal solution for every $\ell \in \{0, \ldots, n\}$. Under this assumption, a *proximity bound* β answering Question 1 can quickly be found using the values A, b, m, and n. In [1], Blair and Jeroslow proved that, provided $k = 0$ or $\ell = 0$, there exists a proximity bound β dependent on A, n, m and the number of integer variables k and ℓ. In [2], Cook et al. improve upon this by providing an explicit bound that depends on only A and n. The result of Cook et al., which is stated below, parametrizes the matrix A using the following determinant-based value

$$\Delta := \max\{|\det(B)| : \ B \text{ is a square submatrix of } A\}.$$

Theorem 1 (Cook et al. [2]). *Let $k, \ell \in \{0, \ldots, n\}$ and assume that $k = 0$ or $\ell = 0$. Let w be an optimal solution to (k-MIP). Then there exists an optimal solution z to (ℓ-MIP) such that $\|w - z\|_\infty \leq n\Delta$.*

In order to complement the upper bound provided in Theorem 1, we provide the following example, which gives a lower bound on β in terms of Δ.

Example 1. Let $\delta \in \mathbb{Z}_+$ and consider the (0-MIP)

$$\max\left\{-x_2 : \begin{bmatrix} -\delta & 0 \\ \delta & -1 \end{bmatrix} x \leq \begin{bmatrix} -1 \\ 0 \end{bmatrix}\right\}$$

with solution $(1/\delta, 1)$. The solution to (2-MIP) is $(1, \delta)$, and for this example, $\Delta = \delta$. Thus, for $k = 0$ and $\ell = 2$, the proximity bound is $\beta = \Delta - 1$. □

Theorem 1 and Example 1 seem to indicate that the parameter Δ may be important in determining proximity bounds. With this in mind, our main result strengthens Theorem 1 by providing a proximity bound in terms of A (through Δ) and the number of integer variables k and ℓ.

Theorem 2. *Let $k, \ell \in \{0, \dots, n\}$. Let w be an optimal solution to (k-MIP). Then there exists an optimal solution z to (ℓ-MIP) such that $\|w - z\|_\infty \leq \max\{k, \ell\}\Delta$.*

Theorem 2 strengthens Theorem 1 in two ways. First, Theorem 2 no longer requires $k = 0$ or $\ell = 0$. While Theorem 1 can be applied twice (in addition with the triangle inequality) to derive proximity bounds when neither k nor ℓ is 0, Theorem 2 provides a tighter upper bound in this setting and a more direct comparison between (k-MIP) and (ℓ-MIP). Second, Theorem 2 shows that proximity does not depend on the ambient dimension n, but rather the number of integer variables k and ℓ.

In the proof of Theorem 2, we address the question of integer feasibility in zonotopes generated by (not necessarily integer) rational vectors. This result is stated as follows.

Lemma 1. *Let $d \in \mathbb{Z}_+$, $\alpha_1, \dots, \alpha_t \in \mathbb{R}_+$ and $u^1, \dots, u^t \in \mathbb{Z}^d$. If $\sum_{i=1}^t \alpha_i > d$, then*

$$\left\{ x \in \mathbb{R}^d : \ x = \sum_{i=1}^t \beta_i u^i, \ 0 \leq \beta \leq \alpha, \beta \neq 0 \right\} \cap \mathbb{Z}^d \ \neq \ \emptyset.$$

The zonotope defined in Lemma 1 always contains 0, and thus, is always integer feasible. However, when 0 is a vertex of the zonotope, then Lemma 1 provides a sufficient condition for the zonotope to contain a *non-zero* integer point. The proof of Lemma 1 uses results from group theory and additive combinatorics, and in particular, results on the Davenport Constant (see, for example, [3]). While the main result presented here is Theorem 2 on proximity of MIP solutions, we believe that this non-zero integer feasibility result along with the connection to group theory may be of independent interest in future research.

References

1. Blair, C., Jeroslow, R.: The value function of a mixed integer program: II. Discrete Mathematics 25, 7–19 (1979)
2. Cook, W., Gerards, A., Schrijver, A., Tardos, E.: Sensitivity theorems in integer linear programming. Mathematical Programming 34, 251–264 (1986)
3. Olson, J.E.: A combinatorial problem on finite abelian groups, i. Journal of Number Theory 1(8-10) (1969)

Breaking symmetries for the UCP by intersecting the full orbitope with an hypercube face

Pascale Bendotti[1,2], Pierre Fouilhoux[2], and Cécile Rottner[1,2]

[1] EDF R&D, 7 Boulevard Gaspard Monge, 91120 Palaiseau, France
pascale.bendotti@edf.fr
cecile.rottner@edf.fr
[2] Sorbonne Université, LIP6 CNRS UMR 7606, 4 Place Jussieu, 75005 Paris
pierre.fouilhoux@lip6.fr

We consider an Integer Linear Program (ILP) of the form

$$\min \left\{ c(x) \mid x \in \mathcal{X} \right\}, \text{ with } \mathcal{X} \subseteq \mathcal{P}(m,n) \text{ and } c : \mathcal{P}(m,n) \to \mathbb{R} \tag{1}$$

where $\mathcal{P}(m,n)$ is the set of $m \times n$ binary matrices. A symmetry is defined as a permutation π of the columns $\{1, ..., n\}$ such that for any solution matrix $x \in \mathcal{X}$, matrix $\pi(x)$ is also solution and has same cost, *i.e.*, $\pi(x) \in \mathcal{X}$ and $c(x) = c(\pi(x))$. The *symmetry group* $\mathcal{G}$ of ILP (1) is the set of all such permutations. Symmetry group $\mathcal{G}$ partitions the solution set $\mathcal{X}$ into *orbits*, *i.e.*, two matrices are in the same orbit if there exists a permutation in $\mathcal{G}$ sending one to the other.

Symmetries arising in ILP can impair the solution process, in particular when symmetric solutions lead to an excessively large branch and bound (B&B) search tree. Various techniques, so called *symmetry-breaking techniques*, are available to handle symmetries in ILP of the form (1). The general idea is, in each orbit, to pick one solution, defined as the *representative*, and then restrict the solution set to the set of all representatives.

A technique is said to be *full-symmetry-breaking* (resp. *partial-symmetry-breaking*) if the solution set is exactly (resp. partially) restricted to the representative set. Moreover, such a technique may introduce some specific branching rules that interfere with the B&B search. This can forbid exploiting a user-defined branching rule or, even, the default solver branching settings. A symmetry-breaking technique is said to be *flexible* if at any node of the B&B tree, the branching rule can be derived from any linear inequality on the variables.

We focus on a particular symmetry group, the *symmetric group* $\mathfrak{S}_n$, which is the group of all column permutations. The most common choice of representative is based on the lexicographical order. The convex hull of all $m \times n$ binary matrices with lexicographically non-increasing columns is called a *full orbitope* [1] and is denoted by $\mathcal{P}_0(m,n)$. No linear description in natural variables is known for the full orbitope.

Special cases of full orbitopes are *packing* and *partitioning orbitopes*, which are restrictions to matrices with at most (resp. exactly) one 1-entry in each row.

For a given face of the (m, n)-dimensional 0/1-cube, an algorithm is proposed in [2] to determine all variables whose values are 0 (resp. 1) in each matrix contained in the intersection of the hypercube face with the partitioning (or packing) orbitope. A symmetry-breaking algorithm, called *orbitopal fixing*, is derived in [2] in order to enumerate only the solutions included in the partitioning (resp. packing) orbitope during the B&B search. The idea is to consider, at each node a, the hypercube face defined by variables fixed to 0 and to 1 as a result of previous branching decisions, and then determine all variables which have constant value 0 (resp. 1) in the intersection of this face with the partitioning (or packing) orbitope. Orbitopal fixing is to fix these variables to 0 (resp. 1) at node a. It is worth noting that orbitopal fixing is flexible, full-symmetry-breaking and does not introduce any additional inequalities. These key features make orbitopal fixing for packing and partitioning orbitopes particularly efficient.

There are many problems whose symmetry group is the symmetric group acting on the columns, or on a subset of the columns, but whose solution space cannot be restricted to a partitioning or a packing orbitope. Examples range from line planning problems in public transports [3] to scheduling problems with a discrete time horizon, like the Unit Commitment Problem (UCP) or its variant, the Min-up/min-down UCP (MUCP) [4].

We propose an orbitopal fixing algorithm for the full orbitope, which handles the symmetries related to the symmetric group arising in the aforementioned problems. This algorithm, which is in linear time, finds all variables whose values are 0 (resp. 1) in each matrix contained in the intersection of any hypercube face with the full orbitope. This is a flexible full-symmetry-breaking method which is computationally efficient. Note that it does not increase the size of the LP solved at each node of the B&B tree. We present numerical experiments on MUCP instances featuring production units with identical characteristics. A comparison with state of the art symmetry-breaking techniques is presented in order to show the effectiveness of our approach.

References

1. Kaibel, V. and Pfetsch, M. *Packing and partitioning orbitopes.* Mathematical Programming, 2008.
2. Kaibel, V., Peinhardt, M. and Pfetsch, M. E *Orbitopal fixing.* IPCO, 2007.
3. R. Borndörfer, M. Grötschel and M. E. Pfetsch *A Column-Generation Approach to Line Planning in Public Transport.* Transportation Science, 2007.
4. P. Bendotti, P. Fouilhoux and C. Rottner *The Min-up/Min-down Unit Commitment Polytope.* Optimization Online `http://www.optimization-online.org/DB_HTML/2016/12/5750.html`, 2016.

Supervised learning with parameter tuning in Branch-and-bound strategies

KABBAJ Mohamed Mustapha[1] and EL AFIA Abdellatif[1]

[1] ENSIAS, Mohamed V University, Rabat, Morocco
mustapha.kabbaj@um5s.net.ma

In real life, MILP has countless applications in different fields like logistics, Finance and transportation... A very common solution technique is Branch-and-Bound. It continues to prove its relevance nowadays.

Branch-and-Bound algorithm that is an iterative algorithm, and at each iteration, we eventually get a feasible or optimal solution of an initial problem. Concretely, the algorithm constructs little by little a tree of nodes, where each node represents an modified version of the original problem. The construction of child nodes is conducted by a variable branching strategy. Another fundamental element in Branch-and-Bound algorithm is Node Selection Strategy that aims to choose among a nodes queue, the one that will speed up the search.

Recently, some works has been trying to identify an analytic approach that decide about strategies described above, given a set of problem features. Authors use likely machine learning techniques. The main remark is that few authors who deal with node selection strategy, and if so, they did not use machine learning framework.

Our contribution is oriented towards learning efficient branch-and-bound strategies. This is the result of a consistent methodology beginning with the collection of the data set, and ending with the test of the final hypothesis. More explicitly, we:

- Define features.
- Collect data set
- Pick the optimal learning model
- Learn the final hypothesis with the chosen model
- Test the final hypothesis

Our methodology allows firstly exploiting information from previous executions of Branch-and-Bound algorithm on other instances. Secondly, it created information channel between node selection strategy and variable branching strategy. And thirdly, it gave good results in term of solving time comparing to standard Branch-and-Bound algorithm. Moreover, it increases machine learning algorithm performance by using cross validation coupled with model selection.

Keywords: Node selection Strategy, Variable Branching Strategy, Branch and Bound; SVM, Cross validation, model selection

Characterization and Approximation of Strong General Dual Feasible Functions

Matthias Köppe[✉] and Jiawei Wang

Dept. of Mathematics, University of California, Davis, USA
mkoeppe@math.ucdavis.edu, jwewang@ucdavis.edu

Dual feasible functions (DFFs) have been used in several combinatorial minimization problems to generate lower bounds efficiently. DFFs are in the scope of superadditive duality theory, and superadditive and nondecreasing DFFs can provide valid inequalities for general integer linear programs. A function $\phi\colon [0,1] \to [0,1]$ is called a classical DFF, if for any finite list of real numbers $x_i \in [0,1]$, $i \in I$, it holds that, $\sum_{i\in I} x_i \leq 1 \Rightarrow \sum_{i\in I} \phi(x_i) \leq 1$. A function $\phi\colon \mathbb{R} \to \mathbb{R}$ is called a general DFF, if the same inequality holds for $x_i \in \mathbb{R}$.

1 Characterizations

A hierarchy on the set of DFFs can be defined to indicate the strength of the corresponding valid inequalities and lower bounds. The point-wise non-dominated DFFs are called *maximal*, and a maximal DFF is said to be *extreme* if it cannot be written as a convex combination of two distinct maximal DFFs. We give a full characterization of maximal general DFFs in terms of the "generalized symmetry condition".

Theorem 1. *A function $\phi\colon \mathbb{R} \to \mathbb{R}$ is a maximal general DFF if and only if the following conditions hold: (i) $\phi(0) = 0$. (ii) ϕ is superadditive. (iii) $\phi(x) \geq 0$ for all $x \in \mathbb{R}_+$. (iv) ϕ satisfies the generalized symmetry condition $\phi(x) = \inf_k\{\frac{1}{k}(1-\phi(1-kx)) : k \in \mathbb{Z}_{++}\}$ for all $x \in \mathbb{R}$.*

Parallel to the restricted minimal and strongly minimal functions in the Yıldız–Cornuéjols model [4], "restricted" and "strongly" maximal general DFFs can be defined by strengthening the notion of maximality. We also give the characterizations of restricted and strongly maximal general DFFs, which replace the "generalized symmetry condition" by a simpler condition.

Definition 1. *We say that a general DFF ϕ is* implied via scaling *by a general DFF ϕ_1, if $\beta\phi_1 \geq \phi$ for some $0 \leq \beta \leq 1$. We call a general DFF $\phi\colon \mathbb{R} \to \mathbb{R}$* restricted maximal *if ϕ is not implied via scaling by a distinct general DFF ϕ_1. We say that a general DFF ϕ is* implied *by a general DFF ϕ_1, if $\phi(x) \leq \beta\phi_1(x) + \alpha x$ for some $0 \leq \alpha, \beta \leq 1$ and $\alpha + \beta \leq 1$. We call a general DFF $\phi\colon \mathbb{R} \to \mathbb{R}$* strongly maximal *if ϕ is not implied by a distinct general DFF ϕ_1.*

Theorem 2. *A function $\phi\colon \mathbb{R} \to \mathbb{R}$ is a restricted maximal general DFF if and only if the following conditions hold: (i) $\phi(0) = 0$. (ii) ϕ is superadditive. (iii) $\phi(x) \geq 0$ for all $x \in \mathbb{R}_+$. (iv) $\phi(x) + \phi(1-x) = 1$ for all $x \in \mathbb{R}$.*

A function $\phi\colon \mathbb{R} \to \mathbb{R}$ is a strongly maximal general DFF if and only if ϕ is a restricted maximal general DFF and $\lim_{\epsilon\to 0^+} \frac{\phi(\epsilon)}{\epsilon} = 0$.

2 Relation to cut-generating functions

Cut-generating functions (CGFs) play an essential role in generating valid inequalities which cut off the fractional basic solution in a simplex-based cutting plane procedure. We study the relation between general DFFs and certain CGFs. Köppe–Wang [3] presented a conversion from minimal Gomory–Johnson CGFs to maximal DFFs. In the single-row Gomory–Johnson model, the basic variables are in $\mathbb{Z}$. Yıldız–Cornuéjols introduced a model generalizing the Gomory–Johnson setting, and considered the basic variables to be in any set $S \subset \mathbb{R}$. General DFFs generate valid inequalities for the Yıldız–Cornuéjols model with $S = (-\infty, 0]$. Jeroslow [2] studied the valid inequalities for a classic model which fits in the Yıldız–Cornuéjols setting where $S = \{0\}$. We introduce a conversion between general DFFs and a family of CGFs which generate valid inequalities for the model where $S = \{0\}$, which lifts valid inequalities generated by CGFs for $S = \{0\}$ to valid inequalities generated by general DFFs for the relaxation $S = (-\infty, 0]$.

3 Two-slope theorem and approximation theorem

Inspired by the famous Gomory–Johnson's 2-slope theorem, we prove a 2-slope theorem for general DFFs. We show that continuous extreme (2-slope) general DFFs are dense in the set of continuous restricted maximal general DFFs. The proof follows a parallel construction by Basu et al. [1], but the details are more complicated since general DFFs are not necessarily bounded and quasiperiodic.

Theorem 3 (Two-Slope Theorem). *Let ϕ be a continuous piecewise linear strongly maximal general DFF with only 2 slope values, then ϕ is extreme.*

Theorem 4 (Approximation Theorem). *Let ϕ be a continuous restricted maximal general DFF, then for any $\epsilon > 0$, there exists an extreme general DFF ϕ_{ext} such that $\|\phi - \phi_{\text{ext}}\|_\infty < \epsilon$.*

References

1. Basu, A., Hildebrand, R., Molinaro, M.: Minimal cut-generating functions are nearly extreme. In: Louveaux, Q., Skutella, M. (eds.) Integer Programming and Combinatorial Optimization: 18th International Conference, IPCO 2016, Liège, Belgium, June 1–3, 2016, Proceedings, pp. 202–213. Springer International Publishing, Cham (2016), http://dx.doi.org/10.1007/978-3-319-33461-5_17
2. Jeroslow, R.G.: Minimal inequalities. Mathematical Programming 17(1), 1–15 (1979)
3. Köppe, M., Wang, J.: Structure and interpretation of dual-feasible functions. Electronic Notes in Discrete Mathematics 62, 153–158 (2017), http://www.sciencedirect.com/science/article/pii/S1571065317302664, LAGOS '17 – IX Latin and American Algorithms, Graphs and Optimization
4. Yıldız, S., Cornuéjols, G.: Cut-generating functions for integer variables. Mathematics of Operations Research 41(4), 1381–1403 (2016), http://dx.doi.org/10.1287/moor.2016.0781

THB2 : Routing Problems

- Maximum concurrent flow with incomplete data
  Pierre-Olivier Bauguion, Claudia D'Ambrosio, Leo Liberti.

- Jointly Optimizing Replica Placement, Requests Distribution and Server Storage Capacity on Content Distribution Networks
  Raquel Gerhardt, Tiago Neves, Luis Rangel.

- The Exponent of Pheromone Level Adaptation for Ant Colony System Algorithm based Hidden Markov Model for TSP Problems
  Safae Bouzbita, Abdellatif El Afia, Rdouan Faizi.

Maximum concurrent flow with incomplete data

Pierre-Olivier Bauguion[1], Claudia D'Ambrosio[2], and Leo Liberti[2]

[1] IRT SystemX, 8 Avenue de la Vauve, 91120 Palaiseau, France
pierre-olivier.bauguion@irt-systemx.fr
[2] CNRS LIX, Ecole Polytechnique, 91128 Palaiseau, France
{dambrosio,liberti}@lix.polytechnique.fr

Abstract. The Maximum Concurrent Flow Problem (MCFP) is often used in the planning of transportation and communication networks. We discuss here the MCFP with incomplete data. We call this new problem the Incomplete Maximum Concurrent Flow Problem (IMCFP). The main objective of IMCFP is to complete the missing information assuming the known and unknown data form a MCFP and one of its optimal solutions. We propose a new solution technique to solve the IMCFP which is based on a linear programming formulation involving both primal and dual variables, which optimally decides values for the missing data so that they are compatible with a set of scenarios of different incomplete data sets. We prove the correctness of our formulation and benchmark it on many different instances.

Keywords: Maximum concurrent flow, multi-commodity flow problems, incomplete data, unknown data, uncertainty, inverse optimization, transportation systems.

1 Introduction

In real-world applications, the available data are often uncertain or incomplete, and their actual values may only be revealed at a time when the overall decision strategy has already been chosen. This is often the case in transportation systems where the parameters are time-dependent and event-sensitive. Statistical inference and data mining represent convenient ways to deal with this uncertainty. One of the best known inference models in transportation systems is the Four Step Model [8], which is an algorithm that iterates over time according to an equilibrium criterion. More recently, a lot of attention has been devoted to machine learning approaches, which generally performs better on large scale datasets. In this context, [10] proposes bayesian networks and [7] uses a deep learning approach to forecast flow in transportation systems.

Optimization methods can also be used to optimally fit experimental measurements. In [6], multi-commodity flow optimization is used to model a gas transportation network while retrieving missing data. The problem discussed in [6] consists in recomposing the flow on each arc, knowing only the global amount of incoming and outgoing flows for each node. The problem of finding a minimal adjustment of the cost function to ensure the optimality of a given solution generated a particular interest with [4] under the label of inverse optimization. For example [1, 11] apply this concept to multi-commodity flow problem (especially min cost flow problem). The survey [5] on this subject includes situations where the inverse problem seeks parameters other than objective function coefficients.

The Maximum Concurrent Flow Problem (MCFP) has been extensively studied over time [9, 3, 2], but in this paper we present a new approach for finding optimal maximum concurrent flows using incomplete data. Our method seeks optimal solutions and completes the partial input. This problem typically arises when we have insights about the global behavior of a system while data are partially unknown [6]. Symmetrically it can validate/invalidate a hypothetical behavior by comparing it with the observed data. This is particularly relevant in transportation when the routing strategy of passengers is known while data are incomplete. We call this problem Incomplete Maximum Concurrent Flow Problem (IMCFP).

References

1. R. K. Ahuja and J. B. Orlin. Inverse optimization, part I: Linear programming and general problem, 1998.
2. P. O. Bauguion, W Ben-Ameur, and E. Gourdin. Efficient algorithms for the maximum concurrent flow problem. *Networks*, 65(1):56–67, 2015.
3. D. Bienstock and O. Raskina. Asymptotic analysis of the flow deviation method for the maximum concurrent flow problem. *Mathematical Programming, Ser. B*, 91:479–492, 2002.
4. D. Burton and Ph. L. Toint. On an instance of the inverse shortest paths problem. *Mathematical Programming*, 53(1):45–61, 1992.
5. C. Heuberger. Inverse combinatorial optimization: A survey on problems, methods, and results. *Journal of Combinatorial Optimization*, 8(3):329–361, 2004.
6. J. N. Hooker. *Inferring network flows from incomplete information with application to natural gas flows.* Oak Ridge National Laboratory, 1980.
7. Y. Lv, Y. Duan, W. Kang, Z. Li, and F. Y. Wang. Traffic flow prediction with big data: A deep learning approach. *IEEE Transactions on Intelligent Transportation Systems*, 16(2):865–873, 2015.
8. M. L. Manheim. *Fundamentals of transportation systems analysis.* Cambridge, Mass.: MIT Press, 1979.
9. F. Shahrokhi and D. W. Matula. The maximum concurrent flow problem. *Journal of the ACM*, 37(2):318–334, 1990.
10. S. Shiliang, Z. Changshui, and Y. Guoqiang. A bayesian network approach to traffic flow forecasting. *IEEE Transactions on Intelligent Transportation Systems*, 7(1):124–132, 2006.
11. J. Zhang and Z. Liu. Calculating some inverse linear programming problems. *Journal of Computational and Applied Mathematics*, 72(2):261 – 273, 1996.

Jointly Optimizing Replica Placement, Requests Distribution and Server Storage Capacity on Content Distribution Networks

Raquel Gerhardt, Tiago Neves, and Luis Rangel

Universidade Federal Fluminense, Escola de Engenharia Industrial Metalúrgica de Volta Redonda, Volta Redonda - RJ, Brazil

Abstract. A Content Distribution Network includes dedicated servers creating an architecture that moves the content to servers that are closer to the user, reducing delays and traffic. In this structure several problems are studied, including the Problem of Allocation of Storage Capacity (SCAP) and the Replica Placement and Request Distribution Problem (RPRDP). This work analyzes these problems in an integrated way and proposes the creation of a new problem named Capacities, Replicas and Requests Distribution Problem (CRRDP), which enables the dynamic allocation of disk space on the servers and distribution of replicas and requests. As main contributions of this work are the creation of a new problem and a new formulation which associates variables and restrictions presents in mathematical formulations for this problems. The Mathematical formulation was analyzed and computational results shows that operational costs can be reduced and that it is possible to disable unused servers over the network.

1 Introduction

In the CDNs context, there are several optimization problems that have already been addressed in multiple ways [2], [1], [3], [4], [5]. This work proposes a new optimization problem, called the Capacities, Replicas and Requests Distribution Problem (CRRDP). This new problem involves the simultaneous optimization of servers disk capacities, replica positioning and the distribution of requests through CDN servers, solving two related optimization problems in a jointly way, the Storage Capacity Allocation Problem (SCAP) and the Replication Replica Placement and Request Distribution Problem (RPRDP). With data volume increasing and content popularity, *CloudCDN* [6] structures bring new insights that utilize network virtualization to facilitate its operations. Such insights are also used in the this paper and will be better explored latter.

2 Problem definition

The main haracteristics of the CRRDP are: 1) Requests are treated individually and can be handled by multiple servers simultaneously; 2) Storage capacity of servers can be changed according to demand fluctuation; 3) New requests and contents may come up and contents can be removed; 4) Network delays can change over the time horizon; 5) Bandwidth constraints are considered for clients and servers; 6) Clients' QoS requirements are fulfilled whenever possible; 7) The problem is Offline, meaning that all changes are known in advance.

3 Formulation Analysis Troughout Computational Tests

- CRRDP Dynamic
- Instances that do not Consider all the Servers as source
- CRRDP with Allocation Cost

4 Conclusions and Future Work

The model used was able to solve two optimization problems jointly (SCAP and RPRDP), obtaining success for most of the instances used. The model was able to prove that the dynamic allocation of servers' disk space was able to indeed reduce the operational cost of a CDN without violating the required quality standards.

References

1. Neves, T.: Redes de distribuição de conteúdos: Abordagens exatas e heurísticas. Technical report, UFF (2011)
2. Uderman, F., Neves, T., Albuquerque, C.: Optimizing server storage capacity on content distribution networks. In: Anais do Simpósio Brasileiro de Redes de Computadores - SBRC2011. (2011)
3. Tang, X., Xu, J.: On replica placement for qos-aware content distribution. In: Proc. of INFOCON2004. (2004) 806–815
4. Aioffi, W., Mateus, G., Almeida, J., Loureiro, A.: Dynamic content distribution for mobile enterprise networks. IEEE Journal on Selected Areas in Communications **23**(10) (2005)
5. Li, W.e.a.: Analysis and performance study for coordinated hierarchical cache placement strategies. 1 edn. Elsevier (2010)
6. Hu, H., Wen, Y.: Joint content replication and request routing for social video distribution over cloud cdn: A community clustering method. IEEE Transactions on Circuits and Systems for Video Technology **26**(7) (Julho 2016) 1320–1333
7. Neves, T., Drummond, L., Ochi, L., Albuquerque, C., Uchoa, E.: Solving replica placement and request distribution in content distribution networks. Electronic Notes in Discrete Mathematics **36** (August 2010) 89–96
8. Gerhardt, R., Neves, T., Albuquerque, C.: Análise de redundância em uma formulação matemática para o problema de posicionamento de réplicas e distribuição de requisições em redes de distribuição de conteúdos. SBPO - Simpósio Brasileiro de Pesquisa Operacional **8** (Agosto 2015)
9. Laboratório de Inteligência Computacional: LABIC. World Wide Web,http://labic.ic.uff.br/ (2005) 02/2009.
10. IBM ILOG CPLEX Optimization Studio (COS): ILOG S.A., CPLEX 10 user's manual (2006)

The Exponent of Pheromone Level Adaptation for Ant Colony System Algorithm based Hidden Markov Model for TSP Problems

Safae Bouzbita[1] Abdellatif El Afia[2] and Rdouan Faizi[3]

ENSIAS - Mohammed V University, Rabat, Morocco,
safae.bouzbita@gmail.com[1]
a.elafia@um5s.net.ma[2]
rdfaizi@gmail.com[3]

The TSP is one of the most complex combinatorial optimization problems [1–3]. Many heuristics and meta-heuristics have been proposed to find near optimal solution to it. The Ant Colony System (ACS) is one of the most interesting variants of Ant Colony Optimization (ACO) meta-heuristic for solving the TSP [4]. However, The performance of ACO algorithms depends strongly on the given values to parameters. Several strategies have been proposed in the literature for adapting parameters while solving a problem. For example, in [5–9] authors have chosen machine learning algorithms to adapt parameters of meta-heuristics at runtime.

In this paper, we propose a learning approach to control the convergence of ants through the dynamic adaptation of the exponent of pheromone level α parameter to two performance measures: Variance and Error while solving some TSP problems. The Ant Colony System (ACS) algorithm was chosen to be the basis of the proposed approach. A Hidden Markov Model (HMM) was built as a classifier method to avoid premature convergence. The implementation was tested on several Travelling Salesman Problem (TSP) instances with different number of cities. The proposed method was compared with the classical ACS and fuzzy logic and has shown encouraging results.

In the proposed method, the hidden states correspond to the state of parameter α, and the observation symbols are the concatenation between the performance measures, such as each measure represented by three symbols L, M, H. In fact, after building a complete solution by ants, the variance and the error calculated and converted into symbols according to some determined intervals. The symbols then combined to build an observation, then the observation sent to the Viterbi algorithm to determine which state is the most likely responsible for producing this observation. The sequence of observation is incremented after each iteration, so the number of elements of a sequence equal to the number of iterations, and the adaptation of the parameter α is done according to the last state in the found sequence. In addition to the Viterbi, we have used the well known Baum-Welch training method to adjust the HMM parameters $\lambda = (A, B, \pi)$ during the run time.

To test the efficiency of the proposed algorithm, we compared it with the standard ACS algorithm and Fuzzy Logic results from [10]. Also, the CPUtime was calculated to determine the convergence speed, and the Wilcoxon statistical test

was used to compare between the algorithms.
We can conclude that the proposed algorithm ACSHMM gives better results in the convergence speed and the solution accuracy compared with both the standard ACS and the proposed one in the literature by fuzzy logic when applied to some TSP instances.

Keywords: exploration and exploitation, parameter adapting, Hidden Markov Model, Travelling Salesman Problems

References

1. Erol, Ayşe Hande, Merve Er, and Serol Bulkan. "Optimizing the ant colony optimization algorithm using neural network for the travelling salesman problem." Actas de la Conferencia Internacional de. 2012.
2. Helsgaun, K., and Ngassa, J. L. ACO and TSP.(2007).
3. Matai, R., Singh, S., and Mittal, M. L. Traveling salesman problem: an overview of applications, formulations, and solution approaches. In Traveling salesman problem, theory and applications. InTech.(2010).
4. Filip, Exnar, and Machac Otakar. "The travelling salesman problem and its application in logistic practice." WSEAS Transactions on Business and Economics 8.4 (2011): 163-173.
5. Li Y., LiW.: Adaptive ant colony optimization algorithm based on information entropy: Foundation and application. Fundamenta Informaticae 77(3):229–242 (2007)
6. Martens, D., De Backer, M., Haesen, R., Vanthienen, J., Snoeck, M., and Baesens, B. Classification with ant colony optimization. IEEE Transactions on Evolutionary Computation, 11(5), 651-665 (2007)
7. Khichane M., Albert P., Solnon C.: An ACO-based reactive framework for ant colony optimization: First experiments on constraint satisfaction problems.In: Stutzle T. (ed.) Learning and Intelligent Optimization, Third International Conference, LION 3, Lecture Notes in Computer Science, vol. 5851, Springer, Heidelberg, Germany, pp. 119–133 (2009)
8. Ling W., Luo H.: An adaptive parameter control strategy for ant colony optimization.In: CIS'07: Proceedings of the 2007 International Conference on Computational Intelligence and Security, IEEE Computer Society, Washington,DC, pp. 142–146 (2007)
9. Melo L., Pereira F., Costa E.: MC-ANT: A multi-colony ant algorithm. In:Artificial Evolution - 9th International Conference, Evolution Artificielle, EA 2009, Lecture Notes in Computer Science, vol. 5975, Springer, Heidelberg, Germany, pp. 25–36 (2009)
10. Neyoy, H., Castillo, O., and Soria, J. Fuzzy logic for dynamic parameter tuning in ACO and its application in optimal fuzzy logic controller design. In Fuzzy Logic Augmentation of Nature-Inspired Optimization Metaheuristics (pp. 3-28). Springer International Publishing. (2015)

THB3 : Graph Structures

- Robust Matching Augmentation
  Moritz Muehlenthaler, Viktor Bindewald, Felix Hommelsheim, Oliver Schaudt.

- Characterising chordal contact B_0-VPG graphs
  Flavia Bonomo, María Pía Mazzoleni, Mariano Rean, Bernard Ries.

- Cluster Editing with Vertex Splitting
  Faisal Abu-Khzam, Judith Egan, Serge Gaspers, Alexis Shaw, Peter Shaw.

- A PPA Episode of Exchange Graphs
  Kathie Cameron, Jack Edmonds.

Robust Matching Augmentation

Viktor Bindewald[1], Felix Hommelsheim[1], Moritz Mühlenthaler[1], and Oliver Schaudt[2]

[1] Department of Mathematics, TU Dortmund University, Germany
{viktor.bindewald,moritz.muehlenthaler,felix.hommelsheim}@math.tu-dortmund.de
[2] RWTH Aachen University, Germany
oliver.schaudt@mathc.rwth-aachen.de

We call a perfect matching in a bipartite graph *k-robust*, if the graph admits a perfect matching after removing at most k of the matching edges. We consider two optimization problems related to k-robust matchings. The first one asks for the maximal number k, such that a given matching in a bipartite graph is k-robust. We show that this problem is $\mathsf{W}[1]$-hard when parameterized by k. The above notion of k-robustness corresponds to *weakly k-robust ∞-recoverable* as studied by Dourado et al. in [1], however, no complexity results for k-robustness were known. The second problem asks for the minimal number ℓ of edges we have to add to the graph, in order to make a given matching k-robust. Clearly, this problem is also hard. It turns out however, that even the problem 1-Robust Matching Augmentation, i.e., the task of making a matching 1-robust with the minimal number of additional edges, is in general as hard as Set Cover. To the best of our knowledge, augmentation of robustness has not been considered explicitly before, although it can be interpreted as a special case of *bulk-robustness* [2, 3].

The task of augmenting a minimal number of edges such that a graph has a certain property has been widely studied, in particular in the context of *connectivity augmentation* [4]. Given a digraph $D = (V, A)$, the problem Strong Connectivity Augmentation asks for a minimum-cardinality arc-set A' such that the digraph $(V, A \cup A')$ has precisely one strong component. Interestingly, 1-robustness is a generalization of strong connectivity in the following sense. Given a simple bipartite graph $G = (V, E)$ and a perfect matching M, we may construct a digraph D by adding an arc uw whenever there is an edge $uv \in M$ and an edge $vw \in E \setminus M$. Now, a matching is 1-robust, if and only if each strong component of D contains at least two vertices. So we have that if D has at least two vertices, then strong connectivity of D implies 1-robustness of M. Note also, that from any digraph D we may construct a corresponding undirected graph G and a perfect matching M of G.

Using the insights described above, we show that the Eswaran-Tarjan algorithm, a polynomial-time algorithm for Strong Connectivity Augmentation from [4], is also useful for 1-robustness augmentation. For instance, it directly leads to a polynomial-time algorithm for 1-Robust Matching Augmentation on trees. With some more work we obtain polynomial-time algorithms for the same problem on graphs with bounded treewidth and convex graphs that again rely on the Eswaran-Tarjan algorithm. On the negative side, we show that the *weighted* version of 1-Robust Matching Augmentation is NP-hard on trees, even if the weight of each edge is either one or two.

References

1. Dourado, M.C., Meierling, D., Penso, L.D., Rautenbach, D., Protti, F., de Almeida, A.R.: Robust recoverable perfect matchings. Networks **66**(3) (2015) 210–213
2. Adjiashvili, D., Stiller, S., Zenklusen, R.: Bulk-robust combinatorial optimization. Mathematical Programming **149**(1-2) (2015) 361–390
3. Adjiashvili, D., Bindewald, V., Michaels, D.: Robust assignments via ear decompositions and randomized rounding. In: ICALP 2016. Volume 55 of Leibniz International Proceedings in Informatics (LIPIcs)., Dagstuhl, Germany (2016) 71:1–71:14
4. Eswaran, K.P., Tarjan, R.E.: Augmentation problems. SIAM Journal on Computing **5**(4) (1976) 653–665

Characterising chordal contact B_0-VPG graphs

F. Bonomo[1], M. P. Mazzoleni[2], M. L. Rean[1], and B. Ries[3]

[1] Universidad de Buenos Aires. Facultad de Ciencias Exactas y Naturales. Departamento de Computación. Buenos Aires, Argentina. / CONICET-Universidad de Buenos Aires. Instituto de Investigación en Ciencias de la Computación (ICC). Buenos Aires, Argentina, fbonomo@dc.uba.ar,marianorean@gmail.com

[2] Univesidad Nacional de La Plata. Facultad de Ciencias Exactas. Departamento de Matemática. La Plata, Argentina. / CONICET, pia@mate.unlp.edu.ar

[3] University of Fribourg, Fribourg, Switzerland, bernard.ries@unifr.ch

Golumbic et al. introduced in [2] the concept of *vertex intersection graphs of paths on a grid* (referred to as *VPG graphs*). An undirected graph $G = (V, E)$ is called a VPG graph if one can associate a path in a rectangular grid with each vertex such that two vertices are adjacent if and only if the corresponding paths intersect on at least one grid-point. It is not difficult to see that VPG graphs are equivalent to the well known class of string graphs, i.e. intersection graphs of curves in the plane (see [2]).

A particular attention was paid to the case where the paths have a limited number of *bends* (a bend is a 90 degrees turn of a path at a grid-point). An undirected graph G is then called a *B_k-VPG graph*, for some integer $k \geq 0$, if one can associate a path with at most *k bends* on a rectangular grid with each vertex such that two vertices are adjacent if and only if the corresponding paths intersect on at least one grid-point. Since their introduction in 2012, B_k-VPG graphs, $k \geq 0$, have been studied by many researchers and the community of people working on these graph classes is still growing (see [1–6, 8, 9]).

These classes are shown to have many connections to other, more traditional, graphs classes such as interval graphs (which are clearly B_0-VPG graphs), planar graphs (recently shown to be B_1-VPG graphs (see [9])), string graphs (as mentioned above equivalent to VPG graphs), circle graphs (shown to be B_1-VPG graphs (see [2])) and grid intersection graphs (GIG) (equivalent to bipartite B_0-VPG graphs (see [2])). Unfortunately, due to these connections, many natural problems are hard for B_k-VPG graphs. For instance, colouring is NP-hard even for B_0-VPG graphs and recognition is NP-hard for both VPG and B_0-VPG graphs [2]. However, there exists a polynomial-time algorithm for deciding whether a given chordal graph is B_0-VPG (see [3]).

A related notion to intersection graphs are *contact graphs*. Such graphs can be seen as a special type of intersection graphs of geometrical objects in which objects are not allowed to cross but only to touch each other. In the context of VPG graphs, we obtain the following definition. A graph G is called a *contact VPG graph* if the vertices can be represented by interiorly disjoint paths on a grid and two vertices are adjacent if and only if the corresponding paths touch. If we limit again the number of bends per path, we obtain *contact B_k-VPG graphs*. These graphs have also been considered in the literature (see [4, 7, 10]). It is known that every planar bipartite graph is a contact B_0-VPG graph [7], and that every K_3-free planar graph is a contact B_1-VPG graph [4]. The authors in [10] consider the case in which whenever two paths touch on a grid point, this grid point has to be the endpoint of one of the paths and belong to the interior of the other path. In this case, the considered graphs must all be planar.

In this paper, we consider *contact B_0-VPG graphs*, ie. intersection graphs in which each vertex is corresponding to a horizontal or vertical path on a grid and the corresponding paths do not cross each other and do not share an edge of the grid. We present a minimal forbidden induced subgraph characterisation of contact B_0-VPG graphs restricted to chordal graphs. This characterisation allows us to derive a polynomial time recognition algorithm for the class of chordal contact B_0-VPG graphs. The algorithm takes a chordal graph as input and returns YES if the graph is contact B_0-VPG and, if not, it returns NO as well as a forbidden induced subgraph. Recall that chordal B_0-VPG graphs can also be recognised in polynomial time (see [3]), even though no structural characterisation of them is known so far.

Our results can be considered as a first step to better understand contact B_0-VPG graphs and their structure. In order to know more about how contact B_0-VPG graphs look like, the study of contact B_0-VPG graphs within other graph classes is needed. It would be interesting to investigate contact B_0-VPG graph from an algorithmic point of view and analyse for instance the complexity of the colouring problem or the stable set problem in that graph class.

References

1. Alcón L., F. Bonomo, and M. P. Mazzoleni, *Vertex intersection graphs of paths on a grid: characterization within block graphs*, Graphs and Combinatorics 33(4) (2017), 653–664.
2. Asinowski, A., E. Cohen, M. Golumbic, V. Limouzy, M. Lipshteyn and M. Stern, *Vertex intersection graphs of paths on a grid*, Journal of Graph Algorithms and Applications 16 (2012), 129–150.
3. Chaplick S., E. Cohen, and J. Stacho, *Recognizing some subclasses of vertex intersection graphs of 0-bend paths in a grid*, Lecture Notes in Computer Science 6986 (2011), 319–330.
4. Chaplick, S., and T. Ueckerdt, *Planar graphs as VPG-graphs*, Journal of Graph Algorithms and Applications 17(4) (2013), 475–494.
5. Cohen, E., M.C. Golumbic, and B. Ries, *Characterizations of cographs as intersection graphs of paths on a grid*, Discrete Applied Mathematics 178 (2014), 46–57.
6. Cohen, E., M.C. Golumbic, W.T. Trotter, and R. Wang, *Posets and VPG Graphs*, Order 33(1) (2016), 39–49.
7. de Fraysseix H., P. Ossona de Mendez, and J. Pach, *Representation of planar graphs by segments*, Intuitive Geometry 63 (1991), 109–117.
8. Golumbic, M.C., and B. Ries, *On the intersection graphs of orthogonal line segments in the plane: Characterizations of some subclasses of chordal graphs*, Graphs and Combinatorics 29(3) (2013), 499–517.
9. Gonçalves, D., L. Insenmann, and C. Pennarum, *Planar Graphs as L-intersection or L-contact graphs*, arXiv:1707.08833v2 (2017).
10. Nieke, A., and S. Felsner, *Vertex Contact Representations of Paths on a Grid*, Journal of Graph Algorithms and Applications 19(3) (2015), 817–849.

Cluster Editing with Vertex Splitting (Short Paper)

Faisal N. Abu-Khzam[1,2], Judith Egan[1], Serge Gaspers[3,4], Alexis Shaw[5], and Peter Shaw[6]

[1] Charles Darwin University, Darwin, Australia
[2] Lebanese American University, Beirut, Lebanon
[3] The University of New South Wales, Sydney, Australia
[4] Data61, CSIRO, Sydney, Australia
[5] The University of Technology Sydney, Sydney, Australia
[6] Massey University, Manawatu, New Zealand

Abstract. We introduce a new variant of the Cluster Editing problem whereby a vertex can be divided into two or more vertices to allow a single vertex to belong to multiple cliques or clusters. This problem has applications in finding correlation clusters in discrete data, including graphs obtained from clinical data. We initiate the study of this new problem and show that it has a quadratic-order kernel.

1 Introduction

Given a graph G and a non-negative integer k, the Cluster Editing problem asks whether G can be turned into a disjoint union of cliques by a sequence of at most k edge-editing operations. The problem is known to be $\mathcal{NP}$-Complete [4], but it becomes fixed-parameter tractable when parameterized by the number of *edge edit operations* [1, 2].

In general, clustering results in a partition of the input graph, thus forcing each data element to be in exactly one cluster. This can be a limitation when a data element plays a role in multiple clusters. In fact, the existence of *hubs* can effectively hide clique-like structures and greatly increase the computational time required to obtain optimum correlation clustering solutions [6, 7].

2 Cluster editing with vertex splitting

We define that an *inclusive vertex split* of a vertex v replaces v with vertices v_1 and v_2 with edges such that $N(v_1) \cup N(v_2) = N(v)$.

The Cluster Editing with Vertex Splitting problem allows an edit operation to be an inclusive vertex split. Formally, given a graph $G = (V, E)$ and an integer k, can a cluster graph G' be obtained from G by a *k-edit-sequence*, $S = e_1 \dots e_k$, where each e_i, $i \in \{1, \dots k\}$, either (i) adds an edge to E; (ii) deletes an edge from E; or (iii) is an inclusive vertex split of some $v \in V$?

We define the *splitting-tree* for vertex v, denoted $T(v)$, as follows: (i) v is the root of $T(v)$; (ii) if an operation e_i for some $i \in \{1, \dots k\}$ inclusively splits a vertex u in $T(v)$, then the two vertices resulting from e_i become children of u.

Our analysis of edit-sequences proves the following:

Theorem 1. *There is an* $O((|V'| - |V|)\Delta(G) + |V| + |E| + |V'| + |E'|)$ *time procedure to determine an optimal edit-sequence* S *which transforms a graph* $G = (V, E)$ *to a given cluster graph* $G' = (V', E')$ *and a vertex relation* $f : V \to 2^{V'}$ *giving the leaves of the splitting-tree* $T(v)$ *for each vertex* $v \in V$.

3 A $4k(k+1)$ vertex kernel

A *critical clique* of a graph $G = (V, E)$ is a maximal induced subgraph C of G such that (i) C is a complete graph, and (ii) there is some subset $U \subseteq V$ such that for every $v \in V(C)$, $N[v] = U$. We prove an analog of the Critical Clique Theorem given by Lin et al. [5]. Using an approach similar to that given in [3] we prove the following.

Theorem 2. *There exists a polynomial-time reduction procedure that takes an arbitrary instance of the* Cluster Editing with Vertex Splitting *problem and produces an equivalent instance whose order is bounded above by* $4k(k+1)$. *In other words, the problem admits a quadratic-order kernel.*

Acknowledgments

Serge Gaspers is the recipient of an Australian Research Council (ARC) Future Fellowship (FT140100048) and acknowledges support under the ARC's Discovery Projects funding scheme (DP150101134). Alexis Shaw is the recipient of an Australian Government Research Training Program Scholarship.

References

1. Cai, L.: Fixed-parameter tractability of graph modification problems for hereditary properties. Information Processing Letters 58(4), 171–176 (May 1996)
2. Gramm, J., Guo, J., Hüffner, F., Niedermeier, R.: Graph-modeled data clustering: Exact algorithms for clique generation. Theory Comput. Syst. 38(4), 373–392 (2005)
3. Guo, J.: A more effective linear kernelization for cluster editing. Theoretical Computer Science 410(8-10), 718 – 726 (2009)
4. Kivnek, M., Morvek, J.: NP-hard problems in hierarchical-tree clustering. Acta Informatica 23 (3), 311–323 (1986)
5. Lin, G.H., Kearney, P.E., Jiang, T.: Phylogenetic k-root and Steiner k-root. In: Algorithms and Computation, pp. 539–551. Springer (2000)
6. Radovanović, M., Nanopoulos, A., Ivanović, M.: Hubs in space: Popular nearest neighbors in high-dimensional data. Journal of Machine Learning Research 11(Sep), 2487–2531 (2010)
7. Tomašev, N., Radovanović, M., Mladenić, D., Ivanović, M.: The role of hubness in clustering high-dimensional data. In: Pacific-Asia Conference on Knowledge Discovery and Data Mining. pp. 183–195. Springer (2011)

A PPA Episode of Exchange Graphs

Kathie Cameron[1] and Jack Edmonds[2]

[1] Department of Mathematics, Wilfrid Laurier University, Waterloo, Canada
kcameron@wlu.ca *
[2] jack.n2m2m6@gmail.com

A *good parity theorem* asserts that, for any instance of the theorem, there exists an even number of easily recognizable (i.e., NP) *desired* structures. Of great computational-complexity interest is the *like-item search problem*: given any instance of a good parity theorem and a desired structure P, find a desired structure different from P.

A good parity theorem is called a *PPA parity theorem* if it can be proved by showing that the desired structures are the odd-degree vertices in some large *exchange graph*, X. Exchange graph proofs are important because Papadimitriou introduced the question of determining the complexity of a like-item search problem which has a polynomial-time algorithm for determining the neighbors of any vertex of X. A natural algorithm for like-item search is to walk in X from an odd-degree vertex to another odd-degree vertex.

We present here a new PPA parity theorem which generalizes attractive known PPA parity theorems. Because the theorem is about the vertices and edges of a given bipartite graph G, we try to avoid some of the inevitable confusion by speaking of vertices and edges of X and nodes and lines of G.

The main interest of our presentation is probably not so much the parity result itself but rather its more intricate than usual exchange-graph technique. Hopefully there are many more PPA parity theorems to be discovered with even more intricate exchange graphs.

Let G be a finite bipartite graph with boy nodes and girl nodes such that each line (edge) of G joins a boy to a girl, and each boy is of even degree in G. Let T^* be a given tree in G which contains all the girls, each girl is met by an odd number of lines of G not in T^*, and each boy of T^* is met by exactly two lines of T^*. A tree T in G is said to be *girlish* (or, more precisely, T^*-*girlish*) when each girl has the same degree in T as in T^* and each boy of T has degree 2 in T. (Note that T and T^* are not necessarily spanning trees of G. There may be different boy nodes in T and T^*.)

Theorem 1. *G has an even number of girlish trees.*

We enable a proof of this theorem by describing an exchange graph X where the odd-degree vertices of X can be shown to be precisely the girlish trees. Hence this good parity theorem is a PPA parity theorem.

It is natural to think of X in terms of an algorithm which enters an edge from one of its vertices and then leaves the edge at its other vertex. Entering

* Research supported by the Natural Sciences and Engineering Research Council of Canada (grant RGPIN-2016-06517).

the edge from the vertex is simply the reverse of leaving the edge at that vertex, and we describe the latter. We do the reverse to prove the parity of the degrees of the vertices of X.

Specify any girl node w of G. The exchange graph X depends on the choice of w. A *skew tree* T differs from a girlish tree only in that node w has degree one higher than its degree in T^* and some other girl node has degree one lower than its degree in T^*. The vertices of the exchange graph X are the girlish trees and the skew trees. The edges of X are of two types: pink and blue. Each edge e of X is a connected subgraph of G with exactly one cycle, say C. Subgraph e has degrees the same as T^* except at two nodes, w and some other node u, which have degree one higher in e than in T^*. The difference between a pink and blue edge of X is that in a pink edge e the node u is a girl and in the blue edge e the node u is a boy (in fact a boy of degree 3 in e). A blue e meets the two vertices of X which are obtained from e by deleting one or the other of the two edges of C which meet boy node u. A pink e (where u is a girl) meets the two vertices of X which are obtained from e by deleting one or the other of the two boy nodes which are adjacent to u in C (deleting a boy node b of C of course means also deleting from e the two edges of C which meet b).

It is a matter of careful book-keeping of all the edges which can be entered from a vertex of X to confirm the Exchange Graph Lemma: The vertices of X corresponding to girlish trees are its odd-degree vertices and the vertices of X corresponding to skew trees are its even-degree vertices. Hence the parity theorem.

FRA1 : Online Algorithms

- Online Firefighting on Trees
  Pierre Coupechoux, Marc Demange, David Ellison, Bertrand Jouve.

- Online correlated orienteering on continuous surfaces
  João Pedro Pedroso.

- Preemptively Guessing the Center
  Christian Konrad, Tigran Tonoyan.

Online Firefighting on Trees

Pierre Coupechoux[1], Marc Demange[2], David Ellison[2⋆], and Bertrand Jouve[3⋆⋆]

[1] LAAS-CNRS, Université de Toulouse, CNRS, Toulouse, France,
[2] RMIT University, School of Science, Melbourne, Australia
[3] FRAMESPA and IMT, CNRS, Université de Toulouse, Toulouse, France

1 Introduction

Introduced by B. Hartnell in 1995 [1], the firefighting problem started as a very simple model for fire spread and containment problems for wildfires. It can also represent any kind of threat able to spread sequentially in a network (diseases, viruses, rumours, flood . . .).

The Firefighter problem is a deterministic discrete-time one-player game defined on a graph. In the beginning, a fire breaks out on a vertex and at each step, if not blocked, the fire spreads to all adjacent vertices. In order to contain the fire, the player is given a number f_i of firefighters at each turn i and can use them to protect vertices which are neither burning nor already protected. The game terminates when the fire cannot spread any further. In the case of finite graphs the aim is to save as many vertices as possible, while in the infinite case, the player wins if the game finishes, which means that the fire is contained.

In a natural variant of the problem, the amount of firefighters available at each turn is any non-negative number and the amount allocated to vertices lies between 0 and 1. A vertex with a protection less than 1 is *partially protected* and its unprotected part can burn partially and transmit only its fraction of fire to the adjacent vertices. Thus, the f_i may take any non-negative value. This defines a variant game called Fractional Firefighter which was introduced in [2].

In this paper, we introduce an online version of both firefighter problems and consider first results on trees. In our model, the graph is known and the sequence of available firefighters is revealed online. We then refer to the usual case where $(f_i)_{i\geq 1}$ is known in advance as *offline*. To our knowledge, this is the first attempt at analysing online firefighter problems. Even though our motivation is mainly theoretical, this paradigm is particularly natural in emergency management where one has to make quick decisions despite lack of information.

An instance of the Fractional Firefighter on trees is defined by a triple $(T, r, (f_i))$, where $T = (V(T), E(T))$ is a tree, $r \in V(T)$ is the root where the fire breaks out and $(f_i)_{i\geq 1}$ is the non-negative *firefighter sequence*. Turn $i = 0$ is the initial state where r is burning and all other vertices are unprotected, and $i \geq 1$ corresponds to the different rounds of the game. At each turn $i \geq 1$ and

⋆ corresponding author.
⋆⋆ We acknowledge the support of GEO-SAFE, H2020-MSCA-RISE-2015 project # 691161.

for every vertex v, the player decides which amount $p(v)$ of protection to add to v. Throughout the game, for every vertex v, the part of v which is burning is denoted by $b(v)$. Let us note that if T is finite, the game will end in at most $h(T)$ turns .

2 Competitive analysis of a Greedy algorithm

On a tree, we define the *weight* of a vertex as the number of its descendants. A greedy algorithm, defined on trees for both firefighter problems and denoted by Gr, maximises at each turn the weight of the vertices protected.

It was shown by Hartnell and Li that the greedy algorithm on trees gives a $\frac{1}{2}$-approximation of the restriction of FIREFIGHTER when a single firefighter is available at each turn. They claim that this approximation ratio remains valid for a fixed number $D \in \mathbb{N}$ of firefighters at each turn. We extend this result to any firefighter sequence $(f_i)_{i\geq 1}$, integral or not. Since Gr is an online algorithm, the performance can also be seen as a competitive ratio for the online version.

Theorem 1. *The greedy algorithm Gr is $\frac{1}{2}$-competitive for both online* FIREFIGHTER *and* FRACTIONAL FIREFIGHTER *on finite trees.*

Corollary 1. *In* FRACTIONAL FIREFIGHTER*, the amount of vertices saved is at most twice the maximum number of vertices saved in* FIREFIGHTER*.*

3 Firefighting on Trees with Linear Growth

In this section, we consider an infinite tree T. The i-th level of T, denoted by T_i, is the set of vertices at distance i from the root. a We say that a rooted tree (T, r) has *linear growth* if the number of vertices per level increases linearly, i.e. $|T_i| = \mathcal{O}(i)$. Note that the linear growth property of T remains if we choose a different root r'. Indeed, if d is the distance between r and r', the set of vertices at distance i from r' is included in $\bigcup_{j=i-d}^{i+d} T_j$, the cardinality of which is a $\mathcal{O}(i)$.

Given two firefighter sequences (f_i) and $(f_i^{'})$, we say that (f_i) is *stronger* than $(f_i^{'})$ if for all k, $\sum_{i=1}^{k} f_i \geq \sum_{i=1}^{k} f_i^{'}$.

Lemma 1. *If the fire can be contained in an instance $(G, r, (f_i^{'}))$ and if (f_i) is stronger than $(f_i^{'})$, then the fire can also be contained in $(G, r, (f_i))$ by an online algorithm that knows (f_i') in advance.*

Theorem 2. *There is an online algorithm for instances $(T, r, (f_i))$ of* FRACTIONAL FIREFIGHTER *where T has linear growth, such that if (f_i) is stronger than some non-zero periodic sequence, the fire will be contained.*

References

1. Hartnell, B., Li, Q.: Firefighting on trees: How bad is the greedy algorithm? Congressus Numerantium (2000) 187–192
2. Fogarty, P.: Catching the Fire on Grids. University of Vermont (2003)

Online correlated orienteering on continuous surfaces

João Pedro PEDROSO

INESC TEC and DCC - Faculdade de Ciências, Universidade do Porto, Portugal
jpp@fc.up.pt

1 Introduction

This work describes a problem with origins in sea exploration, though similar problems arise in other contexts. The identification of the contents of the seafloor is important in view of a possible exploitation of some of these resources. The aim of this problem is to schedule the journey of a ship for collecting information about the composition of the seafloor. We consider a bounded surface $S \subset \mathbb{R}^2$, and, for the sake of simplicity, we consider that the actual resource level at any point $(x, y) \in S$ can be conveyed by a real number $v(x, y)$. This *true value* is initially unknown, except for a limited number N of points in $(\bar{x}_i, \bar{y}_i) \in S$ for which there is previous empirical information $D = \{(\bar{z}_i, \bar{x}_i, \bar{y}_i)\}_{i=1}^{N}$.

Optimal expedition planning involves three subproblems, each corresponding to a different phase in the process: assessment, planning and estimation [1].

Assessment consists of estimating the amount of information that would be conveyed by probing the surface at each point $(x, y) \in S$. This is done by means of an indicator function $h : S \to \mathbb{R}$, given a set of N points for which the true contents $v(x, y)$ were known. Previous work assumed that actual information obtained by probing is not usable at the time of planning; here, we assume that after committing to probing at a certain place, the information obtained can immediately be used to change the course of the following decisions (in particular, set D used for building the indicator function is dynamically expanded).

Planning, the next phase in the solution process, consists of deciding on the position of points to probe until the end of the expedition; the point to probe next is the only one to which we commit. The objective is to maximize the overall informational reward obtained, taking into account that the total duration of the trip is limited to a known bound. Hence, online planning involves using the N previously available points together with the points newly probed in this trip, in order to decide the location of the next point to probe — though an estimation of the whole remaining trip is necessary for correctly taking this decision.

The third subproblem is estimation, which is related to the final aim of the problem: an estimation $w(x, y)$ of the resource level available at any point on the surface S, based on all the information available at the end of the trip.

Acknowledgments. This work was partially supported by project "Coral - Sustainable Ocean Exploitation: Tools and Sensors/NORTE-01- 0145-FEDER-000036", financed by the North Portugal Regional Operational Programme (NORTE 2020), under the PORTUGAL 2020 Partnership Agreement, and through the European Regional Development Fund (ERDF).

2 Tackling the problem

Assessment and estimation. Assessment evaluates how much some arbitrary point, if probed, is expected to improve the quality of solution obtained at the final estimation phase. The final estimation is a regression problem, assigning a real value to any point in S based on the discrete set of points for which such values are known. We propose to use Gaussian processes [2] for assessment and estimation. The measure of the "atractiveness" for probing at a given point is assumed to be given by the standard deviation of a Gaussian process set up with the D available data points. In other words, the assessment model attempts to describe the variance of the conditional distribution $p(z|(x, y))$ based on a set of empirical observations of z on input (x, y), conveyed by the set of triplets $D = \{(\bar{z}_i, \bar{x}_i, \bar{y}_i)\}_{i=1}^m$ observed so far, where m is the current number of samples (including N previous). Similarly, we use the posteriors inferred thought the Gaussian process model with the enlarged data set available at the end of the expedition as a regression for the resource level at any point in S.

Planning. Planning concerns the selection of the next point in S for probing, in a trip whose maximum duration is known beforehand, so as to allow a subsequent estimation as accurate as possible. We are thus in the presence of an orienteering problem [3]. A standard orienteering problem consists of the following: given a graph with edge lengths and a prize that may be collected at each vertex, determine a path of length at most T, starting and ending at given vertices, that maximizes the total prize value for the vertices visited. The problem here is slightly different. Firstly, at any moment we are only committed to the next point to be visited; the following points are planned, but may be changed after the next probing, if it reveals unexpected information, changing the shape of the Gaussian process's regression. Secondly, the graph in our case may consist of any discrete subset of points $V \subset S$, as long as the duration of the expedition does not exceed T. We must take into account the time spend in probing at each vertex for determining a trip's duration. Finally, the correlation between the "prizes" obtained in visited vertices must be taken into account: after probing at a given location, probing other locations in this neighborhood is expected to provide less information than probing at distant points, other factors being equal.

References

1. Kramer, A.V., Zhang, K., Pedroso, J.P.: Orienteering on a continuous surface: A problem in sea exploration. International Network Optimization Conference (February 2017) Lisbon, Portugal.
2. Rasmussen, C.E., Williams, C.K.I.: Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning). The MIT Press (2005)
3. Golden, B.L., Levy, L., Vohra, R.: The orienteering problem. Naval research logistics **34**(3) (1987) 307–318

Preemptively Guessing the Center*

Christian Konrad[1] and Tigran Tonoyan[2]

[1] Department of Computer Science, Centre for Discrete Mathematics and its Applications (DIMAP), University of Warwick, Coventry, UK
c.konrad@warwick.ac.uk
[2] ICE-TCS, School of Computer Science, Reykjavik University, Iceland
ttonoyan@gmail.com

1 Introduction

Online algorithms process their inputs item by item in a linear fashion. They are characterized by the fact that the algorithm's decision as to how to process the current input item is irrevocable. A key difficulty in the design of online algorithms is that they are typically unaware of the length of the input request sequence. Indeed, for many online problems (e.g. problems with a rent or buy flavor such as the ski rental problem), knowing the input length would allow the algorithm to solve the problem optimally. Without knowing the input length, online algorithms are unable to determine the relative position of the current element within the request sequence.

Guessing the Center. In this paper, we ask whether we can nevertheless obtain some sort of orientation within the request sequence. We study the natural task of guessing the central position $n/2$ within a request sequence of unknown length n in an online fashion. In this problem, the online algorithm maintains a guess of the central position while processing the input request sequence. The algorithm is only allowed to update its guess to the position of the current element under investigation. It may thus potentially update the guess many times, however, each update bears the risk that the input sequence may end very soon and the guess is thus far from the center. Such an algorithm follows the following scheme:

Algorithm 1 Scheme for Preemptively Guessing the Center

```
p ← 0 {initialization of our guess}
for each request j = 1, 2, ..., n do {n is unknown}
    if TODO: add condition here then {update guess}
        p ← j
return p
```

We also study a generalization of this problem to weighted requests. This is best modeled as follows. The online algorithm processes a sequence $X = w_1, w_2, \dots, w_n$ of positive integers. Let $W = \sum_{i=1}^{n} w_i$ be the total weight of

* C. Konrad is supported by the Centre for Discrete Mathematics and its Applications (DIMAP) at Warwick University and by EPSRC award EP/N011163/1.

the sequence. We assume that there exists an index m such that $\sum_{i=1}^{m} w_i = \sum_{i=m+1}^{n} w_i$, i.e., the sequence can be split into two parts of equal weight. While processing X, an online algorithm $\mathcal{A}$ maintains a guess p for the position m as in the unweighted case. The objective is to minimize the weight between the guess p and the position m of the central weight, that is, the *deviation*

$$\Delta_{\mathcal{A}}^{X} := \sum_{i=\min\{p,m\}+1}^{\max\{p,m\}} w_i \ ,$$

is to be minimized, where $\mathcal{A}$ refers to the employed algorithm and X is the input sequence. Note that the unweighted version of this problem is obtained by setting $w_i = 1$, for every $1 \leq i \leq n$.

2 Results and Techniques

For unweighted request sequences, we give an optimal randomized preemptive online algorithm for guessing the center. Our algorithm has expected deviation $0.172n$ from the central position $n/2$. Our main result is a lower bound, which shows that this is best possible. We further give a barely random algorithm that uses only a single random bit and reports a position with expected deviation $0.25n$, which is also best possible for the class of algorithms that use a single random bit. For weighted sequences, we give a randomized preemptive online algorithm that reports a position with expected deviation $0.313W$, where W is the total weight of the input sequence. This is complemented by a lower bound of $0.25W$. Closing this gap proves challenging and is left as an open problem.

Techniques. Both our algorithms for unweighted and weighted sequences employ the doubling method with a random seed. In the unweighted case, our algorithm updates its guess to the current position j if $j \in \{\lceil x^{i\delta} \rceil \mid i \in \mathbb{N}\}$ (this condition is slightly different in the weighted case), where $x > 2$ is an optimized parameter that determines the step size between the guesses (this parameter is different for weighted and unweighted sequences), and $\delta \in (0, 1)$ is a seed that is chosen uniformly at random. This technique is well known and has previously been applied for various problems. While our algorithms are extremely simple, their analyses require careful case distinctions.

Our main result is a lower bound for unweighted sequences, which proves that the doubling method is optimal. The argument relies on Yao's Minimax principle. We define a hard input distribution where the probability of a specific input length is inversely proportional to its length. We then argue that a deterministic guessing algorithm, which can be identified by a sequence of increasing positions at which it updates its guess, will in expectation (over the hard input distribution) have a deviation of $0.172n$ from the central position. By Yao's Minimax principle, this implies that our algorithm for unweighted sequences is best possible. This argument is the most technical contribution of the paper. The lower bound for weighted sequences follows the same line, however uses a sequence of exponentially increasing weights.

FRA2 : Traveling Salesman Problem

- On the probabilistic traveling salesman problem
  Soufia Benhida, Arnaud Knippel, Ahmed Mir.

- Multicommodity flow problem with sharing cost
  Pierre Laroche, Sébastien Martin, Mohamed Ould.

- Bounds calculation for the Close Enough Traveling Salesman Problem
  Raffaele Cerulli, Francesco Carrabs, Carmine Cerrone, Bruce Golden.

On the probabilistic traveling salesman problem

Soufia Benhida[12], Arnaud Knippel[2], and Ahmed Mir[1]

[1] GII, ENSA Agadir
s.benhida@gmail.com, mir-ahmed@live.com
[2] LMI, INSA Rouen Normandie
arnaud.knippel@insa-rouen.fr

Abstract. We consider the probabilistic traveling salesman problem, where the clients are present with a certain probability and the tour has to be adapted to skip the absent clients. We present some numerical results for this two-stage stochastic optimization problem for different constraint generation strategies.

Let's consider a complete graph on n nodes representing the clients. For any two nodes i and j there is a distance l_{ij}. Each client i may be present with a probability p_i, but we know which clients are present or absent only after we have chosen a tour. after a tour is chosen and we discover which clients are present, the tour is adapted by skipping the absent clients. The probabilistic traveling salesman problem, introduced in [1], consists in finding an hamiltonian circuit of minimal length expectation.

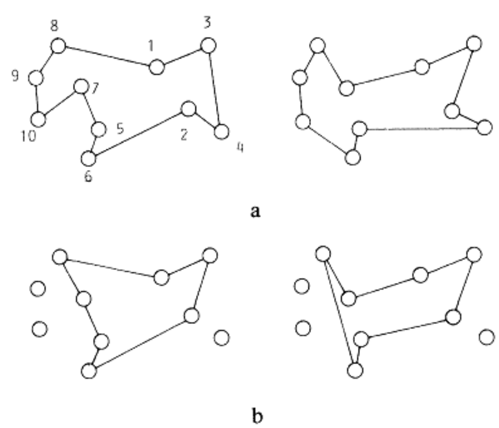


Fig. 1. This shows how the TSP tours (a) are affected by absent client nodes (b)

Finding the best a priori tour is a two-stage stochastic optimization problem. It can be formulated as an integer linear program with binary variables y_{ij} for the nominal tour (on the n nodes) and x_{ij}^k for the adapted tour in any scenario k ($1 \le k \le K$). V_k denotes the set of present nodes in scenario k. A denotes the set of arcs in the complete graph on the n nodes, and A_k the set of arcs in scenario k, once the absent client nodes have been deleted.

$$
\begin{cases}
\min \frac{1}{K} \sum\limits_{1 \le k \le K} \sum\limits_{ij \in A} l_{ij} x_{ij}^k \\
\text{Subject to} \\
\sum\limits_{j \neq i} y_{ij} = 1 \qquad \forall i \in \{1, ..., n\} & (1) \\
\sum\limits_{i \neq j} y_{ij} = 1 \qquad \forall j \in \{1, ..., n\} & (2) \\
\sum\limits_{j \neq j} x_{ij}^k = 1 \qquad \forall i \in V_k \qquad \forall k \in \{1, ..., K\} & (3) \\
\sum\limits_{i \neq j} x_{ij}^k = 1 \qquad \forall j \in V_k \qquad \forall k \in \{1, ..., K\} & (4) \\
\sum\limits_{i,j \in S} y_{ij} \le |S| - 1 \qquad \forall S \in \{1, ..., n\}, 2 \le |S| \le n-2 & (5) \\
\sum\limits_{i,j \in S} x_{ij}^k \le |S| - 1 \qquad \forall S \in V_k, 2 \le |S| \le n-2 \qquad \forall k \in \{1, ..., K\} & (6) \\
x_{ij}^k \ge y_{ij} \qquad \forall ij \in A_k \qquad \forall k \in \{1, ..., K\} & (7) \\
y_{ij} \in \{0, 1\} \qquad \forall ij \in A & (8) \\
x_{ij}^k \in \{0, 1\} \qquad \forall ij \in A_k \qquad \forall k \in \{1, ..., K\} & (9)
\end{cases}
$$

This formulation is based on the classical formulation for the TSP (with (7) being the coupling constraints) and asks for constraint generation. We discuss several strategies and give numerical results.

References

1. Patrick Jaillet, Probabilistic Traveling Salesman Problems, *PhD thesis, MIT*, 1985.
2. Patrick Jaillet, A priori solution of a travelling salesman problem in which a random subset of the customers are visited. *Operations Research*, (vol.36, No 6)pp. 929-936,1988.
3. O. Berman, D. Simchi-Levi, Finding the optimal a priori tour and location of a traveling salesman with nonhomogeneous customers, *Transportation Science* 22 (2), 148-154, 1988.
4. D.J. Bertsimas, L. Howell, Further results on the probabilistic traveling salesman problem, *European Journal of Operational Research* 65 (1), 68-95, 1995.
5. G. Laporte, F.V. Louveaux, and H. Mercure. A priori optimization of the probabilistic traveling salesman problem, *Operations Research* 42, 543-549, 1994.

Multicommodity flow problem with sharing cost

Pierre Laroche[1], Sébastien Martin[1], and Mohamed Ould[2]

[1] Université de Lorraine, LCOMS, Metz, France surname.name@univ-lorraine.fr
[2] ADIAS France mohamed.ould@adias.fr

Abstract. In this article we present a variant of the multicommodity flow problem, where the flow is 1 for each commodity, and two paths carrying flow can share the same arc and then reduce their costs. The considered variant has applications in freight transport where pooling several demands in one truck is possible. Given costs, the problem we consider here is to find a multicommodity flow with min cost. We propose an integer programming formulation for this problem and describe valid inequalities. We also devise heuristics for solving this problem. We present a comparison of the results obtained by Cplex and those obtained by the heuristics.

Keywords: multi-commodity path · integer linear program · heuristic.

The problem we consider is a restricted case of the well known multicommodity flow problem[1]. Let $D = (V, A)$ be a directed graph, where V is the set of vertices and A is the set of arcs. Each arc $a \in A$ has a capacity c_a and a weight w_a. Suppose, there are k commodities $S = \{(s_1, t_1), ..., (s_k, t_k)\}$ where s_i and t_i are the source and destination of commodity i, respectively.

The multicommodity flow where the flow is 1 for each commodity problem consists in finding k paths of minimum weight between all commodities. It is equivalent to the following integer linear program[1].

$$x^i_{uv} = \begin{cases} 1 & \text{if the arc } (u,v) \text{ is used by the path } i \\ 0 & \text{otherwise,} \end{cases} \qquad \forall (u,v) \in A, i \in \{1, ..., k\}$$

$$x_{uv} = \begin{cases} 1 & \text{if the arc } (u,v) \text{ is used by at least one path,} \\ 0 & \text{otherwise,} \end{cases} \qquad \forall (u,v) \in A.$$

$$(P) \quad \begin{aligned} & \min \sum_{(u,v)\in A} w_{uv} x_{uv} && && (1) \\ & x^i_{uv} \le x_{uv}, && \forall (u,v) \in A, i \in \{1, ..., k\} && (2) \\ & \sum_{i\in\{1,...,k\}} x^i_{uv} \le c_{uv}, && \forall (u,v) \in A && (3) \\ & \sum_{(u,v)\in\delta^+(u)} x^i_{uv} = \sum_{(u,v)\in\delta^-(u)} x^i_{uv}, && \forall i \in \{1, ..., k\} \forall u \in V \setminus \{s_i, t_i\} && (4) \\ & \sum_{(u,v)\in\delta^+(s_i)} x^i_{uv} = 1, && \forall i \in \{1, ..., k\} && (5) \end{aligned}$$

$$\sum_{(u,v)\in\delta^-(t_i)} x^i_{uv} = 1, \forall i \in \{1,...,k\} \tag{6}$$

$$x_{uv} \in \{0,1\}, \qquad \forall (u,v) \in A, \tag{7}$$

$$x^i_{uv} \in \{0,1\}, \qquad \forall (u,v) \in A, \forall i \in \{1,...,k\}, \tag{8}$$

where inequalities (2) ensure that if an arc is used by at least one path then the arc is considered by the objective function. Inequalities (3) guarantee that the number of paths which use an arc (u,v) cannot exceed the capacity c_{uv} of this arc. Inequalities (4-6) are the flow equations.

If we consider the freight problem where trucks are used to deliver some commodities from one point to another, then the multicommodity flow problem with sharing cost consists in finding k paths minimizing the total cost of the used trucks.

Let p_i be the size of the load associated with the path i. For each arc $(u,v) \in A$ we associate a size of a truck. We consider the variable z_{uv}, for each $(u,v) \in A$, corresponding to the number of trucks on the arc (u,v).

Replace inequalities (3) by the following inequalities

$$\sum_{i\in\{1,...,k\}} p_i x^i_{uv} \leq z_{uv} c_{uv}, \forall uv \in A, \tag{9}$$

and replace the objective function (1) by $\min \sum_{(u,v)\in A} p_{uv} z_{uv}$, where p_{uv} is the price of one truck on the arc (u,v) and p_i the size of the commodity.

Considering a set of paths I_{uv} using the arc (u,v) such that $\sum_{i\in I_{uv}} x^i_{uv} > c_{uv}$ we can deduce the following inequality

$$\sum_{i\in I_{uv}} x^i_{uv} \leq |I_{uv}| - 2 + 2z_{uv} + z_{uv}. \tag{10}$$

Now we can extend these inequalities by considering a set I_{uv} of paths using arc (u,v), such that $\sum_{i\in I_{uv}} p_i > \xi c_{uv}$ where ξ is the number of trucks needed to transport the demands associated with the paths of I_{uv} and deduce the following valid inequality

$$\sum_{i\in I_{uv}} x^i_{uv} \leq |I_{uv}| - (\xi + 1) + \xi z_{uv}, \tag{11}$$

if $p_i \leq c_{uv}$ for all $i \in I_{uv}$.

We propose other valid inequalities and discuss the efficiency of each class of them.

We also present a greedy algorithm, a local search algorithm and a genetic algorithm for this problem.

References

1. Lorenzo Brunetta, Michele Conforti, Matteo Fischetti. A polyhedral approach to an integer multicommodity flow problem. Discrete Applied Mathematics,101, 13-36, (2010)

Bounds calculation for the Close Enough Traveling Salesman Problem

F. Carrabs, C. Cerrone, R. Cerulli, B. Golden

Abstract The close-enough traveling salesman problem (CETSP) is a variant of the Euclidean traveling salesman problem in which the traveler visits a node if it passes through the neighborhood set of that node. We compute a lower bound of the optimal solution by discretizing the neighborhoods and by solving the generalized traveling salesman problem on the discretized graph. In order to reduce the impact on the lower bound value of the discretization error, that occurs for each discretized neighborhood, we adaptively select the neighborhoods to discretize by using a Carousel Greedy approach. It is worth noting that the computation of the lower bound is carried out by selecting only a subset of neighborhoods and, above all, this computation defines a visiting sequence of these neighborhoods. When the visiting sequence is fixed, it is possible to solve the CETSP in polynomial time by using a second order cone programming model. If the solution found by this model touches all the neighborhoods then we obtained a feasible solution for the starting problem and then we got an upper bound, otherwise one of uncovered neighborhoods is added to the sequence. The process is repeated until a feasible solution is found. The preliminary results, carried out on benchmark instances, show that our approach often overcomes the other approached proposed in literature in terms of both computational time and quality of the bounds.

Key words: Close-enough,Traveling salesman problem, Discretization scheme

F. Carrabs, R. Cerulli,
Department of Mathematics, University of Salerno, Giovanni Paolo II, 132, 84084, Fisciano (SA), Italy, e-mail: {raffaele, fcarrabs}@unisa.it

C. Cerrone
Department of Biosciences and Territory, University of Molise, Italy, e-mail: carmine.cerrone@unimol.it,

B. Golden
Robert H. Smith School of Business, University of Maryland, USA, e-mail: bgolden@rhsmith.umd.edu,

FRA3 : Mixed Integer Programming II

- Robust Queue Constrained Packing
  Christopher Bayliss, Christine Currie, Antonio Martinez-Sykora, Julia Bennell.

- An Integer Programming Approach to the Student-Project Allocation Problem with Preferences over Projects
  David Manlove, Duncan Milne, Sofiat Olaosebikan.

- A Sum-of-Squares Approach to Fairness in Combinatorial Optimisation
  Tolga Bektas, Adam Letchford.

Robust Queue Constrained Packing

Dr Christopher Bayliss[1], Dr Christine Currie[1], Dr Antonio Martinez-Sykora[2] Professor Julia A Bennell[2]

[1] Mathematical Sciences, University of Southampton, SO17 1BJ, UK
[2] Southampton Business School, University of Southampton, SO17 1BJ, UK
lncs@springer.com

Keywords: Queue constrained packing, uncertain arrival times.

We consider the problem of loading vehicles onto a ferry, where the arrival order of vehicles is stochastic and vehicles join one of a fixed number of queues on the dockside prior to loading. Vehicles vary in size from small motorbikes to large freight vehicles, but as tickets are bought in advance, the ferry operator knows how many of each vehicle type to expect. There is a penalty for each vehicle that has booked but will not fit onto the ferry and is left behind on the dockside. The objective of the optimisation is to minimize the total penalties.

The contributions of this work are as follows: an efficient and easy to implement packing methodology which addresses the queue constraints; a sample-based approximation algorithm for deriving terminal queueing policies under random vehicle arrival orders; and a new approach to obtain efficient lower bounds on the wasted space that is generated in the ferry.

The vehicle ferry loading process is divided into two stages:

Stage 1: On departure day customers who have purchased tickets in the selling season arrive at the ferry terminal at random times, close to the departure time. Upon arrival vehicles are allocated to one of a set of parallel lanes in the terminal, where they wait to embark the ferry through a single ramp entrance. The terminal queueing policies allow vehicles to be pre-sorted according to their size and any special loading requirements (for example some passengers need unrestricted access to the lifts).

Stage 2: Vehicles are loaded onto the ferry, where the order is dictated by the vehicles' positions in the queue as only vehicles currently at the front of a queue are available for loading. Where the number of queues is less than the number of vehicle types, the queue constraint usually results in more wasted space on board and hence higher penalties.

The packing methodology used is a Sequential Guillotine Cut Knapsack (SGCKS) approach, which addresses the structural constraints of the problem. In SGCKS queueing policies and packing solutions are encoded as integer strings: queueing policy strings define the target vehicle dimensions for each terminal lane, whilst packing solutions define a sequence of horizontal or vertical cuts that are then packed as rows or columns of vehicles using vehicles from the fronts of the terminal queues.

With full knowledge of the set of vehicles that will arrive and a random sample of arrival orders for these vehicles as input (the uncertainty set *S*), we formulate the problem as a two-stage stochastic optimisation, where the two stages of the problem mimic those of the process. The first stage finds the terminal queueing policy that maximises the total revenue of the vehicles that can successfully be packed in each scenario in *S*. In the second stage the yard queueing policy is fixed and new random arrival orders are realised. The packing problem can then be re-optimised with a fixed set of terminal queues as the input and an objective of minimising penalty payments for any vehicles that cannot be loaded.

We propose an iterative metaheuristic approach to solve the first stage problem. The iterative metaheuristic alternates between packing iterations and terminal queueing policy iterations. In packing iterations the incumbent queueing policy is held fixed and the metaheuristic searches the space of packing solutions for each of the arrival scenarios in *S*. In queueing policy iterations the packing solutions are held fixed and the metaheuristic searches the space of terminal queueing policies.

In the first stage problem the decision maker is provided with two main levers: the uncertainty set size *S* and a subset size *w*. Setting $w<|S|$ corresponds to maximising the revenue of vehicles that can be loaded in *w* out of the arrival scenarios in *S*. Based on this we consider two objective functions for the first stage: 1) (EXP) maximise the expected revenue of the vehicles that can be loaded in *w* out $|S|$ scenarios; and 2) (MAXIMIN) maximise the revenue of a set of vehicles that can be loaded in each of *w* out of the scenarios in *S*. For MAXIMIN we also have to define a “vehicle mix intersection” operation which tells us how many of each type of vehicle can be loaded in each and every scenario. The most basic definition of a vehicle mix intersection is the minimum number of vehicles of each type that could be packed in each scenario, hence the name MAXIMIN.

In experiments we show that the best choices for *w* relative to $|S|$ depend upon the nested vehicle size relation structure of the particular vehicle demand scenario under consideration. In particular, little or no nested vehicle sizes require $|B|\approx 0.5|S|$, whilst “Russian Doll” instances require $|B|\approx|S|$. We also demonstrate how the proposed MAXIMIN formulation leads to more robust yard policies which work better in the unseen second stage scenarios than those derived from the EXP objective function.

To illustrate the quality of the packing solutions obtained from the SGCKS packing methodology a lower bound waste formulation is proposed. This formulation is based on calculating the most efficient horizontal and vertical patterns of vehicles that fit within the ferry, and packing these on the ferry in a manner that relaxes the non-overlapping constraints. An efficient width pattern is a set of vehicles whose widths sum as close to, but not greater than, the width of the ferry. An efficient length pattern has a similar definition. An estimate of the lower bound of the wasted space can be computed from the wasted lengths associated with the most efficient width and length patterns. Such an approach can be applied iteratively to generate the efficient patterns and subtracting the vehicles in the patterns from those considered in the next iteration. An upper bound ferry utilisation can then be computed from the lower bound on the waste. Using this approach a variant of the proposed SGCKS methodology achieved an average optimality gap of 2.71% compared with the calculated upper bound, based on 300 generated 2-d rectangle packing problems.

An Integer Programming Approach to the Student-Project Allocation Problem with Preferences over Projects

David Manlove*, Duncan Milne, and Sofiat Olaosebikan**

School of Computing Science, University of Glasgow,
e-mail: David.Manlove@glasgow.ac.uk, Duncan.Milne1@gmail.com,
s.olaosebikan.1@research.gla.ac.uk

Introduction. Matching problems, which generally involve the assignment of a set of agents to another set of agents based on preferences, have wide applications in many real-world settings; for example, in an educational setting where university departments seek to allocate students to projects [1,3]. Here, we study the *Student-Project Allocation problem with preferences over Projects* (SPA-P), which involves sets of students, projects and lecturers; lecturer preferences over their proposed projects; student preferences over a subset of these projects; and the capacities of projects and lecturers (i.e., the maximum number of students that each project and lecturer can accommodate). In this context, we seek a stable matching of students to projects (and lecturers).

Informally, a *matching* is a set of acceptable (student, project) pairs such that each student is assigned at most one project, and the capacities of projects and lecturers are not exceeded; whilst a *stable matching* ensures that (i) no student and lecturer who are not matched together would rather be assigned to each other than remain with their current assignment, and (ii) no group of students acting together could undermine the integrity of the matching by swapping their assigned projects, in order to be better off.

It was shown in [4] that stable matchings in an instance of SPA-P can have different sizes, and the problem of finding a maximum size stable matching, denoted MAX-SPA-P, is NP-hard. There are two known approximation algorithms for MAX-SPA-P in the literature, with performance guarantees of 2 [4] and $\frac{3}{2}$ [2]. Moreover, it was shown in [2] that MAX-SPA-P is not approximable within $\frac{21}{19} - \varepsilon$, for any $\varepsilon > 0$, unless P = NP. In this paper, we describe an Integer Programming (IP) formulation to enable MAX-SPA-P to be solved to optimality.

An IP approach to MAX-SPA-P. Given an instance I of SPA-P, we give a general construction of an IP model J of I as follows: (i) create binary-valued variables to represent the assignment of students to projects; (ii) enforce constraints to ensure that the assignment is a matching, and that the matching is

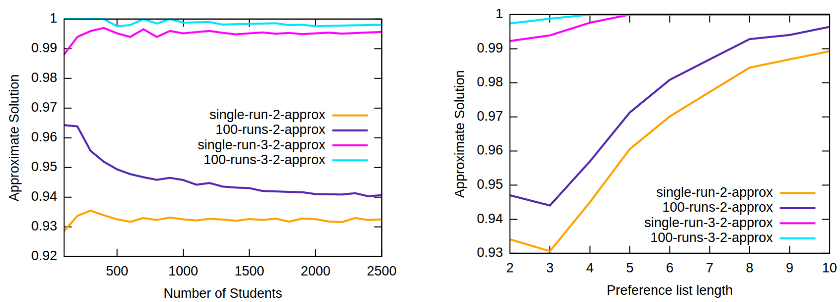
* Supported by grant EP/P028306/1 from the Engineering and Physical Sciences Research Council.

** Supported by a College of Science and Engineering Scholarship, University of Glasgow.

stable; finally (iii) describe an objective function to maximize the size of the stable matching. We present the following result regarding the correctness of J.

Theorem 1. *A feasible solution to* J *is optimal if and only if the corresponding stable matching in* I *is of maximum size.*

Empirical Analysis, Discussions and Concluding Remarks. We carried out an empirical analysis that investigates how the matchings produced by the approximation algorithms compare to optimal solutions obtained from our IP model, with respect to the size of the stable matchings constructed, on instances that are both randomly-generated and derived from real datasets. Our main finding, illustrated in Fig. 1, is that as we increase the number of students, projects and lecturers, and the length of the students' preference lists, each of the approximation algorithms finds stable matchings that are close to having maximum size, outperforming their approximation factor. Perhaps most interesting is the $\frac{3}{2}$-approximation algorithm, which finds stable matchings that are very close in size to optimal.



(a) (b)

Fig. 1. Ratio of the average size of a stable matching with respect to the optimal solution. (a) Each student's preference list contained a minimum of 2 and a maximum of 5 projects, and the number of students varied from 100 to 2500, in increments of 100. (b) The number of students was fixed at 1000, and the length of students' preference lists was varied between 2 and 10.

References

1. M. Chiarandini, R. Fagerberg, and S. Gualandi. Handling preferences in student-project allocation. *Annals of Operations Research,* to appear, 2018.
2. K. Iwama, S. Miyazaki, and H. Yanagisawa. Improved approximation bounds for the student-project allocation problem with preferences over projects. *Journal of Discrete Algorithms*, 13:59–66, 2012.
3. A. Kwanashie, R.W. Irving, D.F. Manlove, and C.T.S. Sng. Profile-based optimal matchings in the Student–Project Allocation problem. In *Proceedings of IWOCA '14: the 25th International Workshop on Combinatorial Algorithms*, volume 8986 of *Lecture Notes in Computer Science*, pages 213–225. Springer, 2015.
4. D.F. Manlove and G. O'Malley. Student project allocation with preferences over projects. *Journal of Discrete Algorithms*, 6:553–560, 2008.

A Sums-of-Squares Approach to Achieving Fairness in Combinatorial Optimisation

Tolga Bektaş[1] and Adam N. Letchford[2]

[1] Department of Decision Analytics and Risk, University of Southampton, Southampton SO17 1BJ, United Kingdom
T.Bektas@soton.ac.uk
[2] Department of Management Science, Lancaster University, Lancaster LA1 4YX, United Kingdom
A.N.Letchford@lancaster.ac.uk

Consider the following generic combinatorial optimisation problem. We have a set W of *workers* and a set T of *tasks*. Each task must be assigned to a worker. If worker $w \in W$ is assigned the set $S \subseteq T$ of tasks, then a cost $c_w(S)$ is incurred. The cost $c_w(S)$ is not given explicitly, but must be computed by solving a (smaller and simpler) combinatorial optimisation problem. We seek an assignment of tasks to workers that minimises some non-decreasing function of the $c_w(S)$ values.

Although apparently simple, this general scheme covers a surprisingly wide variety of important combinatorial optimisation problems, including problems in, e.g., vehicle routing, network design, facility location and machine scheduling. In all of these cases, a natural formulation of the problem is:

$$\begin{aligned} \min\ & \textstyle\sum_{w\in W} c_w(S_w) \\ \text{s.t.}\ & \textstyle\bigcup_{w\in W} S_w = T \\ & S_w \cap S_u = \emptyset \quad \big(\{w,u\} \subset W\big), \end{aligned} \tag{1}$$

where S_w is the set of tasks assigned to worker w.

A drawback of this formulation is that the optimal solution(s) may be *unfair*, in the sense that one worker has a significantly higher workload than another. To alleviate this problem, one can replace the min-sum objective function with an alternative function (such as min-max) and/or include additional constraints (such as lower and upper bounds on the workloads). Unfortunately, this may have unintended consequences, in terms of both computational difficulty and the quality of the solution obtained.

Of course, the issue of fairness has already been studied, not only by the combinatorial optimisation community, but also from the perspective of many other disciplines, such as computer science [4], economics [6], marketing [8], operational research [2], psychology [5] and recreational mathematics [3]. Nevertheless, we make four new contributions in this talk:

1. We show that, for several combinatorial problems of interest, the above formulation frequently yields very unfair solutions. (We use benchmark instances of the capacitated vehicle routing, capacitated minimum spanning tree and p-median problems.)

2. We attempt to explain theoretically why such unfair solutions tend to arise, with the help of some results in [1, 7].
3. To remedy this situation, we propose to minimise the sum of the squared workloads instead. That is, we minimise

$$\sum_{w \in W} \big(c_w(S_w)\big)^2.$$

4. We present some theoretical and empirical evidence that this modified objective function tends to lead to solutions that are significantly fairer than those obtained with the min-sum approach, yet perform well in terms of the min-sum and min-max objectives simultaneously.

References

1. J. Beardwood, J.H. Halton & J.M. Hammersley (1959) The shortest path through many points. *Proc. Camb. Phil. Soc.*, 55, 299–327.
2. D. Bertsimas, V.F. Farias & N. Trichakis (2011) The price of fairness. *Oper. Res.*, 59, 17–31.
3. L.E. Dubins & E.H. Spanier (1961) How to cut a cake fairly. *Amer. Math. Monthly*, 68, 1–17.
4. R.K. Jain, D.-M.W. Chiu & W.R. Hawe (1984) A quantitative measure of fairness and discrimination for resource allocation In shared computer systems. *Tech. Report DEC-TR- 301*, Digital Equipment Corporation.
5. O. Janssen (2000) Job demands, perceptions of effort-reward fairness and innovative work behaviour. *J. Occup. Organ. Psychol.*, 73, 287–302.
6. M. Rabin (1993) Incorporating fairness into game theory and economics. *Amer. Econ. Rev.*, 83, 1281–1302.
7. J.M. Steele (1981) Subadditive Euclidean functionals and nonlinear growth in geometric probability. *Ann. Prob.*, 9, 365–376.
8. L. Xia, K.B. Monroe & J.L. Cox (2004) The price is unfair! A conceptual framework of price fairness perception. *J. Market.*, 68, 1–15.

FRB1 : Graph Partitionning

- Graph Partitioning for System On Chip emulation on Multi FPGA Platforms
  Lilia Zaourar, Francois Galea.

- Compact MILP formulations for the p-center problem
  Zacharie Ales, Sourour Elloumi.

- Integer Metric Polyhedra: MIP formulations and application to Max-Cut
  Viet Hung Nguyen, Michel Minoux.

- New Partition Inequalities for the Unsplittable flow problem
  Sonia Haddad Vanier, Arnaud Knippel.

Graph Partitioning for System On Chip emulation on Multi FPGA Platforms

Lilia Zaourar✉, François Galea

CEA, LIST, Computing and Design Environment Laboratory
91191 GIF SUR YVETTE CEDEX, FRANCE.

Keywords: Partitioning, Hypergraph, Integrated Circuit

1 Introduction

Prototyping over Field Programmable Gate Arrays (FPGA) is commonly used in order to reduce verification time when designing Systems on Chip (SoC). Nowadays, the size of the SoC to be verified often exceeds the logical capacity of a single FPGA. To handle this limitation, it is possible to partition the SoC between several FPGAs and then route the logic signals using the interconnections between these FPGA.

The automation of inter-FPGA partitioning presents a significant technical challenge. The goal is to effectively split the SoC into several parts by reducing the communications between them and the length of logical paths through different partitions while respecting limited resource capacities over each FPGAs. In addition, it is suitable to reduce the number of incoming and outgoing signals of each partition. These objectives are very important and have a great influence on the performance of the emulation of SoC that we want to prototype.

This problem can be represented as a graph partitioning problem, which is a well known combinatorial optimization problem, with additional constraints. In fact, graph partitioning is one of the fundamental $NP-complete$ [1] problems and is widely studied in this community.

In this work, we present a detailed mathematical modeling of the problem taking into account all the hardware constraints of the multi-FPGA platform. We then propose several strategies to effectively solve the problem of partitioning. Finally, a study of the results obtained on real instances of Integrated Circuits (IC) will be presented.

2 Mathematical Formulation and Resolution

An IC is represented by a *netlist*, which is the description of logic cells and their interconnection by nets. This netlist is represented by a non-oriented hypergraph (V, E). The set of vertices $V = \{v_1, v_2, ..., v_N\}$ represents logic cells, and the set of hyperedges $E = \{e_1, e_2, ..., e_M\}$ corresponds to nets of the IC. It is a hypergraph because a wire can connect two or more cells.

The hypergraph is defined by its incidence matrix H such that each coefficient h_{ij} is equal to 1 if v_i is connected to hyperedge e_j, and 0 otherwise. Each hyperedge e_j is weighted by a value w_j between 0 and 1, which is the maximum criticality of all registry paths through e_j.

The problem is to place each vertex v_i on a partition p_k while minimizing a cost that depends on the fragmentation on the different partitions of each hyperedge. Let x_{ik} be a binary decision variable equal to 1 if vertex v_i is placed on partition p_k and 0 otherwise and y_{jk} an intermediate variable equal to 1 if hyperedge e_j has at least one of its ends in partition p_k and 0 otherwise. The model has multiple resource capacity constraints; each partition p_k has a c_{kr} capacity for the resource r.

The problem can be formulated as follows:

$$(P)\begin{cases} \min \sum\limits_{j} \sum\limits_{k} \sum\limits_{k'>k} w_j y_{jk} y_{jk'} & \\ \text{s.t.} \sum\limits_{k} x_{ik} = 1 & \forall i \\ h_{ij} x_{ik} \leq y_{jk} & \forall i,j,k \\ \sum\limits_{i} q_{ir} x_{ik} \leq c_{kr} & \forall k,r \\ x_{ik} \in \{0,1\}, y_{jk} \geq 0 & \forall i,j,k \end{cases}$$

We can not afford to solve the problem using exact method since our mathematical model (P) is quadratic and the size of instances can reach several millions of vertices (currently, IC contain millions of cells).

This problem is a generalization the *min-cut* problem of hypergraph partitoning with cut minimization, as treated by the *hMetis* tool [3]. It is therefore possible to obtain a first solution using *hMetis* with additional treatments to obtain a feasible solution.

Once an initial solution is obtained, several solution approaches could be considered. To this end, we used *LocalSolver 7.5* [2], developed by *Innovation24*, for combinatorial optimization entirely based on local search with an extension to mixed-variables optimization. It implements a modified local search heuristic, offering the power of local search through a model-and-run solver for large-scale $0-1$ nonlinear programming. We also implemented a simulated annealing heuristic [4]. It is a multilevel method, in which a clustering method is recursively applied to reduce the size of the problem, then simulated annealing is applied to each level.

Preliminary numerical results will be presented based on real IC instances.

References

1. M.R. Garey, D.S. Johnson. Computers and intractability, 1979.
2. Innovation 24. LocalSolver. `http://www.localsolver.com/`.
3. George Karypis and Vipin Kumar. hMetis 1.5: a hypergraph partitioning package. Technical report, Department of Computer Science, University of Minnesota, 1998. `http://www.cs.umn.edu/metis`, 1998.
4. Scott Kirkpatrick, C. Daniel Gelatt, and Mario P. Vecchi. Optimization by Simulated Annealing. *Science*, 220(4598):671–680, 1983.

Compact MILP formulations for the p-center problem

Zacharie Ales[1], Sourour Elloumi[1]

ENSTA-ParisTech / UMA, 91762 Palaiseau, France
Laboratoire CEDRIC, Paris, France
{zacharie.ales, sourour.elloumi}@ensta-paristech.fr

We consider N clients $\{C_1, ..., C_N\}$ and M potential facility sites $\{F_1, ..., F_M\}$. Let d_{ij} be the distance between C_i and F_j. The objective of the p-center problem is to open up to p facilities such that the maximal distance (called *radius*) between a client and its closest selected site is minimized.

This problem is very popular in combinatorial optimization and has many applications. We refer the reader to the recent survey [1]. Very recent publications include [5] which provides heuristic solutions and [2] on an exact solution method.

Let $\mathcal{M}$ and $\mathcal{N}$ respectively be the sets $\{1, ..., M\}$ and $\{1, ..., N\}$. The most classical formulation, denoted by (P_1), for the p-center problem (see for example [3]) considers a variable R equal to the value of a radius, the binary variables y_j equal to 1 if and only if F_j is open and the binary variables x_{ij} equal to 1 if and only if C_i is assigned to F_j.

$$
(P_1)\begin{cases}
\min R & & \text{(1a)}\\
\text{s.t.} \sum_{j=1}^{M} y_j \le p & & \text{(1b)}\\
\sum_{j=1}^{M} x_{ij} = 1 & i \in \mathcal{N} & \text{(1c)}\\
x_{ij} \le y_j & i \in \mathcal{N}, j \in \mathcal{M} & \text{(1d)}\\
\sum_{j=1}^{M} d_{ij}\, x_{ij} \le R & i \in \mathcal{N} & \text{(1e)}\\
x_{ij}, y_j \in \{0,1\} & i \in \mathcal{N}, j \in \mathcal{M} & \\
r \in \mathbb{R} & &
\end{cases}
$$

Constraint (1b) ensures that no more than p facilities are opened. Each client is assigned to exactly one facility through Constraints (1c). Constraints (1d) link variables x_{ij} and y_j while (1e) ensure the coherence of the objective.

A more recent formulation of the p-center problem, denoted by (P_2), was proposed in [4]. Let $D^0 < D^1 < ... < D^K$ be the different d_{ij} values $\forall i \in \mathcal{N}\ \forall j \in \mathcal{M}$. Note that, if many distances d_{ij} have the same value, K may be significantly lower than $M \times N$. Let $\mathcal{K}$ be the set $\{1, ..., K\}$. Formulation (P_2) is based on the variables y_j, previously introduced, and one binary variable z^k, for each $k \in \mathcal{K}$, equals to 1 if and only if the optimal radius is greater than or equal to D^k:

$$
(P_2)\begin{cases}
\min\ D^0 + \sum_{k=1}^{K}(D^k - D^{k-1})\ z^k & & (2a)\\
\text{s.t. } 1 \le \sum_{j=1}^{M} y_j \le p & & (2b)\\
\quad z^k + \sum_{j\,:\,d_{ij}<D^k} y_j \ge 1 & i \in \mathcal{N}, k \in \mathcal{K} & (2c)\\
\quad y_j, z^k \in \{0,1\} & j \in \mathcal{M}, k \in \mathcal{K} &
\end{cases}
$$

Constraints (2c) ensure that if no facility located at less than D^k of client C_i is selected, then the radius must be greater than or equal to D^k.

This formulation has been proved to be tighter than (P_1). However, its size strongly depends on the value K (*i.e.,* the number of distinct distances d_{ij}).

In this work we first prove that a large part of constraints (2c) are redundant and can be removed without affecting the quality of the linear relaxation. This leads to a formulation (CP_1) with $\mathcal{O}(\min(NM, NK))$ constraints instead of $\mathcal{O}(NK)$.

Then, we introduce (CP_2) which is the most compact formulation currently known for this problem. It is obtained by replacing the K variables z^k of (CP_1) by a unique variable r which represents the index of a radius. We prove that the linear relaxation of (CP_1) is stronger than the one of (CP_2).

We besides introduce an iterative algorithm which enables us to reduce the number of clients and facilities as well as to compute strong bounds which significantly reduce the size of formulations (P_2), (CP_1) and (CP_2).

Finally, the efficiency of the iterative algorithm and the proposed formulations are compared in terms of quality of the linear relaxation and computation time over instances from OR-Library.

References

1. Hatice Calik, Martine Labbé, and Hande Yaman. *p-Center Problems*, pages 79–92. Springer International Publishing, Cham, 2015.
2. Hatice Calik and Barbaros C Tansel. Double bound method for solving the p-center location problem. *Computers & Operations Research*, 40(12):2991–2999, 2013.
3. M Daskin. Network and discrete location analysis. *ed: John Wiley and Sons, New York*, 1995.
4. Sourour Elloumi, Martine Labbé, and Yves Pochet. A new formulation and resolution method for the p-center problem. *INFORMS Journal on Computing*, 16(1):84–94, 2004.
5. D. Ferone, P. Festa, A. Napoletano, and M. G. C. Resende. On the fast solution of the p-center problem. In *2017 19th International Conference on Transparent Optical Networks (ICTON)*, pages 1–4, July 2017.

Integer Metric Polyhedra: MIP formulations and application to Max-Cut

Viet Hung Nguyen and Michel Minoux

Sorbonne Université, CNRS, Laboratoire d'Informatique de Paris 6, LIP6, F-75005 Paris, France.

1 Introduction

Let $G = (V, E)$ be an undirected graph with $n = |V|$ and $m = |E|$. We denote by ij, the edge between the two nodes i and j of V. A *chordless cycle* C in G is a cycle whose induced subgraph is the cycle itself. Let $\mathcal{C}$ be the set of the chordless cycles in G. Let $\mathbb{R}^E$ be the real space of dimension $|E|$ indexed by the edges in E. For a vector $x \in \mathbb{R}^E$, x_e with $e \in E$ denotes the component of x associated with the edge $e \in E$ and for any subset $F \subseteq E$, let $x(F) = \sum_{e \in E} x_e$.
The *(semi)-metric polytope* METP(G) associated with G in $\mathbb{R}^E$, which can be defined as follows:

$$x(F) - x(C \setminus F) \leq |F| - 1, \quad \forall C \in \mathcal{C} \text{ and } F \subseteq C \text{ with } |F| \text{ odd}, \tag{1}$$

$$x_e \geq 0 \quad \forall e \in E \text{ s.t. } e \text{ does not belong to any triangle}$$

$$x_e \leq 1 \quad \forall e \in E \text{ s.t. } e \text{ does not belong to any triangle} \tag{2}$$

Note that the inequalities (1) are called *cycle inequalities*. Inequalities (2) are applied only for the edges in G which do not belong to any triangle as those for the other edges can be derived from the cycle inequalities. These inequalities were introduced in the seminal paper by Barahona et Mahjoub [2] on the cut polytope.

Note that since there is a priori no known polynomial upper bound (in terms of n and m) on the number of chordless cycles and there may be also an exponential number of possible choices for the set F given a chordless cycle C, the above formulation METP(G) have a priori an exponential number of inequalities. However, METP(G) has polynomial size extended formulations [1], [6], called METP(K_n), which consists of $O(n^2)$ variables where additional variables correspond to the additional edges which complete G to K_n. This extended formulations involve the following so-called *triangle inequalities*:

$$x_{ij} + x_{ik} + x_{jk} \leq 2 \text{ for all } i, j, k \in \mathcal{T}. \tag{3}$$

$$x_{ij} - x_{ik} - x_{jk} \leq 0, \tag{4}$$

$$x_{ik} - x_{ij} - x_{jk} \leq 0, \tag{5}$$

$$x_{jk} - x_{ij} - x_{ik} \leq 0 \text{ for all } i, j, k \in \mathcal{T}. \tag{6}$$

where $\mathcal{T}$ is the set of alll (unordered) triples of distinct nodes $i, j, k \in V$ such that at least ij, ik or jk is an edge in E. The semi-metric polytope are the core of the linear programming relaxations for many fundamental combinatorial optimization problems such as MaxCut. Moreover, when G is sparse, it is well known that the relaxations given by semi-metric polytope is very good. In this case, for MaxCut problem, branch-and-cut algorithms based on the integer formulation obtained from the cycle inequalities and the $0/1$ constraints remain the best approaches so far. Hence, optimizing over the semi-metric polytope appears to be of key importance. This is achieved either by cutting-planes algorithm (using the polynomial time separation algorithm for the cycle inequalities given in [2]) or by solving the compact extended formulation. However, the latter is very hard [5]) as the linear program to be solved turn out to be highly degenerate and its size could be very big even when the above compact extended formulation is used.
In this paper, we propose a way of bypassing to this difficulty. The idea is that instead of handling separately the cycle inequalities and the $0/1$ constraints over the variables x_{ij} for all $ij \in E$, we try to represent both by means of extra $0/1$ variables and a small set of constraints. Precisely, we define the *Integer Metric Polyhedra* (IMP) which are the semi-metric polytope (and the metric cone) with integrality constraints over the variables $x_{ij}$for all $ij \in E$. In particular, the *integer semi-metric polytope* coincides with the MaxCut polytope [2]. In this paper, we give Mixed Integer Programming (MIP) formulations for IMP which features only n $0/1$ variables and m continuous variables together with $O(m)$ constraints. As a consequence, it gives a very compact MIP formulation of linear size and only n $0/1$ variables for MaxCut and Graph Partitioning, etc. Computational results based on this new formulation on a series of MaxCut instances on sparse graphs involving up to 10000 nodes are discussed and shown to be competitive as compared with the best existing approaches for MaxCut.

References

1. Barahona, F.: On cuts and matchings in planar graphs. Math. Prog. 60, 53–68 (1993)
2. Barahona, F., Mahjoub, A.R.: On the cut polytope. Math. Prog. 36, 157–173 (1986)
3. Chopra, S.: The graph partitioning polytope on series-parallel and 4-wheel free graphs. SIAM J. Discret. Math. 7(1), 16–31 (1994)
4. Deza, M., Laurent, M.: Applications of cut polyhedra I. Journal of Compu. and Appl. Math. 55(2), 191 – 216 (1994)
5. Frangioni, A., Lodi, A., Rinaldi, G.: New approaches for optimizing over the semi-metric polytope. Math. Prog. 104(2-3), 375–388 (2005)
6. Nguyen, V.H., Minoux, M., Nguyen, D.P.: Reduced-size formulations for metric and cut polyhedra in sparse graphs. Networks 69(1), 142–150 (2017)

New Partition Inequalities for the Unsplittable flow problem

Sonia Haddad-Vanier[1] and Arnaud Knippel[2]

[1] SAMM, Université Paris1 Panthéon-Sorbonne
sonia.vanier@univ-paris1.fr
[2] LMI, INSA Rouen Normandie
arnaud.knippel@insa-rouen.fr

Abstract. We study some polyhedral aspects of the polytope of the minimum cost unsplittable flow problem (MCUFP). We first extend the classical cover and cut inequalities to introduce valid inequalities for the (MCUFP) polytope. Then we introduce new classes of valid inequalities, and give separation algorithms for a branch-and-cut framework.

Given a network $G = (V, E)$ defined by a set of nodes V and a set of arcs E, each arc has a capacity y_{ij}. Let D denote the set of commodities. Each commodity has an origin s, a destination t and a flow value to route d_{st}. We would like to concurrently route every demand on a single path from s to t without violating the capacities. The unsplittable flow problem has been proven to be NP-hard as a generalization of the Partition or Bin Packing problems. This combination of routing and bin packing makes the unsplittable flow problem particularly difficult.

The (MCUFP) problem can be described using a binary flow variable x_{ij}^{st} for each commodity st and arc ij that takes value of 1 if the commodity uses the arc ij, 0 otherwise. For each arc of the network there is a capacity constraint of the form : $\sum_{st \in D} d_{st} \quad x_{ij}^{st} \leq y_{ij}$. For each commodity $st \in D$ we have a flow conservation constraint. Let K_{ij}^{st} denote a unit flow cost for arc ij routing a commodity st. Let $\wp^{st}$ be the set of all possible simple paths for commodity st in the graph G. The minimum cost unsplittable flow problem can be formulated as the following integer linear program:

$$\begin{cases} \min \sum_{st \in D} \sum_{ij \in E} K_{ij}^{st} x_{ij}^{st} \\ \text{Subject to} \\ \sum_{st \in D} d_{st} \quad x_{ij}^{st} \leq y_{ij} \qquad \forall ij \in E \\ \sum_{j \in V} x_{ij}^{st} - \sum_{j \in V} x_{ji}^{st} = \begin{cases} +1 & \text{if } i = s \\ -1 & \text{if } i = t \\ 0 & \text{otherwise} \end{cases} \qquad \forall i \in V, \forall st \in D \\ x_{ij}^{st} \in \{0, 1\} \qquad \forall st \in D, \forall ij \in E \end{cases} \tag{1}$$

Let $(S, \bar{S})$ define a cut in the network, $\delta(S)$ arc set of $(S, \bar{S})$, uv an arc of the cut so $uv \in \delta(S)$. Let C be a demand cover set such as $\sum_{st \in C} d_{st} > y_{uv}$, and let CE be an extended cover set : $CE = \{st : st \in C\} \cup \{st : d_{st} \geq d_j, d_j \in C\}$. Then the following cover and extended cover inequalities are valid for (MCUFP):

$\sum_{st \in C} x_{uv}^{st} \leq |C| - 1$ and $\sum_{st \in CE} x_{uv}^{st} \leq |C| - 1$.

Now for all the other arcs $ij \in \delta(S) - \{uv\}$ we have the valid inequalities: $\sum_{ij \in \delta(S) - \{uv\}} \sum_{st \in C} x_{ij}^{st} \geq 1$ and $\sum_{ij \in \delta(S) - \{uv\}} \sum_{st \in CE} x_{ij}^{st} \geq |CE| - [|C| - 1]$.

Let $\delta_H(S)$ be the set of all the demands through the cut $(S, \bar{S})$. Then the 2-partition inequalities $\sum_{st \in \delta_H(S)} \sum_{uv \in \delta_(S)} x_{uv}^{st} \geq |\delta_H(S)|$ as well as the extended cover inequalities $\sum_{ij \in \delta(S) - \{uv\}} \sum_{st \in CE} x_{ij}^{st} \geq 1 + |CE| - |C|$ are valid for our problem.
We extended those results to introduce new multi-partition inequalities.
Let $S = (S_1, S_2, ..., S_q)$ be a q-partition in G. The following multi-partition inequalities are valid: $\sum_{t_1=1}^{q-1} \sum_{t_2=t_1+1}^{q} \sum_{ij \in \delta(S_{t_1}, S_{t_2})} y_{ij} \geq \sum_{t_1=1}^{q-1} \sum_{t_2=t_1+1}^{q} \sum_{ij \in \delta_H(S_{t_1}, S_{t_2})} \alpha_{ij} d_{ij}$
and $\sum_{t_1=1}^{q-1} \sum_{t_2=t_1+1}^{q} \sum_{ij \in \delta(S_{t_1}, S_{t_2})} \sum_{st \in D} x_{ij}^{st} \geq \sum_{t_1=1}^{q-1} \sum_{t_2=t_1+1}^{q} \sum_{ij \in \delta_H(S_{t_1}, S_{t_2})} \alpha_{ij}$
and $\sum_{ij \in \delta(S) - \{uv\}} \sum_{st \in CE} x_{ij}^{st} \geq 1 + \sum_{st \in CE - C} \alpha_{st}$.
Where $\alpha_{st} = distance_{G(S)}(S(s), S(t))$.

We will discuss a summary of computational experiments with a branch-ancut algorithm to test the effectiveness of our results.

References

1. Alper Atamtürk, Deepak Rajan, On splittable and unsplittable flow capacitated network design arc–set polyhedra, Mathematical Programming, Volume 92, Issue 2, 315–333, 2002.
2. Cynthia Barhart, Christopher A. Hane, Pamela H. Vance, Using Branch-And-Cut to Solve Origin Destination Integer Multicommodity Flow Problems. *Operations Research, INFORMS*, (vol.48, No 2), 318-326, 2000.
3. Meriema Belaidouni, Walid Ben-Ameur, On the minimum cost multiple-source unsplittable flow problem, Volume 41, Issue 3 (Journées Polyèdres et Optimisation Combinatoire), 253-273, 2007.
4. Yefim Dinitz, Naveen Garg, Michel X.Goemans, On the Single-Source Unsplittable Flow Problem. *Combinatorica*, 19 (1), 17-41, 1999.
5. Jon M. Kleinberg, Single-source unsplittable flow, FOCS '96 Proceedings of the 37th Annual Symposium on Foundations of Computer Science, 14 - 16, 1996.

FRB2 : Polyhedral Approaches III

- Alternating current optimal power flow with generator selection
  Esteban Salgado, Andrea Scozzari, Fabio Tardella, Leo Liberti.

- On the split-rank of the facet defining inequalities of mixed-integer bilinear covering set
  Hamidur Rahman, Ashutosh Mahajan.

- New polyhedral approach for the minimum energy symmetric network connectivity problem in wireless sensor network
  Fatiha Bendali, Jean Mailfert, Mourad Baiou, Salsabil Grouche.

Alternating current optimal power flow with generator selection*

Esteban Salgado[1], Andrea Scozzari[2], Fabio Tardella[3], and Leo Liberti[1]

[1] CNRS LIX, École Polytechnique, 91128 Palaiseau, France
esteban.salgado@polytechnique.edu, liberti@lix.polytechnique.fr
[2] Facoltà di Economia, Università degli Studi Niccolò Cusano-Telematica, Roma, Italy andrea.scozzari@unicusano.it
[3] MEMOTEF, Facoltà di Economia, Università "La Sapienza", Roma, Italy
fabio.tardella@uniroma1.it

Keywords: smart grid, semidefinite programming, diagonal dominance, dimensionality reduction.

Abstract.

The Alternating Current Optimal Power Flow (ACOPF) problem is as follows: given an electric power network consisting of nodes (called *buses*) and links (called *lines*) one seeks an optimal generation and distribution plan of active and reactive power under physical constraints (Ohm's and Kirchhoff's laws), and subject to power generation, voltage magnitude and current bounds on each line.

Not every bus can produce power. Those which can are called *generators*. There is often a planning issue related to their activation and deactivation. Modelling this choice implies the addition of binary variables to the model, which yields a Mixed-Integer Quadratically Constrained Quadratic Programming (MIQCQP) problem.

In this paper we study the ACOPF with selection of generators (ACOPFG). Based on the ideas in [1, 2, 3], we derive Mixed-Integer Linear Programming (MILP) formulations using Diagonally Dominant Programming (DDP) for inner and outer approximations for the ACOPF with binary variables.

The ACOPFG can be modeled as:

$$\left.\begin{array}{rr} \min\limits_{v\in\mathbb{R}^n,\, z\in\{0,1\}^g} & \langle C, V\rangle + c^\top z \\ \forall k\in\mathcal{E} & \langle A^k, V\rangle = a_k \\ \forall \ell\in\mathcal{I} & \langle B^\ell, V\rangle \le b_\ell \\ \forall w\in\mathcal{Z} & \langle Q^w, V\rangle - q_w^{\max} z_{\lceil w/2\rceil} \le q_w \\ \forall w\in\mathcal{Z} & \langle Q^w, V\rangle - q_w^{\min} z_{\lceil w/2\rceil} \ge q_w \\ & V = vv^\top \end{array}\right\} \text{(ACOPFG)}$$

Given that solving these formulations is **NP**-hard [7], one common approach is to relax the last constraint into $V \succeq 0$ ($\text{ACOPFG}_{\text{MISDP}}$).

* This work was partially supported by the PGMO grant "Projet 2016-1751H", by a Siebel Energy Institute Seed Grant, and by Università di Roma "La Sapienza" under a visiting grant for the last author.

SDP relaxations for the continuous ACOPF have been studied before [4,5,6]. Unfortunately SDPs have scalability issues and for real-life size instances it is not a suitable approach.

Based on [1, 2] we propose iterative procedures to derive inner and outer approximations for ACOPFG approximating the SDP cone by the cone of Diagonally Dominant matrices $\mathcal{D}_n = \{M \in \mathcal{S}_n \mid \forall i \;\; M_{ii} \geq \sum_{j \neq i} |M_{ij}|\}$ and its dual $\mathcal{D}_n^*$. Given that $\mathcal{D}_n \subsetneq \mathcal{S}_n \subsetneq \mathcal{D}_n^*$ and $\mathcal{D}_n$ and $\mathcal{D}_n^*$ can be described linearly, we solve

$$\left.\begin{array}{rr} \min\limits_{v \in \mathbb{R}^n,\, z \in \{0,1\}^g} & \langle C, V \rangle + c^\top z \\ \forall k \in \mathcal{E} & \langle A^k, V \rangle = a_k \\ \forall \ell \in \mathcal{I} & \langle B^\ell, V \rangle \leq b_\ell \\ \forall w \in \mathcal{Z} & \langle Q^w, V \rangle - q_w^{\max} z_{\lceil w/2 \rceil} \leq q_w \\ \forall w \in \mathcal{Z} & \langle Q^w, V \rangle - q_w^{\min} z_{\lceil w/2 \rceil} \geq q_w \\ & V \in \mathcal{K} \end{array}\right\} \text{(MI-DDP)}$$

where $\mathcal{K}$ is either $\mathcal{D}_n(U) = \{M \in \mathcal{S}_n \mid \exists\, A \in \mathcal{D}_n \text{ s.t. } M = U^\top A U\}$ or $\mathcal{D}_n^*(U)$

For the classical formulation we derive the bounds through an approach using the primal and dual formulations of ($\mathsf{ACOPFG}_{\mathsf{MISDP}}$). Given that we cannot describe a dual for the mixed-integer formulation, we add cuts with respect to eigenvectors associated to negative eigenvalues of the solutions at each step of our procedure.

Empirically we observed that these bounds are tight for ACOPFG and give promising results. We also tested ways to retrieve feasible solutions for ACOPF from our bounds with positive results on some instances, but we have failed, so far, to do it consistently.

References

1. Ahmadi, A., Hall, G.: Sum of squares basis pursuit with linear and second order cone programming. Tech. Rep. 1510.01597v1, arXiv (2015)
2. Ahmadi, A., Hall, G.: Sum of squares basis pursuit with linear and second order cone programming. Contemporary Mathematics (to appear)
3. Ahmadi, A., Majumdar, A.: Dsos and sdsos optimization: more tractable alternatives to sum of squares and semidefinite optimization. Tech. Rep. 1706.02586v1, arXiv (2017)
4. Bai, X., Wei, H., Fujisawa, K., Wang, Y.: Semidefinite programming for optimal power flow problems. Electrical Power and Energy Systems 30, 383–392 (2008)
5. Chen, C., Atamtürk, A., Oren, S.: Bound tightening for the alternating current optimal power flow problem. IEEE Transactions on Power Systems 31(5), 3729–3736 (2016)
6. Lavaei, J., Low, S.: Zero duality gap in optimal power flow problem. IEEE Transactions on Power Systems 27(1), 92–107 (2012)
7. Vavasis, S.: Quadratic programming is in NP. Information Processing Letters 36, 73–77 (1990)

On the split-rank of the facet defining inequalities of mixed-integer bilinear covering set

Hamidur Rahman and Ashutosh Mahajan

Indian Institute of Technology Bombay, Powai, Mumbai - 400 076, India.
{hrahman,amahajan}@iitb.ac.in

We consider the following mixed-integer bilinear covering set

$$S = \left\{ (x, y) \in \mathbb{Z}_+^n \times \mathbb{R}_+^n : \sum_{i=1}^{n} x_i y_i \geq r \right\},$$

where $r > 0$. This set appears in real life applications like the trim loss (or cutting stock) problem [1]. The set S is nonconvex, even its continuous relaxation S_C is nonconvex for $n \geq 2$. The linear inequality description of the convex hull of the set S, denoted as $conv(S)$ is derived by Tawarmalani et al. [2] using orthogonal disjunctive procedure. The description of $conv(S)$ consists of countably infinite number of facet defining inequalities, and is therefore not a polyhedron.

In this article we focus mainly on deriving the split-rank of the facet defining inequalities of $conv(S)$. We derive the disjunctions from which the facet defining inequalities can be constructed, and consequently, we provide an alternative proof of the validity of the inequalities derived by Tawarmalani et al. [2].

The continuous relaxation S_C of S is not a convex set for $n \geq 2$. To apply the concept of split cut, we need a closed convex relaxation of the original set. So, we take the convex hull of S_C, say S_{CH}, the tightest outer approximation of S_C. The set S_{CH} is closed and it is defined as $conv\,(S_C) = S_{CH} = \left\{(x, y) \in \mathbb{R}_+^n \times \mathbb{R}_+^n : \sum_{i=1}^{n} \sqrt{\frac{x_i y_i}{r}} \geq 1\right\}$ [2]. We study the facet defining inequalities of $conv(S)$ as split or disjunctive cuts with respect to the closed convex set S_{CH}.

We show that for $n = 1$, each facet defining inequality of $conv(S)$ has split-rank one. To prove this result we showed that simple variable disjunctions are sufficient to establish the claim.

For $n \geq 2$, we identify the facet defining inequalities having split-rank one. To establish our claim, we derive the split disjunctions for which these facet defining inequalities of $conv(S)$ is valid. For the rest of the facet defining inequalities we show that there does not exist any split-disjunction for which the ineaualities are valid, and consequently they have split-rank at least two . We prove this result mainly for $n = 2$ and extend for general positive integer n.

For the facet defining inequalities with split-rank at least two, we study them as disjunctive cuts and derive the disjunctions for which they are valid.

We then study the gap between the set S^1 (say) that is constructed considering only the rank one facet defining inequalities of $conv(S)$, and $conv(S)$. Since both the sets S^1 and $conv(S)$ are unbounded, we can not compare them in terms of their volumes. We compare them in terms of the difference between the minimum values of a given objective function over these two sets. Let $c^T x + d^T y$ be a given objective function. Assume that

$$Z_{CV} = \min_{(x,y)\in conv(S)} c^T x + d^T y, \text{ and } Z_{C1} = \min_{(x,y)\in S^1} c^T x + d^T y.$$

Here, by the gap between S^1 and $conv(S)$ we mean the difference between the values of Z_{CV} and Z_{C1}. Clearly $Z_{C1} \leq Z_{CV}$ as $conv(S) \subset S^1$.

We provide the necessary and sufficient condition for which the gap between Z_{CV} and Z_{C1} is zero. We also show with an example that this gap can be arbitrary large.

To show the effectiveness of the rank one facet defining inequalities, we did some computational experiments on some instances of cutting stock problems. We theoretically show that for a relaxation of the cutting stock problems (without the Knapsack constraints), only rank one inequalities are sufficient to give the same bound as adding all the facet defining inequalities of each bilinear constraint that is present in the cutting stock formulation. To check whether the result holds with the knapsack constraints also, we performed a computational experiment on benchmark problems. We compare the bounds obtained using the above two approaches and the number of steps taken. The results show that the optimization over S^1 gives the same bound as that over the convex hull in much fewer iterations for all input problems. Below we preset the computational results for few instances that we used for our experiments to see the effectiveness of the rank one cuts.

Instances	n	Using inequalities for S^1 only			Using inequalities for $conv(S)$		
		Iterations	Termination	Lower Bound	Iterations	Termination	Lower Bound
Fiber-10-9080	10	6	Yes	3.8505	253	Yes	3.8505
Fiber-14-9080	14	5	Yes	1.9006	470	Yes	1.9006
Fiber-15-5180	15	6	Yes	3.7394	549	Yes	3.7394
Fiber-15-9080	15	6	Yes	2.1147	475	Yes	2.1147
Fiber-16-5180	16	5	Yes	5.1701	633	Yes	5.1701
Fiber-16-9080	16	6	Yes	2.9283	800	No	2.9283
CutGen-01-02	10	5	Yes	0.9744	245	Yes	0.9744
CutGen-01-25	10	5	Yes	0.9984	218	Yes	0.9984
CutGen-02-40	10	5	Yes	10.4068	248	Yes	10.4068
CutGen-02-60	10	5	Yes	10.0957	259	Yes	10.0957
Rand-15	15	5	Yes	650.29	440	Yes	650.29
Rand-20	20	5	Yes	624.92	800	No	624.92

Table 1: Comparison of iterations : Over the inequalities of S^1 and $conv(S)$.

References

1. Harjunkoski, I., Westerlund, T., Porn, R., Skrifvars, H.: Different transformations for solving nonconvex trim-loss problems by MINLP. European Journal of Operetions Research 105, 594 – 603 (1998)
2. Tawarmalani., M., Richard, J.P., Chung, K.: Strong valid inequalities for orthogonal disjunctions and bilinear covering sets. Mathematical Programming B 124, 481 – 512 (2010)

New polyhedral approach for the minimum energy symmetric network connectivity problem in wireless sensor network

Mourad Baïou, Fatiha Bendali, Salsabil Grouche, Jean Mailfert

Laboratoire LIMOS-CNRS UMR 6158, Campus universitaire des Cézeaux,1 rue de la Chébarde. TSA 60125.CS 60026. 63177 Aubière Cedex, France.
{baiou,bendali,grouche,mailfert}@isima.fr,

1 Introduction

A wireless sensor network (WSN) is composed by a large number of autonomous units. These units work together for performing a common task as data recovery or message delivery. Communications between sensors are made by radio. But as no wired backbone infrastructure is installed, these devices have to be organized in a coherent network.

In the *minimum energy symmetric network connectivity* problem (MESNCP for short), we consider a set of sensors in the plane, and we aim to determine the transmission power associated with each device such that there exists at least one communication path between each pair of sensors, and the overall sum of power consumption is minimized. As the *MESNCP* is *NP-hard* [4,5], some authors have proposed approaches based on integer programming formulations and approximation algorithms for this problem [1,2,6,7].

A Wireless sensor network is modeled by an undirected graph $G = (V, E)$, where each sensor is associated to a node in V and an edge $e = (u, v)$ in E represents a virtual wire between sensors u and v. For any edge $e = (u, v)$, a positive real $c(e)$ indicates the necessary power that u and v must use for a direct radio communication between each other. In the following, we suppose that all $c(e)$ are distinct positive numbers.
A *power* assignment in G is any function $p : V \to \mathbb{R}_+$. Indeed, if the power $p(u)$ assigned to a sensor u exceeds $c(u, v)$ for some $v \in V$, then the sensor v is able to receive the signal broadcasted by u. Thus, the function p controls the connectivity of G.

2 Description

For a subset $F \subseteq E$, its *incidence vector* $x^F \in \mathbb{R}^{|\mathbb{E}|}$ is defined by $x^F(e) = 1$ if $e \in F$ and $x^F(e) = 0$ otherwise, for $e \in E$.

Denote by $\mathcal{F}$ the set of subsets F of edges of E such that the spanning subgraph (V, F) is connected. Given $F \in \mathcal{F}$, define the power assignment vector p_F induced by F, as

$$p_F(u) = \max\{c(e) : e \in F \cap \delta(u)\}, \text{ for all } u \in V.$$

Let $\Pi(G, c) \subseteq \mathbb{R}^{|E|\times|V|}$ be the polyhedron

$$\text{conv}\{(x^F, p_F) : F \in \mathcal{F}\} + \text{cone}\{(0_E, r_1), (0_E, r_2), \ldots, (0_E, r_{|V|})\},$$

where 0_E is the zero vector in $\mathbb{R}^{|E|}$ and $r_i \in \mathbb{R}^{|V|}$, $i = 1, \ldots, |V|$, are the basic unit vectors, that is $r_i(j) = 1$ if $i = j$ and 0 otherwise.

Then, the minimum energy symmetric network connectivity problem can be formulated as follows:

$$MESNC(G, c) = \min\{\sum_{u \in V} p(u) : (x, p) \in \Pi(G, c)\}.$$

We will describe a linear relaxation of $\Pi(G, c)$. Several valid inequalities will be given. Partition inequalities will express the connectivity of a solution and a new family of valid inequalities is proposed for the power assignments. Dimension and facets of $\Pi(G, c)$ are studied.

A Branch-and-Cut algorithm based on partition and power inequalities is presented. It is known that the separation problem of partition inequalities has been reduced to $|V|$ min-cut problems in [3]. We show that the separation of the new power inequalities may be solved by a procedure based on Bellman shortest path algorithm. Some computational results will conclude this presentation.

References

1. E. Althaus, G. Calinescu, I.I. Mandoiu, S. Prassad, N. Tchervenski, and A. Zelikovsky, *Power efficient range assignment in ad-hoc wireless networks.* Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC'03), p1889-1894, 2003.
2. E. Althaus, G. Calinescu, I.I. Mandoiu, S. Prassad, N. Tchervenski, and A. Zelikovsky, *Power efficient range assignment for symmetric connectivity in static ad-hoc wireless networks.* Wireless Networks, 12(3), p287-299, 2006.
3. F. Barahona *Separating from the dominant of the spanning tree polytope.* Operations Research letters, 12, p201-203 1992.
4. M. Cagalj, J.P. Hubaux, C. Enz, *Minimum energy broadcast in all wireless networks : NP-completeness and distribution issues*, ACM Mobicom, 2002.
5. B. Fuchs, *On the hardness of range assignment problems.* Networks 52(4), p183-195, 2008.
6. X. Li, Y.P. Aneja, and J. Huo *A robust branch-and-cut approach for the minimum energy symmetric networks connectivity problem*, Omega 40, p 210-217, 2012.
7. R. Montemanni, and L.M. Gambardella. *Exact algorithms for the minimum power symmetric connectivity networks problem in wireless networks*, Computers and Operations Research, p2891-2904, 2005.

FRB3 : Mixed Integer Programming III

- An SDP approach for minimizing convex ordered median location problems in finite dimension and with general l_t-norms.
  Safae Elhaj-Ben-Ali..

- A multigraph formulation for the Generalized Minimum Spanning Tree Problem
  Ernando Gomes De Sousa, Rafael Andrade, Andréa Cynthia Santos.

- A Generalization of the Minimum Branch Vertices Spanning Tree Problem
  Massinissa Merabet, Jitamitra Desai, Miklos Molnar.

An SDP approach for minimizing convex ordered median location problems in finite dimension and with general ℓ_τ-norms

Safae ELHAJ-BEN-ALI

ENSA-Fés

Abstract. The main goal of this work is to design a common approach to solve the continuous l_p minisum location problem and moreover, all the class of convex ordered location problems, for different distances and in any finite dimension. We prove that this approach has a polynomial worse case complexity. Thus, providing a unifying new algorithmic paradigm for this class of location problems. First, it avoids the problems of limit convergence proven for the Weiszfeld type algorithms. Then, it can be applied to any convex ordered median problem, even with mixed norms, in any dimension. Moreover, we show an explicit reformulation of these problems as SDP problems which enables the usage of standard free source solvers (SEDUMI, SDPT3,...) to solve them up to any degree of accuracy. This is essentially the second goal of this work, it was already known that convex location problems with l_1 norm were reducible to linear programming. This work proves that most convex continuous location problems with l_p norms are reducible to SDP programming showing the similarities existing between all this class of problems.

A Multigraph Formulation for the Generalized Minimum Spanning Tree Problem

Ernando Gomes de Sousa[1], Rafael Castro de Andrade[2] and Andréa Cynthia Santos[3]

[1] Federal Institute of Education, Science and Technology of Maranhão
BR-230, Km 319, Zona Rural, CEP: 65840-000, São Raimundo das Mangabeiras - Maranhão - Brazil
ernando.sousa@ifma.edu.br
[2] Department of Statistics and Applied Mathematics
Federal University of Ceará
Campus do Pici, Bloco 910. CEP 60.455-900 - Fortaleza, Ceará - Brazil
rca@lia.ufc.br
[3] Charles Delaunay Institute, LOSI
Technological University of Troyes
12, rue Marie Curie, CS 42060, 10004, Troyes Cedex, France
andrea.duhamel@utt.fr

The Generalized Minimum Spanning Tree Problem (GMSTP) is defined in a connected, undirected and m-partite complete graph $G = (V, E)$. Its vertex set V is partitioned into m clusters, with $V = V_1 \cup V_2 \cup ... \cup V_m$ and $V_r \cap V_q = \emptyset$, for all $r \neq q$ with $r, q \in M = \{1, ..., m\}$. Its edge set E is given by $E = \{\{i, j\} \mid i \in V_r,\ j \in V_q\}$, for all $r \neq q$ with $r, q \in M$, for which $c_e \in R_+$ denotes the edge cost of $e = \{i, j\} \in E$. GMSTP consists in finding a minimum cost tree, spanning a unique vertex in each cluster, with exactly $m - 1$ edges connecting the m clusters. GMSTP extends the so-called Minimum Spanning Tree (MST), which belongs to the NP-hard class of problems [5] and has application in network design, irrigation agriculture, smart cities, data science, among others.

In regarding mathematical models, Myung, Lee and Tcha [5] presented four formulations for GMSTP. One formulation has a polynomial number of constraints and variables, while the others have an exponential number of constraints. In the study of Feremans, Labb and Laporte [2], four formulations have an exponential number of constraints. In addition, a mathematical formulation proposed by Pop [7] makes use of a graph G and a support graph G', built as follows. Each vertex of G' represents a cluster of G, and the edges of G' correspond to connections between two clusters of G. The idea is to address the spanning tree constraints in G', while specific GMSTP constraints are handled using G.

The GMSTP formulation proposed here is inspired on the formulation of Andrade [1]. We propose a novel mathematical formulation based on multigraph, which performs very well on known instances from the literature. Given a graph G, previously defined, the multigraph $H(G) = (V', E')$ is obtained as follow. Each cluster of G is considered as a vertex of V' and each edge $e \in E$ corresponds to exactly an edge $e' \in E'$ of same cost c_e, with $|E| = |E'|$.

Experiments were performed with the goal of evaluate the performance of the proposed model ($\mathcal{P}_{andr}$) compared to the following GMSTP formulations from literature: the polynomial multicommodity flow formulation of Myung, Lee and Tcha [5] and the formulation of Pop [7], refereed respectively here as ($\mathcal{P}_{myung}$) and ($\mathcal{P}_{pop}$).

A benchmark set of 40 instances is used in the experiments: 20 instances are original from Öncan, Cordeau and Laporte [6] and the 20 remaining were created by taking a subset of vertices of instances in [6]. Initially, we apply the preprocessing procedure of Ferreira et. al. [4] that allowed to fix out of the solution about 85% of the edges from the original graphs.

Considering the set of 40 instances tested, the proposed formulation ($\mathcal{P}_{andr}$) proved optimality for all instances, while ($\mathcal{P}_{myung}$) proved optimality for 32 instances and the model ($\mathcal{P}_{pop}$) proved optimality for 17 instances. The formulation ($\mathcal{P}_{andr}$) had an average runtime of 668.05 seconds, whereas ($\mathcal{P}_{myung}$) had an average of 1423.65 seconds and ($\mathcal{P}_{pop}$) of 2323.45 seconds. The formulations reached an average GAP of 0.00, 45.00 and 9.66 for the ($\mathcal{P}_{andr}$), ($\mathcal{P}_{myung}$), and ($\mathcal{P}_{pop}$) models, respectively.

In summary, the proposed formulation presents a competitive performance with the models of ($\mathcal{P}_{myung}$) and ($\mathcal{P}_{pop}$). In particular, by founding better lower bounds and by proving optimality for larger instances in the time limit than the other models. The originality of the proposed formulation by considering a multigraph opens new research directions in terms of adaptation for related problems, as well as for its use in cutting plane and decomposition approaches. We intend to develop a Benders decomposition and a B&C using the proposed model, and to develop valid inequalities to strengthen the model.

References

1. Andrade, R. C.: Appointments on the spanning tree polytope. Annals of the ALIO/EURO - Montevideo (2014)
2. Feremans, C., Labbé, M. and Laporte, G.: A comparative analysis of several formulations for the generalized minimum spanning tree problem. Networks, 39(1), 29-34 (2002)
3. Feremans, C., Labbé, M. and Laporte, G.: The generalized minimum spanning tree problem: Polyhedral analysis and branch-and-cut algorithm. Networks, 43(2), 71-86 (2004)
4. Ferreira, C. S., Ochi, L. S., Parada, V. and Uchoa, E.: A GRASP-based approach to the generalized minimum spanning tree problem. Expert Systems with Applications, 39(3), 3526-3536 (2012)
5. Myung, Y. S. and Lee, C. H. and Tcha, D. W.: On the generalized minimum spanning tree problem. Networks, 26(4), 231-241, (1995)
6. Öncan, T., Cordeau, J. and Laporte, G.: A tabu search heuristic for the generalized minimum spanning tree problem. European Journal of Operational Research, 191(2), 306-319 (2008)
7. Pop, P. C. The generalized minimum spanning tree problem. Ph.D. Thesis, University of Twente, Netherlands (2002)

A Generalization of the Minimum Branch Vertices Spanning Tree Problem

Massinissa Merabet[1], Jitamitra Desai[2], and Miklos Molnar[3]

[1] SAMOVAR lab
National School of Computer Science for Industry and Business (ENSIIE), France
massinissa.merabet@ensiie.fr
[2] School of Mechanical and Aerospace Engineering
Nanyang Technological University (NTU), Singapore
jdesai@ntu.edu.sg
[3] The Montpellier Laboratory of Informatics, Robotics and Microelectronics
University of Montpellier, France
miklos.molnar@lirmm.fr

Abstract. The *k-Minimum Branch Vertices Spanning Tree* (k-MBVST) problem is to find a spanning tree of graph with the minimum number of k-branch vertices. Well-developed applications related to routing in optical networks are known. We propose Proofs of NP-hardness and non-inclusion in the APX class as well as an ILP formulation of the k-MBVST problem. Computational results based on randomly generated graphs show the efficiency of your resolution method.

Keywords: Spanning Tree, Minimization of Branch Vertices, Integer Linear Programming (ILP), MBVST, k-MBVST, Optical Networks.

1 Introduction and motivation

Given a connected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, a vertex $v \in \mathcal{V}$ is defined to be a *branch vertex* in a spanning tree if its degree (denoted $d_{\mathcal{G}}(v)$) is strictly greater than two, i.e., $d_{\mathcal{G}}(v) > 2$. The MBVST problem is to find a spanning tree of graph $\mathcal{G}$ with the minimum number of branch vertices. This NP-hard and non-APX problem has been well-studied in the literature [Mar15]. The most widespread application of such MBVST problems arises in multicast routing protocols in WDM networks. From a computational viewpoint, they are mainly based on *light-trees*, which require intermediate nodes to have the ability to split and direct the input signal to multiple outputs as and when necessary. Such a node is equipped with a light-splitters which are rather expensive devices. Moreover, if a light signal is split into k copies, then the signal power of each resultant copy is reduced by, at least, a factor of $1/k$ of the original signal power. If k is too large, then the information cannot be deciphered at the destinations due to the signal strength dropping below the minimum threshold value, and therefore, k functions as a limiting (tolerance) parameter. A k-branch vertex is a vertex with degree strictly greater than $k + 2$ in the spanning tree. It is useful to look for a light-tree in the WDM network with the minimum number of k-branch vertices. This leads to our problem :

Definition 1. *Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a graph. The k-MBVST problem consists of finding a spanning tree $\mathcal{T}$ of $\mathcal{G}$ such that the number of k-branch vertices in $\mathcal{T}$ is minimized.*

2 NP-hardness and negative approximability

$s_k(\mathcal{G})$ is the smallest number of k-branch vertices in any spanning tree of $\mathcal{G}$.

Theorem 1. *Let r be a fixed non-negative integer. It is NP-complete to decide whether a given graph $\mathcal{G}$ satisfies $s_k(\mathcal{G}) \leq r$ for any value of k.*

The proof is based on a reduction from the *Hamiltonian problem* to the k-MBVST problem (see the regular paper).

Theorem 2. *The k-MBVST problem is not in APX for any value of k.*

Proof. The proof is based on an AP-reduction from the *Minimum Set Cover problem* to the k-MBVST problem (see the regular paper).

3 ILP formulation and computational results

The formulation of the k-MBVST problem as an *integer linear program* (**ILP**) derived in this paper is predicated on the single balance commodity flow formulation proposed in [CGI09]. However, it is worthwhile to exploit the underlying graph to ascertain which vertices must necessarily be, can never be, or could possibly be k -branch vertices in the optimal solution [LMSP17]. Computing a tighter upper bound on the maximum quantity of flow transiting on the graph edges and deploying a tight constraint to check if a vertex is k-branch or not, we propose a significantly improved version and compar it to the classical formulation in the regular paper.

Computational results based on randomly generated graphs show that the number of k-branch vertices included in the spanning tree increases with the size of the vertex set $\mathcal{V}$, but decreases with k as well as graph density. We also show that when $k \geq 4$, the number of k-branch vertices in the optimal solution is close to zero, regardless of the size and the density of the underlying graph.

4 Conclusion

Due to its importance, we propose a generalization of the MBVST problem by introducing the notion of the k-branch vertex. Our new parametrized problem (k-MBVST) aims to find a spanning tree of $\mathcal{G}$ with the minimum number of k-branch vertices. For any non-negative integer r, we proved that it is NP-complete to decide whether a graph can be spanned by a tree with at most r k-branch vertices. Furthermore, we also established that the k-MBVST is hard to approximate. We also proposed an integer linear programming formulation based on a single commodity flow balance constraints. Tests on random graphs allowed us to evaluate the number of k-branch vertices in the optimal solution.

References

[CGI09] R. Cerulli, M. Gentili, and A. Iossa. Bounded-degree spanning tree problems: models and new algorithms. *Computational Optimization and Applications*, 42:353–370, April 2009.

[LMSP17] M. Landete, A. Marín, and J. Sainz-Pardo. Decomposition methods based on articulation vertices for degree-dependent spanning tree problems. *Computational Optimization and Applications*, 68(3):749–773, Dec 2017.

[Mar15] A. Marín. Exact and heuristic solutions for the minimum number of branch vertices spanning tree problem. *European Journal of Operational Research*, 245(3):680 – 689, 2015.

Index