Welcome to ORBEL 32!

It is with great pleasure that we welcome you in Liège for the 32nd edition of the Belgian Operational Research Society's annual conference. This year's edition offers a rich program with three keynote speakers and more than 80 talks contributed by authors and coauthors from Belgium and from Algeria, France, Germany, Iran, Italy, Luxembourg, the Netherlands, Switzerland, and Turkey. We are grateful to our three distinguished keynote speakers for agreeing to speak at our conference: we are sure that you will enjoy their talks. Thanks to all of you, we will have 26 parallel contributed sessions, one of which is dedicated to the talks of the 2018 ORBEL Award candidates. The ORBEL Award is handed out each year to the best student thesis in operations research and is sponsored by OM Partners. The winner of the award will be announced during the ORBEL Award ceremony on Friday, after which you will be kindly invited to the closing cocktail.

ORBEL 32 is organized by the HEC Liège Centre for Quantitative Methods and Operations Management (QuantOM). More than a formal research institute, QuantOM is an open association of scientists who work under a common label in order to promote and to stimulate the development of research conducted at HEC Liège (or more broadly, within the ULiège) in the field of quantitative methods and of their applications to supply chain management, logistics, operations, and other areas of management and economics. The organization of this conference would not have been possible without the enthusiasm and the dedication of its members, and the help of the administrative staff at HEC Liège.

The conference is hosted in the main building of HEC Liège, Management School of the University of Liège, one of the leading Belgian university business schools for undergraduate and graduate programs with more than 115 full-time faculty members and researchers and more than 2600 students. The school's mission is to train creative managers who will be responsible for building the future of businesses and organizations in a cross-cultural world. The international vision of HEC Liège translates into multiple research activities in management and economics, numerous partnerships with worldwide companies and universities, and growing internationalization of its programs and faculty. HEC Liège is EQUIS accredited. Its English-speaking Masters' programs in Management and in Business Engineering are EPAS-accredited. Its PhD program was first worldwide to receive an EPAS quality label. HEC Liège is also member of the "Conférence des Grandes Ecoles" based in France.

We wish you a very interesting and fruitful meeting and a pleasant stay in Liège.

ORBEL 32 Organizing Committee

Organizing committee
Arda, Yasemin (Chair)
Bay, Maud
Crama, Yves
François, Véronique
Hoffait, Anne-Sophie
Limbourg, Sabine
Michelini, Stefano
Paquay, Célia
Pironet, Thierry
Rodríguez-Heck, Elisabeth
Schyns, Michaël
Smeulders, Bart

Scientific committee
Aghezzaf, El Houssaine
Arda, Yasemin
Beliën Jeroen
Bisdorff, Raymond
Blondeel, Wouter
Caris, An
Chevalier, Philippe
Crama, Yves
David, Benoît
De Baets, Bernard
De Causmaecker, Patrick
De Smet, Yves
Fortz, Bernard
Goossens, Dries
Janssens, Gerrit
Kunsch, Pierre
Labbé, Martine
Leus, Roel
Mëlot, Hadrien
Pirlot, Marc
Sartenaer, Annick
Schyns, Michaël
Sörensen, Kenneth
Spieksma, Frits
Vanden Berghe, Greet
Vansteenwegen, Pieter
Van Utterbeeck, Filip
Van Vyve, Matthieu
Wittevrongel, Sabine

Plenary session I: Thursday 9:30 - 10:30

Prof. Dominique Feillet
Ecole des Mines de Saint-Etienne and LIMOS
Head of the Logistics and Manufacturing Sciences department

Vehicle routing problems with road-network information

Since the introduction of the Vehicle Routing Problem more than 50 years ago, routing is defined as "the process of selecting best routes in a complete graph". However, there are several situations where the abstraction of the road network into a complete graph (the so-called customer-based graph) is at best disputable, at worst can lead to a bad optimization of vehicle routes. In this presentation we will discuss about the possibility of considering more information from the road network. We will review the literature, detail several situations where additional information from the road network could be helpful and describe some consequences on exact or heuristic solution approaches.

Dominique Feillet works as a Professor of Operations Research at Mines Saint-Etienne, a leading engineering school in France. He received an engineering degree in Computer Sciences from ENSIMAG (Grenoble, France) and holds a phD in Industrial Engineering from Ecole Centrale Paris (France). He joined Mines Saint-Etienne in 2008, where he is head of the Department on Manufacturing Sciences and Logistics since 2013. He entered the LIMOS laboratory, attached to the French National Center for Scientific Research (CNRS), in 2014 and is heading the research team on Decision-making tools for Production and Logistics since 2015.
His primary research interest concerns the development of relevant discrete optimization models and methods with regards to new practices in transportation and distribution. He is particularly interested in vehicle routing optimization, but has also been involved in several collaborative projects with railway or shipping industries.
His research has resulted in more than 50 publications in first-rank journals like Transportation Science, EURO Journal on Transportation and Logistics, EJOR, Networks or Computers & OR. He was finalist of the VEROLOG Solver challenge in 2014 and winner of the Scientific Prize of the EURO/ROADEF challenge in 2016. He is associate editor of EJTL and member of the advisory board of Computers & OR. He is a former secretary of the French Operations Research association (ROADEF). He has been involved in the organization of several national and international events as ROADEF'2003, NOW'2006, ROADEF'2010 or Odysseus 2015. He was invited several times in summer schools to give talks on vehicle routing or column generation.

Plenary session II: Friday 9:30 - 10:30

Prof. Martin Savelsbergh
H. Milton Stewart School of Industrial & Systems Engineering, GeorgiaTech
Chair and Co-Director of Supply Chain & Logistics Institute

Recent Advances in Criterion Space Search Algorithms for Multi-objective Mixed Integer Programming

Multi-objective optimization problems are pervasive in practice. In contrast to single-objective optimization, the goal in multi-objective optimization is to generate a set of solutions that induces the Pareto front, i.e., the set of all nondominated points. A nondominated point is a vector of objective function values evaluated at a feasible solution, with the property that there exists no other feasible solution that is at least as good in all objective function values and better in one or more of them. Recently, criterion space search algorithms, in which the search for the Pareto front takes place in the space of the vectors of objective function values, i.e., the criterion space, have gained in popularity. These methods exploit the advances in single-objective optimization solvers, since they repeatedly solve single-objective optimization problems. We will introduce and discuss criterion space search algorithms for both pure and mixed multi-objective integer programs.

Martin Savelsbergh is James C. Edenfield Chair and Professor at the H. Milton Stewart School of Industrial & Systems Engineering. Martin is an optimization and logistics specialist with over 20 years of experience in mathematical modeling, operations research, optimization methods, algorithm design, performance analysis, logistics, supply chain management, and transportation systems. He has published over 150 research papers in many of the top optimization and logistics journals and has supervised more than 25 Ph.D. students. Martin has a track record of creating innovative techniques for solving large-scale optimization problems in a variety of areas, ranging from supply chain master planning and execution, to world-wide tank container management, to service network design, to production planning, and to vehicle routing and scheduling. Martin has given presentations and short courses on optimization, transportation, and logistics in more than a dozen countries around the world.

Ongoing research projects that Martin is pursuing include innovations in last-mile delivery, advances in dynamic ride-sharing, methods for multi-objective optimization, and dynamic management of time-expanded networks. Martin Savelsbergh is the Editor-in-Chief of Transportation Science, the flagship journal of INFORMS in the area of transportation and logistics. He has served as Area Editor for Operations Research Letters and as Associate Editor for Mathematics of Operations Research,

Operations Research, Naval Logistics Research, and Networks. Martin has served as president of the Transportation and Logistics Society of INFORMS and is an INFORMS Fellow.

Plenary session III: Friday 15:30 - 16:30

Prof. Michel Bierlaire
EPFL Ecole Polytechnique Fédérale de Lausanne

Modeling advanced disaggregate demand as MILP

Choice models are powerful tools to capture the detailed choice behavior of individuals, characterizing the demand at a disaggregate level. Although many such models are available in the literature, they are very rarely integrated in optimization models. The main reason is that they are non-linear, non-convex and often not available in closed form. We propose a modeling framework that allows to represent (almost) any choice model as constraints for MILP. The talk describes the framework and illustrates its relevance on some examples.

Michel Bierlaire holds a PhD in Mathematical Sciences from the Facultés Universitaires Notre-Dame de la Paix, Namur, Belgium (University of Namur). Between 1995 and 1998, he was research associate and project manager at the Intelligent Transportation Systems Program of the Massachusetts Institute of Technology (Cambridge, Ma, USA). Between 1998 and 2006, he was a junior faculty in the Operations Research group ROSO within the Institute of Mathematics at EPFL. In 2006, he was appointed associate professor in the School of Architecture, Civil and Environmental Engineering at EPFL, where he became the director of the Transport and Mobility laboratory. Since 2009, he is the director of TraCE, the Transportation Center. From 2009 to 2017, he was the director of Doctoral Program in Civil and Environmental Engineering at EPFL. In 2012, he was appointed full professor at EPFL. Since September 2017, he is the head of the Civil Engineering Institute at EPFL.

His main expertise is in the design, development and applications of models and algorithms for the design, analysis and management of transportation systems. Namely, he has been active in demand modeling (discrete choice models, estimation of origin-destination matrices), operations research (scheduling, assignment, etc.) and Dynamic Traffic Management Systems.

As of December 2017, he has published 113 papers in international journals (including Transportation Research Part B, the transportation journal with the highest impact factor), 4 books, 39 book chapters, 170 articles in conference proceedings, 160 technical reports, and has given 187 scientific seminars. His ISI h-index is 25. His Google Scholar h-index is 51.

He is the founder, organizer and lecturer of the EPFL Advanced Continuing Education Course "Discrete Choice Analysis: Predicting Demand and Market Shares".

He is the founder and the chairman of hEART: the European Association for Research in Transportation. He is the Editor-in-Chief of the EURO Journal on Transportation and Logistics. He is an Associate Editor of Operations Research and of the Journal of Choice Modelling. He is the editor of two special issues for the journal Transportation Research Part C. He has been member of the Editorial Advisory Board (EAB) of Transportation Research Part B since 1995, of Transportation Research Part C since January 1, 2006, and of the journal "European Transport" since 2005.

Sponsors

We gratefully acknowledge support from our partners:

HEC-Liège		http://www.hec.ulg.ac.be
ICRON		http://icrontech.com
N-SIDE		http://www.n-side.com
OM Partners		http://ompartners.com
ULiège		http://www.uliege.be

Practical information

The whole conference takes place at the main building of HEC Liège, Rue Louvrex 14, 4000 Liège. The building is located close to the city centre and 15 minutes by walk from the train station.

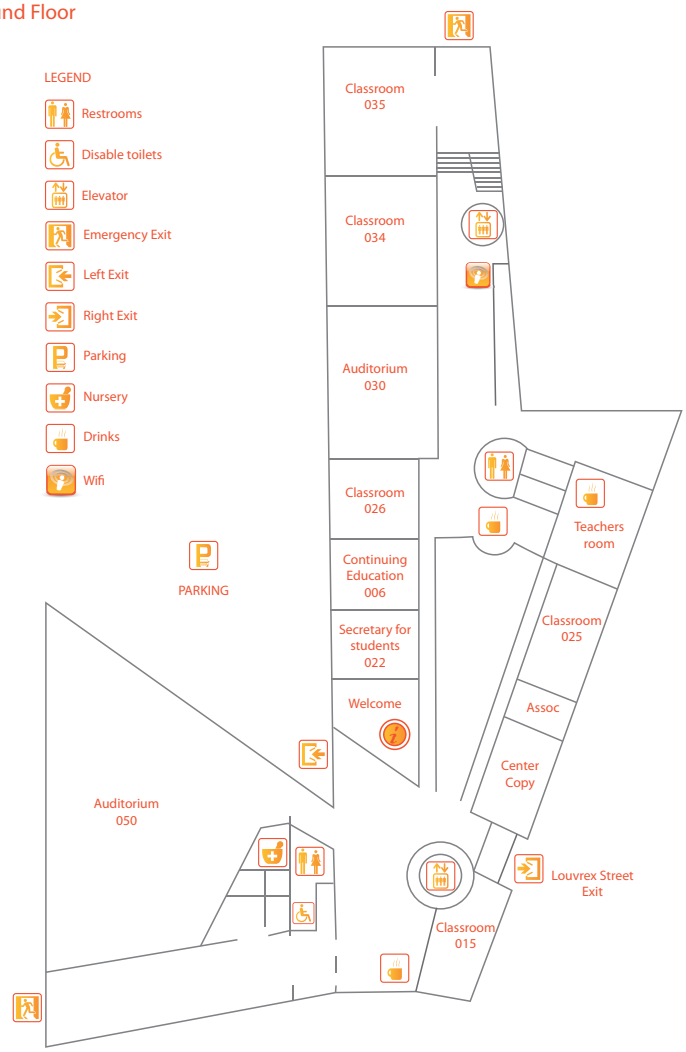
The registration desk will be located at the meeting venue, where you will be provided with your name badge and registration pack for the event. Registration will be open from 8:30 AM to 9:15 AM, on February 1, 2018.

The plenary sessions take place in room 030 and the parallel sessions in rooms 120, 126, 130, and 138. The conference rooms are equipped with an overhead projector, a video projector, and a computer. We suggest that you bring your own computer and/or transparencies as a backup. Each talk is allocated 20 minutes, including discussions. Please note that we are running on a very tight schedule. Therefore, it is essential that you limit your presentation to the assigned time. In order to allow inter-session hopping, we ask the chairpersons to follow the planned schedule as faithfully as possible. The two buffet lunches and coffee breaks take place in the lunchroom (room 135).

The conference dinner will take place in the Balcon de l'émulation Place du Vingt Août 16, 4000 Liège.

HEC ULg
Management School - University of Liège

Building N1
Ground Floor




HEC ULg
Management School - University of Liege
Building N1
1st Floor
101 - Financial Analysis
102 - Management Control Systems
103 - C.E.P.E
104 - Accenture Research Chair
105 - Finance
106 - Communication & External Affairs
107 - Finance
109 - Finance
111 - Financial Management
LEGEND
Restrooms
Disable toilets
Elevator
Emergency Exit
Left Exit
Right Exit
Parking
Nursery
Drinks
Classroom 115
Classroom 119
Classroom 120
111
109
107
106
105
104
103
102
101
Classroom 126
Classroom 130
Ml th Building ground floor
Library
Cafetaria
Lunchroom
Classroom 138

General schedule

Thursday, February 1

08:30-09:15	Welcome			
09:15-09:30	Opening session			
09:30-10:30	Plenary session I - Dominique Feillet *030* *Chairperson: Yasemin Arda*			
10:30-11:00	Coffee break			
11:00-12:20	Parallel sessions - TA			
	TA 1 Routing problems *Room 138*	TA 2 Emergency operations scheduling *Room 130*	TA 3 Algorithm design *Room 126*	TA 4 Multiple objectives *Room 120*
12:20-13:30	Lunch			
12:25-13:25	ORBEL board meeting *130*			
13:30-14:50	Parallel sessions - TB			
	TB 1 Integrated logistics *Room 138*	TB 2 Person transportation *Room 130*	TB 3 Continuous models *Room 126*	TB 4 Integer programming *Room 120*
14:50-15:20	Coffee break			
15:20-16:20	Parallel sessions - TC			
	TC 1 Material handling and warehousing 1 *Room 138*	TC 2 Operations management *Room 130*	TC 3 Matrix factorization *Room 126*	
16:30-17:10	Parallel sessions - TD			
	TD 1 Material handling and warehousing 2 *Room 138*	TD 2 Routing and local search *Room 130*	TD 3 Traffic management *Room 126*	TD 4 Pharmaceutical supply chains *Room 120*
17:15-18:15	ORBEL general assembly *130*			
18:30	Conference dinner *Balcon de l'émulation*			

Friday, February 2

09:30-10:30	Plenary session II - Martin Savelsbergh *030* *Chairperson: Yves Crama*			
10:30-10:50	Coffee break			
10:50-12:10	Parallel sessions - FA			
	FA 1 Optimization in health care *Room 138*	FA 2 Network design *Room 130*	FA 3 Local search methodology *Room 126*	FA 4 ORBEL Award *Room 120*
12:10-13:00	Lunch			
13:00-14:00	Parallel sessions - FB			
	FB 1 Production and inventory management *Room 138*	FB 2 Logistics 4.0 *Room 130*	FB 3 Data clustering *Room 126*	FB 4 Collective decision making *Room 120*
14:10-15:10	Parallel sessions - FC			
	FC 1 Sport scheduling *Room 138*	FC 2 Discrete choice modeling *Room 130*	FC 3 Data classification *Room 126*	
15:10-15:30	Coffee break			
15:30-16:30	Plenary session III - Michel Bierlaire *030* *Chairperson: Michaël Schyns*			
16:30-16:45	ORBEL award and closing session			
16:45-18:00	Closing cocktail			

Detailed program

A comparative study of branch-and-price algorithms for a vehicle routing problem with time windows and waiting time costs

Stefano Michelini
QuantOM, HEC Liège - Management School of the University of Liège
e-mail: stefano.michelini@uliege.be

Yasemin Arda
QuantOM, HEC Liège - Management School of the University of Liège
e-mail: yasemin.arda@uliege.be

Hande Küçükaydın
MEF University, Istanbul
e-mail: hande.kucukaydin@mef.edu.tr

Branch-and-price is one of the most commonly used methodologies for solving routing problems. In recent years, several studies have investigated advanced labeling algorithms to solve the related pricing problem, which is usually a variant of the elementary shortest path problem with resource constraints. Being able to solve this subproblem efficiently is crucial, since it is a major bottleneck for the performance of the branch-and-price procedure. Such labeling algorithms include relaxation methods like decremental state space relaxation [1], *ng*-route relaxation [2], and hybrids of these two [3], [4]. The cited relaxation methods mainly focus on how to treat efficiently the elementarity constraints, since these tend to make label domination difficult and consequently generate a higher number of labels, which translates to longer computational time and higher memory usage. In this study, we investigate the performances of these methods in a branch-and-price framework. The problem under consideration is a variant of the vehicle routing problem with time windows in which waiting times have a linear cost.

In order to perform the comparisons, we first parametrize several algorithmic components. Then, we search for good parameter configurations for each algorithm with `irace` [5]. The `irace` package is a tool for automated parameter tuning that generates and runs a very high number of configurations on a set of tuning instances, which in our case is a sample of the Gehring and Homberger instances [6] for the vehicle routing problem with time windows. `Irace` uses statistical tests to determine the best performing configurations. We select a single such configuration for each algorithm, called final configuration. Finally, we run all final configurations on the benchmark instances created by Solomon [7] and analyze the results with statistical tests. Our results show in particular that a class of hybrid algorithms with certain features based on *ng*-route relaxation outperforms all the others.

References

[1] Righini, G. and M. Salani. *New dynamic programming algorithms for the resource constrained elementary shortest path problem.* Networks, 2008.

[2] Baldacci, R., A. Mingozzi, and R. Roberti. *New route relaxation and pricing strategies for the vehicle routing problem.* Operations research, 2011.

[3] Martinelli, R., D. Pecin, and M. Poggi. *Efficient elementary and restricted non-elementary route pricing.* European Journal of Operational Research 2014.

[4] Dayarian, I, Crainic, T., Gendreau, M., and W. Rei. *A column generation approach for a multi-attribute vehicle routing problem.* European Journal of Operational Research, 2015.

[5] López-Ibáñez, M., Dubois-Lacoste, J, L. Pérez Cáceres, T. Stützle, and M. Birattari. *The irace package: Iterated Racing for Automatic Algorithm Configuration.* Operations Research Perspectives, 2016.

[6] Gehring, H., and J. Homberger. *A parallel two-phase metaheuristic for routing problems with time windows.* Asia-Pacific Journal of Operational Research, 2001.

[7] M. M. Solomon. *Algorithms for the Vehicle Routing and Scheduling Problems with Time Window Constraints* Operations Research, 1987.

A model for container collection in a local recycle network

Jens Van Engeland
KU Leuven, CEDON
e-mail: jens.vanengeland@kuleuven.be

Carolien Lavigne
KU Leuven, CEDON
e-mail: carolien.lavigne@kuleuven.be

As moving towards a more circular economy has become an increasingly important topic in waste management policy, the recovery of products, parts and materials will gain in importance. Additionally, the EU proximity principle related to waste management and emissions generated by transporting large amounts of end-of-life products, shift the attention to local recovery networks. The Flemish inter-municipal cooperation for domestic waste management Meetjesland (IVM) is currently examining the set-up of such local recovery networks. More specifically, container collection at Civic Amenity (CA) sites and local valorization of municipal waste is being investigated.

This paper proposes a Mixed Integer Linear Programming (MILP) model to optimize a container collection problem. Residents of a municipality unload their waste in these containers, which causes them to gradually fill up during a planning horizon. If a container is (almost) full, it has to be collected, replaced by an empty container and transported to a processor. Each day of the planning horizon, the different crews and vehicles stationed at a central depot leave to collect containers at the different sites. A vehicle can load up to two containers, which can differ in size. The crews start and end their day at the depot and work in shifts. During the collection process, the crew has to respect the opening hours of the CA sites. After unloading the container(s) at the processor, the crew can continue visiting CA sites or it can return to the depot to end its collection work for that day. Moreover, each crew should take a break per day, approximately halfway their shift.

The problem can be modelled as a variant to the PVRP-TW (Periodic Vehicle Routing Problem with Time Windows). In the container collection problem, a vehicle can maximum load two containers, hence only two subsequent site visits are possible. Therefore, we propose a solution method based on "trips" which are constructed a priori and contain all possible site combinations. Trips are the main building blocks of the model and are combined in order to make a feasible schedule for a crew on a day. Accordingly, in contrast to *operational* routing problems the model builds *tactical* waste collection schemes. The goal is to a find an optimal schedule describing which crew should visit which CA site on what day. The obtained solution should exhibit minimal route and truck costs.

The model is generic and applicable to different regions and material flows. For the case of IVM, the collection of bulky waste and polyvinyl chloride (PVC) is modelled. Bulky waste is collected either in small ($30\,m^3$) or large containers ($40\,m^3$) and transported to a nearby thermal treatment plant. PVC on the other hand is collected in smaller containers ($12\,m^3$) and is brought to a local recycler.

The MILP model was implemented in IBM ILOG CPLEX Optimization Studio. The solver obtained feasible collection schemes within a reasonable amount of computation time. As a result, useful suggestions concerning the crew schedule and number of vehicles needed can be suggested. Additional optimization runs will be performed to evaluate relevant policy and investment decisions such as alternative truck and container capacity investments and the effects of the possibility for working overtime.

Improving the Last Mile Logistics in a City Area by Changing Time Windows

Corrinne Luteyn
KU Leuven, Leuven Mobility Research Center - CIB
e-mail: corrinne.luteyn@kuleuven.be

Pieter Vansteenwegen
KU Leuven, Leuven Mobility Research Center - CIB
e-mail: pieter.vansteenwegen@kuleuven.be

In the last decades, the number of vehicles delivering products in city areas has increased enormously. Not only the shops and companies in the area require a delivery of their ordered goods, but also the residents of the city which ordered their new products online. All these customers require a fast delivery within a tight time window. These time windows for the customers are adding an extra complexity to the routing planning problem of the delivery companies. Besides that, these time windows will often lead to an increase in transportation costs. For instance, when customers, which are located closely to each other, have very different time windows, the delivery company has to visit almost the same location twice, at different moments. This leads to extra transportation costs for the delivery company.
In this research, we investigate the possible savings that can be obtained when a delivery company has the ability to discuss possible changes in time windows with their customers. By tuning the time windows of customers, which are closely located to each other, the delivery company can save transportation costs. However, if the company changes some of the time windows, it might loose some goodwill by its customers. In this research, this is modeled by a fixed cost for changing a time window. There is a given budget or a given number of time windows that can be changed.
Furthermore, in this research, it is assumed that a set of customers is given. All these customers have every day a non-negative demand and, therefore, require to be served. This non-negative demand of a customer varies from day to day. Since the capacity of the available vehicles is limited, the routes for the vehicles will also differ from day to day. Moreover, the required service time window for a given customer is the same every day. This means that changes in the time windows will influence the routes of all days in the considered group of days.
The daily routing planning problem of the delivery company can be modeled by a Vehicle Routing Problem with Time Windows (VRPTW)[1]. In this problem, a set of customers with a non-negative demand is given and each of these customers requires a visit of one of the available vehicles within a given time window. To the best of our knowledge, the possibility of changing the customers' time windows is not yet studied in the literature. We call this new variant of the VRPTW, the Vehicle Routing Problem with Changed Time Windows (VRPCTW). The objective of this

new problem is to determine the best fixed number of time windows changes, such that all customers are served on each day of the considered group of days at minimal total transportation cost.
To determine the best time windows to change, we present a Two-Phase Heuristic, which can be extended with a Testing Stage. The two phases of the heuristic are the construction phase and the analysis phase. In the construction phase, for each day in the group of days, routes for the vehicles are constructed. This means that a VRPTW is solved for each day. To construct the routes for the vehicles, a Variable Neighborhood Search (VNS) is used. The applied VNS is based on the basic VNS of Hansen and Mladenovic [2]. In this construction phase, the time windows of the customers are assumed to be soft time windows. This means that customers can be served outside their given time window at a certain penalty cost. This penalty cost consists of a fixed part, which is equal to the fixed cost for changing a time window, and a variable part, which is based on the deviation from the service time of the given time window.
In the analysis phase of the heuristic, the constructed routes are analyzed to determine the best time windows to change. This determination is based on the penalty costs which are incurred in the constructed routes of the previous phase. Customers where large penalty costs are incurred at several days, are good candidates for changes in their time windows. Due to the interference of time windows of different customers, the Two-Phase Heuristic can be extended with a testing stage. In this testing stage, promising combinations of changes in time windows are tested by applying the VNS of the construction phase to the set of customers with adjusted time windows.
The presented solution approach is tested on a set of new benchmark instances based on the Solomon instances for the VRPTW [3]. Preliminary results show that by changing only a small number of time windows, the total transportation costs for the vehicles during the considered group of days can be decreased by around 3%.

Acknowledgments

This research was funded by a Ph.D. grant of the Agency for Innovation by Science and Technology in Flanders (IWT).

References

[1] G. Desaulniers, O.B.G. Madsen and S. Ropke. The vehicle routing problem with time windows. In *Vehicle routing: Problems, methods, and applications. Second Edition*. MOS-SIAM Series on Optimization (SIAM, Philadelphia), 119–159, 2014.

[2] P. Hansen and N. Mladenovic. Variable neighborhood search: Principles and applications. *European Journal of Operational Research*. 130(3):449–467, 2001.

[3] M. M. Solomon. Algorithms for the vehicle routing and scheduling problems with time window constraints. *Operations Research*. 35(2):254–265, 1987.

A buffering-strategy-based solution method for the dynamic pickup and delivery problem with time windows

Farzaneh Karami[a,*], Wim Vancroonenburg[a,b], Greet Vanden Berghe[a]

[a]KU Leuven, Department of Computer Science, CODeS & imec

[b]Research Foundation Flanders - FWO Vlaanderen

[*]E-mail: farzaneh.karami@kuleuven.be

Pickup and delivery problems with time windows (PDPTWs) are ubiquitous throughout the logistics sector, with them involving the picking-up of items from one location and their delivery to another. Requests require handling before a certain time and thus there are two so-called time windows associated with each request: a pickup time window and a delivery time window.

Several applications such as resource allocation, scheduling, robotics and vehicle routing present situations wherein PDPTWs must be solved without complete knowledge of the associated requests in advance, but where they are announced as time goes by. Such problems are called *dynamic* PDPTWs.

Dynamic PDPTW instances can be characterized by two parameters: their dynamism and urgency [1]. Under van Lon's definition, dynamism relates to the continuity of request arrivals, whereas urgency relates to the minimum time available to handle a request. Optimization decisions are derived from the data available at a particular moment in time and are likely to deteriorate in quality as new information becomes available. This gradual release of information over time therefore necessitates the decision-making process to be repeated at certain time intervals.

This paper proposes a buffering-strategy-based optimization approach to the *dynamic* PDPTW, which seeks to group sets of dynamic arrival requests, allocate them to available vehicles, and minimize requests' tardiness, vehicles' travel time and vehicles' overtime as objective while respecting several requests' and multiple vehicles' constraints iteratively. Thus, the decision-making process repetitively re-optimizes wherein the input only consists of a small, but certain part of the known future requests at that time. This dynamic optimization approach employs a cheapest insertion procedure and a local search algorithm as two heuristic-based *scheduling policies*. The objective value of servicing all requests, obtained at the end of the planning horizon, is called the solution cost.

The proposed approach is assessed in a computational experiment on existing instances [1] by comparing the solution costs against the best solutions found by Mixed Integer Linear Programming (MILP) model. This dataset considers different configurations of dynamism degrees and urgency levels of request arrivals. It then enables evaluating the proposed approaches in distinct situations.

The results illustrate how the performance of the proposed approach is impacted by

the urgency level, the dynamism degree and the re-optimization frequency. Results indicate that any dynamism increment will positively decrease the solution costs, whereas urgency is negatively correlated with solution costs. In addition, it is noteworthy that solution quality (relative gap to the best found MILP solution for which all information is known in advance) is only slightly affected by changing dynamism and urgency characteristics of request arrivals.

Acknowledgements

W. Vancroonenburg is a postdoctoral research fellow at Research Foundation Flanders - FWO Vlaanderen. Editorial consultation provided by Luke Connolly (KU Leuven).

References

[1] van Lon, R., Ferrante, E., Turgut, A. E., Wenseleers, T., Vanden Berghe, G., Tom Holvoet, T., *Measures of dynamism and urgency in logistics*, European Journal of Operational Research, 253 (3), (2016) pp. 614–624.

Input data quality assessment in the context of emergency department simulations

L. Vanbrabant [a, b], N. Martin[a], K. Ramaekers[a], and K. Braekers[a]

(a) Hasselt University, (b) Research Foundation Flanders (FWO)

e-mail: lien.vanbrabant@uhasselt.be, niels.martin@uhasselt.be, katrien.ramaekers@uhasselt.be, kris.braekers@uhasselt.be

Emergency departments (EDs) are confronted with the problem of crowding. Crowding occurs when the demand for emergency services exceeds the available resources, and is mainly caused by an increase in patient visits without a proportionate capacity expansion. It undermines the ability of an ED to provide timely and qualitative care to patients. Therefore, hospitals aim to relieve the pressure on EDs by improving their operational efficiency [3, 4].

Simulation techniques are frequently used to model, analyse and optimise ED operations. The characteristics of simulation make this technique highly suitable to investigate the complex and stochastic environment of an ED [2, 4]. Simulation results can be valuable to hospital managers, leading to the successful implementation of operational improvements to reduce the negative effects of crowding [3]. Nevertheless, simulation is only advantageous if the model closely resembles the real system. The garbage in, garbage out principle states that the results of a simulation analysis are only as reliable as the model and its inputs. The basis of any simulation study is the construction of a realistic simulation model, and accurate data on the real system are a prerequisite within this regard [1, 2]. In general, one third of the total simulation modelling process time should consist of input data management and analysis. However, most ED simulation studies do not consider input data quality, or the quality assessment process is not discussed. This is remarkable, since electronic health records (EHRs) are the main source of input data in ED simulations and despite their wide availability, the quality and suitability of this data for research purposes can be questionable [5].

In order to obtain high quality input data, the identification and assessment of data quality problems is an important initial step in the simulation modelling process. Several classification frameworks for data quality problems have been provided in literature, both general and healthcare specific, but they are not adjusted to the context of reusing electronically recorded data for operations research purposes. In addition to the frameworks, methods for data quality assessment in a simulation context are proposed in literature. However, these methods are not frequently used and they are not capable of objectively and unambiguously assessing data quality in a standardised way [1, 2, 5].

In our research, a comprehensive data quality framework is developed that focuses on data quality problems present in EHR data used as input for ED simulations. The main classification of the framework divides data quality problems into missing

and not-missing data. These categories are further subdivided until fourteen specific data quality problems are identified. For all these problems, assessment techniques to investigate their presence in an objective way are described. In addition, the detection and quantification of several data quality problems is automated by the development of quality assessment functions in the R programming language. The functions are defined in a generic way, with each function having function-specific arguments. These arguments enable to adjust the function to the ED and dataset under study.

Input data analysis is a time-consuming task, but the framework and implemented assessment techniques standardise and facilitate the process of identifying and quantifying potential data quality problems. The application of the implemented assessment functions to a real-life case study demonstrates the importance and benefits of evaluating input data quality. The results confirm that EHR data may contain many quality issues and provide useful insights on the specific problems that are present and their extent. This information can be used to clean the dataset, or to adjust the modelling approach or level of detail taken into account in the simulation model.

The next step, after data quality assessment and data cleaning, is the construction of a simulation model. By investigating input data before use, the introduction of bias into the simulation model is limited. As a result, the simulation model better reflects reality, which in turn implies that effective solutions to the crowding problem can be identified by means of what-if analysis or simulation-optimisation techniques. In summary, the complete simulation analysis depends on the quality of input data, emphasising the fundamental nature of data quality assessment. A formal data quality assessment process enhances the reliability of simulation results.

References

[1] Kelton, W.D., Sadowski, R.P., Zupick, N.B. (2015). Chapter 4 Modeling Basic Operations and Inputs. In *Simulation with Arena*. McGraw-Hill Education, 6, 121-205.

[2] Liu, Z., Rexachs, D., Epelde, F., Luque, E. (2017). A simulation and optimization based method for calibrating agent-based emergency department models under data scarcity. *Computers & Industrial Engineering*, 103, 300-309.

[3] Paul, S.A., Reddy, M.C., DeFlitch, C.J. (2010). A Systematic Review of Simulation Studies Investigating Emergency Department Overcrowding. *Simulation*, 86 (8-9), 559-571.

[4] Saghafian, S., Austin, G., Traub, S.J. (2015). Operations research/management contributions to emergency department patient flow optimization: Review and research prospects. *IIE Transactions on Healthcare Systems Engineering*, 5(2), 101-123.

[5] Skoogh, A., Perera, T., Johansson, B. (2012). Input data management in simulation - Industrial practices and future trends. *Simulation Modelling Practice and Theory*, 29, 181-192.

Expanding the SIMEDIS simulator for studying the medical response in disaster scenarios

S. Koghee
Royal Military Academy, Department of Mathematics
e-mail: selma.koghee@mil.be

F. Van Utterbeeck
Royal Military Academy, Department of Mathematics
e-mail: filip.vanutterbeeck@rma.ac.be

M. Debacker
Vrije Universiteit Brussel, Research Group on Emergency and Disaster Medicine

I. Hubloue
Vrije Universiteit Brussel, Research Group on Emergency and Disaster Medicine

E. Dhondt
Belgian Armed Forces, COMOPSMED, Medical Component

The SIMEDIS (Simulation for the assessment and optimisation of Medical Disaster management) simulator has been developed to study the pre-hospital medical response in disaster scenarios and to provide evidence for best-practice recommendations. The advantage of using simulations over real-life practise runs is that the medical response can be tested much more extensively. In the simulation runs, many variations of the parameters are covered for a specific scenario. The results from the simulation are used to determine under which conditions the number of diseased victims is minimized, since the aim of the medical response is to save as many victims as possible and to have their quality of life as highly as possible.
The process that is studied is the pre-hospital medical response according to the Belgian regulations. This starts with the self-evacuation of the uninjured and mildly injured victims, who will be treated separately from the severely injured victims. The severely injured victims are rescued by fire fighters and brought to the casualty collection point (CCP). Here, the victims are assigned a priority, based on their current health state, by a doctor or a nurse. The order in which the victims are treated and transported is based on these priorities. The next steps depend on the chosen policy, which can be Scoop and Run or Stay and Play. Scoop and Run is aimed to bring the victims to a hospital as soon as possible. On the other hand, the strategy of Stay and Play is to first stabilize the victims at the scene before transporting them to a hospital. In the case the Scoop-and-Run policy is chosen, the victims are evacuated directly from the CCP to a hospital. The first treatment or stabilization will be applied during transport by the supervising medical personnel. In case the Stay-and-Play policy is followed, the victims are first transported to the forward medical post (FMP), were the victims receive the first medical treatment

before they are transported to a hospital.
The simulator is a discrete-even simulator, where the involved entities are the victims, medical personnel, ambulances and hospitals. The parameters under variation are a combination of variables related to the circumstances, for example the number of victims and the initial hospital capacity, and others that correspond to the decisions made by the medical director who is in charge of the medical resources during the rescue operation, such as the policy and the supervision level during transport to the hospital.
This model has fist been developed in Arena, a computer program to simulate processes, and was used to study an aeroplane crash at Zaventem Airport. Now the model has been re-created in Julia, a high-level open source programming language, improved with extra features, and generalized to be applicable to different scenarios. A first test was to recreate the Zaventem scenario. The obtained results are in line with previous research for the same number of victims ([1]), even when random variation in the event times is included. Furthermore, for victim numbers which are 25% higher or lower the qualitative effects of the parameters remain the same to a large extend. Lastly, registration of the patient waiting times and the resource effectiveness help to suggest possible improvements to reduce the death toll even further. Nevertheless, such suggestions should be examined first with new simulations.

References

[1] Debacker, M., Van Utterbeeck, F., Ullrich, C., Dhondt, E., and Hubloue, I., 'SIMEDIS: a discrete event simulation model for testing responses to mass causality Incidents.' In Journal of Medical Systems **40**, 273 (2016).

Prioritization between boarding patients and patients currently waiting or under treatment in the emergency department

Kim De Boeck
KU Leuven, Research Centre for Operations Management
e-mail: kim.deboeck@kuleuven.be

Raïsa Carmen
KU Leuven, Research Centre for Operations Management
e-mail: raisa.carmen@kuleuven.be

Nico Vandaele
KU Leuven, Research Centre for Operations Management
e-mail: nico.vandaele@kuleuven.be

A serious issue often encountered in hospital emergency departments (EDs) is crowding. Crowding occurs when the demand for emergency resources exceeds the resources available in the ED ([2]). This can cause several negative effects including longer waiting time and length of stay (LOS), lost hospital revenue, increased number of ambulances that cannot immediately unload the patient at the ED, more patients that leave without being seen (LWBS) and even higher mortality rate ([1]). One of the main contributors to crowding is inpatient boarding which results from the inability to transfer patients from the ED to the inpatient ward (IW) ([2]). Although the importance of inpatient boarding is stated in many operations research (OR) and operations management (OM) literature, very little research focuses on this topic (Saghafian, Austin & Traub, 2015). Furthermore, when boarding is included in the analysis, it is often assumed that boarding patients occupy beds in the ED, but do not require any other resources. However, in reality, ED physicians are responsible for these patients as long as they are not transferred to the IW.

Our paper investigates the effect of this simplification by incorporating the additional check-ups required by boarding patients into the analyzed system. Adding these required check-ups increases the workload of the physicians, but also raises an additional question. That is, after finishing a treatment, the physician needs to decide whether he or she will see a boarding patient or one of the other patients currently waiting or under treatment in the ED (which we refer to as test patients). Since the treatment characteristics of these two kinds of patients differ, this decision has an important impact on system performance.

The main contribution of our paper is the examination of different control policies for the physicians in an ED. Using discrete-event simulation, three static control policies (first-come, first-served and always prioritizing either boarding or test patients) and two dynamic control policies (using threshold values and accumulating priorities) are studied.

When only looking at operational system performance measures like waiting times and utilization rates, the recommended control policy is simple and straightforward; always prioritize test patients. However, in an emergency department setting, we also need to consider health-related performance measures; physicians cannot disadvantage one type of patients in favour of operational system performance. The most important health-related performance measure taken into account in our paper is the boarding patients' risk percentage which we define as the percentage of boarding patients in the physicians' queue that has to wait longer than 15 minutes before seeing a physician.
The result is a trade-off between the two types of performance measures; operational system performance can only be improved at the expense of boarding patients' risk percentage. This trade-off leads us to conclude that defining a control policy for the ED physicians is far from trivial and that no single optimal control policy exists. Rather, the optimal control policy depends on the target boarding patients' risk percentage the hospital is willing to accept.
Another important finding is that it is not always necessary to use a complicated dynamic control policy. Indeed, a simple static first-come, first-served policy turns out to perform extremely well for a wide range of imposed targets on boarding patients' risk percentage. Only when an extremely high or low value of target boarding patients' risk percentage is desired, it is worthwhile to put effort into the implementation of a dynamic control policy.

Acknowledgments

We gratefully acknowledge financial support from the GlaxoSmithKline Research Chair on Re-Design of Healthcare Supply Chains in Developing Countries to increase Access to Medicines.

References

[1] Hoot, N. R., & Aronsky, D. (2008). Systematic Review of Emergency Department Crowding: Causes, Effects, and Solutions. Annals of Emergency Medicine, 52(2), 126-136.

[2] Moskop, J.C., Sklar, D.P., Geiderman, J.M., Schears, R.M., & Bookman, K.J. (2009). Emergency Department Crowding, Part 1 - Concept, Causes, and Moral Consequences. Annals of Emergency Medicine, 53(5), 605-611.

[3] Saghafian, S., Austin, G., & Traub, S.J. (2015). Operations research/management contributions to emergency department patient flow optimization: review and research prospects. IIE Transactions on Healthcare Systems Engineering, 5(2), 101-123.

Protecting operating room schedules against emergency surgeries: outlook on new stochastic optimization models

M. Vandenberghe, S. De Vuyst, E.H. Aghezzaf
Ghent University, Departement of Industrial Systems Engineering and Product Design
e-mail: Mathieu.Vandenberghe;Stijn.DeVuyst;ElHoussaine.Aghezzaf@ugent.be

H. Bruneel
Ghent University, Department of Telecommunications and Information Processing
e-mail: Herwig.Bruneel@ugent.be

For most hospitals, operating rooms constitute both a major source of revenue and their largest cost category [1]. Problems in the operating room also trickle down to other departments, which creates a strong desire to increase efficiency while minimizing problems. On the scheduling front, one particular problem is that of emergency patients arriving and requiring rapid attention. Because hospitals strive to keep operating room occupancy high, free rooms are typically not available: thus the earliest possibility for the emergency to 'break into' the schedule is when any room finishes its surgery. This leads to a novel characteristic for a good schedule: minimizing the maximum time between these potential 'break-in-moments ', a problem defined in a deterministic context by [2].

As surgery durations are often highly uncertain, we have recently focused on the Stochastic Break-In-Moment (SBIM) Problem. We have proposed solution methods [4] for the more general scheduling problem where the surgery durations are stochastic, but emergency arrivals are essentially unknown. The problem involves M surgeries with independent stochastic durations $\mathbf{P} = (P_1, \ldots, P_M)$. The surgeries are pre-assigned to K operating rooms, and a global schedule $\pi = (\pi^1, \ldots, \pi^K) \in \Pi$ determines their order within each room. Any particular schedule π results in a set of completion times $C_i, \quad i \in \mathcal{I} = \{1, \ldots, M\}$. We then define the BIM sequence B_i as the completion times C_i in non-decreasing order. The intervals between these BIMs form the Break-In-Intervals (BII) $S_i = B_i - B_{i-1}$, $B_0 = 0$, $i \in \mathcal{I}$. The objective is to find a schedule π that minimizes the expected value of the largest BII length:

$$v^* = \min_{\pi \in \Pi} g(\pi)\,, \qquad \text{with} \qquad g(\pi) = \mathrm{E}[\max_{i \in \mathcal{I}} S_i(\pi, \mathbf{P})]\,, \tag{1}$$

and where the expectation is taken over the joint distribution of $\mathbf{P}$. Here the best strategy was always to spread the break-in-moments as uniformly as possible across the session interval, because no assumptions were made regarding the emergency arrivals. The optimal value v^* in (1) is the expected waiting time of an emergency if it should arrive at the worst possible time during the session.

Our goal now is to step away from only considering the absolutely worst-case situation and explicitly take knowledge about the arrival pattern of emergencies into

account. We therefore propose a reformulation of (1) where the number of emergencies and their arrival times have a known distribution. This new problem, referred to as the Stochastic Emergencies variant (SE), calls for new objectives as well as a series of possible design decisions. A relevant instance is where we assume a Poisson distributed amount J of emergencies of which the arrival times $\mathbf{A} = (A_1, \ldots, A_J)$ and surgery durations $\mathbf{P}_e$ are independent with known distributions. The new objective is to find the schedule that minimizes the total time from each emergency arrival until the next BIM:

$$v^* = \min_{\pi \in \Pi} g(\pi)\,, \qquad \text{with} \qquad g(\pi) = \mathrm{E}[\sum_{j \in \mathcal{J}} S_j(\pi, \mathbf{P}, \mathbf{P}_e, \mathbf{A})]\,, \tag{2}$$

where $\mathcal{J} = \{1, \ldots, J\}$ and $g(\pi)$ now only refers to the J waiting intervals of the emergencies.
Because there are now various stochastic components (the surgery durations, the time of emergency arrivals, as well as the number of emergency arrivals), it becomes much more difficult to accurately assess (let alone optimize) the objective value of the schedule. This calls for a larger scenario set to optimize, which in turns requires more sophisticated solution methods.
We will present the current state and challenge of this SE variant, and show the potential of this stochastic extension compared to the regular SBIM (in terms of its VSS). We also propose solution methods, in particular an implementation of Sample Average Approximation proposed in [3], specifically for problems where adding extra scenarios is computationally challenging.

References

[1] Macario, A., Vitez, T.S., Dunn B., and McDonald T. Where are the costs in perioperative care? Analysis of hospital costs and charges for inpatient surgical care. *Anesthesiology*,. 83(6):1138-44, 1995.

[2] Van Essen, J.T., Hans, E.W., Hurink, J.L., and Oversberg, A. Minimizing the waiting time for emergency surgery *Operations Research for Health Care* 1 (2), 34-44, 2011.

[3] Kleywegt, A., Shapiro, A., and Homem-de-Mello T. The Sample Average Approximation Method for Stochastic Discrete Optimization *SIAM J. Optimization*, 12(2), 479-502, 2002.

[4] Vandenberghe, M., De Vuyst, S., Aghezzaf, E-H., and Bruneel, H. Increasing the Robustness of Surgery Schedules against Emergency Break-ins *European Journal of Europeans Research*, submitted, under review.

Towards Algorithm Selection for the Generalised Assignment Problem

H. Degroote and P. De Causmaecker
KU Leuven Department of Computer Science, CODeS & Imec-ITEC
e-mail: hans.degroote@kuleuven.be

J.L. González-Velarde
Tecnologico de Monterrey, Escuela de Ingeniería y Ciencias

Many problems in Operations Research (OR) are NP-complete. Nevertheless, many practical problems are solved efficiently by employing powerful heuristics. Such heuristics work well in some cases, but not in others. The goal of algorithm selection is to use machine learning techniques to predict which algorithm is best suited for solving a given instance, based on a set of cheaply-computable features characterising the instance and based on performance data of the algorithms on a set of training instances. Algorithm selection has been mainly applied to decision problems in a competition setting. This extended abstract discusses some of the challenges that arise when applying algorithm selection to an OR problem, outside of a competition setting. The generalised assignment problem (GAP) is used as a case study.

In the GAP, each of n jobs should be assigned to exactly one of m agents, where each agent has a maximal resource capacity. Assigning job j to agent i consumes r_{ij} of the agent's resources and incurs a cost of $c_{i,j}$. The goal of the generalised assignment problem is to find a feasible assignment of jobs to agents that minimises the total cost.

Many powerful algorithms have been developed for the GAP, often using metaheuristics. No algorithm seems to outperform all other algorithms on all instances, which indicates a potential for algorithm selection. Furthermore, in many applications the GAP should be solved quickly and repeatedly, either because it directly models a problem at the operational level (f.e. multi processor task scheduling) or because it is solved repeatedly as a sub problem to a more complicated problem (f.e. vehicle routing applications). Since algorithm selection is most applicable in a setting with limited computational resources, it is a good fit for the GAP.

To perform algorithm selection, first algorithms must be collected. Most algorithms published within OR are not open source, which complicates this step. One solution is to implement the algorithms based on the description in their respective publications, but this is prone to misinterpretation, and is a lot of work regardless. Preferable is to request the source code from the creators. However, in this case it is typically required that the source code cannot be distributed further (otherwise it would have been made open source). This means that an algorithm system itself cannot be made open source unless all of the algorithms it uses are also open source. Even sharing privately with other people is complicated, as this requires permission from the owners of each algorithm. This issue is hard to solve unless OR algorithms more often become open source.

Four algorithms were obtained for the GAP (out of over ten requested by sending e-mails to the contact author), as well as an integer programming formulation that can be read by CPLEX. One of the four algorithms could not be made to work, resulting in a total of three heuristic algorithms: two tabu searches with ejection chains [3] [2] and a tabu search with path relinking [1], and one exact solver: CPLEX.
A benchmark library of GAP instances exists, consisting of 56 instances of five types (A,B,C,D and E), differentiated by the way in which the resource capacities and the cost and resource matrices are generated. These are used in literature to evaluate new algorithms against. Furthermore, an additional 230 instances of types C, D and E were obtained (types A and B are considered too easy to be interesting in literature). Instance size is measured in amount of agents times amount of jobs, and it ranges from 5×100 to 80×1600.
Before applying algorithm selection must be decided how to measure performance. In a competition setting this is a given, but for OR this is application dependent. The best way of precisely defining a performance measure is to consider the actual application and how the quality of the solution influences decision making and accompanying revenue, but this is extremely low level. This issue illustrates the difficulty of defining a general algorithm selection system for the GAP, or indeed any OR problem: relative algorithm quality depends on the performance measure, and the performance measure depends on the specific application. In the experiment here reported, performance is measured as the objective value of the best solution found within 5 seconds, with a value of infinity if none was found.
Preliminary experiments show that algorithm selection indeed has potential for the GAP: on 124 out of the 286 instances (43.36%) the best algorithm was not the single best solver (the algorithm with best performance on average). The single best solver was the ejection chain based tabu search of [2], but both CPLEX and the path relinking based tabu search regularly outperformed it.
Future work is to develop features for the GAP and to apply algorithm selection, to verify that the potential is indeed met.
This work is supported by the Belgian Science Policy Office (BELSPO) in the Interuniversity Attraction Pole COMEX (http://comex.ulb.ac.be).

References

[1] M. Yagiura, T. Ibaraki and F. Glover.A path relinking approach with ejection chains for the generalized assignment problem. European journal of operational research, 169(2): 548–569, 2006

[2] M. Yagiura, T. Ibaraki and F. Glover. An ejection chain approach for the generalized assignment problem. INFORMS Journal on Computing, 16(2), 133-151, 2004

[3] M. Laguna, J.P. Kelly, J.L. González-Velarde and F. Glover. Tabu search for the multilevel generalized assignment problem. European journal of operational research, 82(1), 176-189, 1995

Searching the Design Space of Adaptive Evolutionary Algorithms

Ayman Srour
KU Leuven, CODeS,
e-mail: aymanih.srour@kuleuven.be

Patrick De Causmaecker
KU Leuven, CODeS,
e-mail: patrick.decausmaecker@kuleuven.be

The success of Evolutionary Algorithms (EAs) in solving many optimization problems is due to several adjustable parameter configurations which provides EAs a flexible problem adaptation. However, this flexibility also encounters with the difficulty of defining optimal parameter configuration. The problem of finding optimal parameter configuration for an algorithm is a nontrivial task. Adaptive parameter control (APC) [1] have been intensively used in literature to find the optimal parameters configuration automatically and in an online manner and during the algorithm execution by considering feedback from an EA run. In the context of EA, more sophisticated literature reviews, such as [2, 3], have focused on studying the design structure of the existing APC methods in Adaptive Evolutionary Algorithms (AdEA) and tried to decompose the APC process into several elements or components. The components that play a role to generalize the design strategies of the most existing APC methods

Aleti et al. [2] proposed a generic model of the state-of-the-art APC methods based on a comprehensive review of the literature. The authors distinguish between the optimization process and the control process of the Adaptive Evolutionary Algorithms (AdEA). The control process is further divided into four components, namely, feedback collection, effect assessment, quality attribution, and selection. Each component represents a stage of the parameter control accompanying with several possible adaptive strategies for each stage. Considering this model, the architecture of every AdEA can be mapped into the four components. Thus, a complete design on any AdEA should have at least one strategy for each component, and each strategy can be implemented by using one of many predefined rules or methods. In fact, we can utilize the model to facilitate the design of any new AdEA, but it is still time-consuming to perform this task manually. As developing a robust AdEA leads us to consider several issues such as the nature of a problem, the difficulty of target instances and the decision which APC strategy for a specific parameter in AdEA should be adopted, which is, of course, affecting the performance of the AdEA. All of these issues also make the design of AdEA a difficult task. The difficulty is because several APC strategies exist today and deciding the optimal among many alternatives, for a specific problem or even for a specific problem instance, is infeasible by considering the human design alone. Using an automatic algorithm design is an

affordable alternative approach that not only helps to reduce the difficulty of selecting the best among several algorithm components but also can help to generate a novel algorithmic design which is in many cases superior to the standard algorithms. More recently, GE has been used to evolve the design of EAs for solving Royal Road Functions [4], Integration and Test Order problems [5]. In this work, we propose a Grammatical Evolution framework to evolve the design of AdEA. GE automates the design of AdEA by defining a grammar guiding the selection of an APC strategy for different EA parameters and defining the corresponding implementations. The main aim of the automatic design of AdEA is to surrogate the human design leading to significant performance improvement. The results of the experiments using several traveling salesman problem instances confirmed that GE-AdEA has a potential to improve the adaptive parameter control strategy for crossover and mutation operators and their rates. The results also revealed that up to 20 % of the generated AdEAs outperform a tuned EA. Full computational results are available in [6].

References

[1] Eiben, Ágoston E., Robert Hinterding, and Zbigniew Michalewicz. Parameter control in evolutionary algorithms. *IEEE Transactions on evolutionary computation* 3.2 (1999): 124-141.

[2] Aleti, A. and I. Moser, A Systematic Literature Review of Adaptive Parameter Control Methods for Evolutionary Algorithms. *ACM Computing Surveys (CSUR)*, 2016. 49(3): p. 56.

[3] Corriveau, Guillaume, et al. Bayesian network as an adaptive parameter setting approach for genetic algorithms. *Complex & Intelligent Systems* 2.1 (2016): 1-22.

[4] Lourenço, N., F. Pereira, and E. Costa. Evolving evolutionary algorithms. in *Proceedings of the 14th annual conference companion on Genetic and evolutionary computation.* 2012. ACM.

[5] Mariani, T., et al. A grammatical evolution hyper-heuristic for the integration and test order problem. in *Genetic and Evolutionary Computation Conference.* 2016.

[6] A. Srour and De Causmaecker P. Evolving Adaptive Evolutionary Algorithms in *Proceedings of the MISTA* - Multidisciplinary International Scheduling Conference, Kuala Lumpur, 2017.

Analysis of algorithm components and parameters using surrogate modelling: some case studies

Nguyen Dang
KU Leuven, Department of Computer Science, CODeS & imec-ITEC
e-mail: nguyenthithanh.dang@kuleuven.be

Patrick De Causmaecker
KU Leuven, Department of Computer Science, CODeS & imec-ITEC
e-mail: patrick.decausmaecker@kuleuven.be

Modern well-performing optimisation algorithms are usually highly parametrised. Algorithm developers are often interested in the questions of how different algorithm components and parameters influence performance of algorithms. Insights into how an algorithm works may produce useful knowledge that can be transferred back into the algorithm development procedure in an iterative manner, leading to higher performing algorithms. The knowledge may also provide suggestions for better algorithm design in other application domains or similar contexts [1].
We focus on a group of recently proposed algorithm parameter analysis techniques that utilise surrogate modelling [2] [3], namely forward selection [4], fANOVA [5] and ablation analysis with surrogates [6]. Given a performance dataset of an algorithm, these methods build random-forest regression models to predict the algorithm performance over the whole configuration space. The models are then used to analyse the influence of the algorithm parameters on the predicted performance. These methods allow all types of parameters, including the categorical one, and have been tested on a number of cases with high-dimensional algorithm configuration space [4] [5] [6]. Each one addresses a different analysis aspect. The question when applying these methods is what kind of information we want to gain for a deeper understanding of how the analysed algorithm works. It might not always be straightforward for an algorithm developer to interpret analysis results and decide what to use for improving his/her algorithm. In this work, we illustrate the applications of these methods on a number of case studies in optimisation algorithms and discuss the advantages of combining them for thorough insights into the algorithms under study.
The implementation of all the three analyse methods used in this work is from the PIMP package (https://github.com/automl/ParameterImportance), provided by the *ML4AAD Group* [1].

Acknowledgements

This work is funded by COMEX (Project P7/36), a BELSPO/IAP Programme.

[1] http://www.ml4aad.org/

References

[1] Chris Fawcett and Holger H Hoos. Analysing differences between algorithm configurations through ablation. *Journal of Heuristics*, 22(4):431–458, 2016.

[2] Kevin Leyton-Brown, Eugene Nudelman, and Yoav Shoham. Empirical hardness models: Methodology and a case study on combinatorial auctions. *Journal of the ACM (JACM)*, 56(4):22, 2009.

[3] Frank Hutter, Lin Xu, Holger H Hoos, and Kevin Leyton-Brown. Algorithm runtime prediction: Methods & evaluation. *Artificial Intelligence*, 206:79–111, 2014.

[4] Frank Hutter, Holger H Hoos, and Kevin Leyton-Brown. Identifying key algorithm parameters and instance features using forward selection. In *International Conference on Learning and Intelligent Optimization*, pages 364–381. Springer, 2013.

[5] Frank Hutter, Holger Hoos, and Kevin Leyton-Brown. An efficient approach for assessing hyperparameter importance. In *International Conference on Machine Learning*, pages 754–762, 2014.

[6] André Biedenkapp, Marius Thomas Lindauer, Katharina Eggensperger, Frank Hutter, Chris Fawcett, and Holger H Hoos. Efficient parameter importance analysis via ablation with surrogates. In *AAAI*, pages 773–779, 2017.

Automatic configuration of metaheuristics for the Q3AP

I. Ait Abderrahim, L. Loukil
University of Oran 1 Ahmed Ben Bella, Algeria,
e-mails: aitabderrahim_i@hotmail.fr, loukil.lakhdar@univ-oran1.dz

T. Stützle
Université libre de Bruxelles (ULB)
e-mail: stuetzle@ulb.ac.be

The quadratic three-dimentional assignment problem (Q3AP) is a generalization of the standard quadratic assignment problem (QAP). Q3AP involves the simultaneous and independent assignment of two different entities to a common location. It is an NP-hard permutation problem that has gained interest for its application in the hybrid automatic repeat request protocol (Hybrid ARQ) in wireless communication systems. When applied to Hybrid ARQ protocol, solving the Q3AP consists in finding an optimal symbol mapping over two vectors (one vector for the original transmission and the second for the repeat transmission) so as to minimize errors in the received transmission of data. High-quality solutions to the Q3AP can significantly increase throughput and reduce the cost for providing reliable digital transmission over noisy fading channels.

The Q3AP was introduced by Pierskalla [1] in 1967 in a technical memorandum. Thirty six years later, the Q3AP was re-discovered by Hahn et al. [2] while working on a problem arising in data transmission system design. Several approaches have been proposed to solve this problem, which can be classified into three classes: the first class is exact methods. The team of Hahn [2], which is one of the first to work on this problem, implemented a Branch and Bound algorithm(B&B). Later, Galea et al. [3] have proposed a parallel B&B and Mezmaz et al. [4] proposed a grid based parallel B&B algorithm. Another recent work due to Mittelmann [5] uses a branch and cut method that was able to solve optimally a Q3AP instance of size 16.

The second class are metaheuristic algorithms where we find the work of Hahn et al. [2] who adapted four approximate solution methods (metaheuristics) that have achieved widespread success in solving QAP for solvig the Q3AP. Iterated Local Search is considered in [2] as globally the best performing metaheuristic among the four implemented ones to reach the optimal solution, Another contribution was presented by Luong et al. [6], where they proposed a GPU-based parallel iterative tabu search.

The third class are hybrid methods to solve the Q3AP. Loukil et al. [7] have proposed a two-level parallel hybrid genetic algorithm (GA), P. Lipinski [8] suggests a hybrid algorithms combining an evolutionary algorithm (EA) with a local search (LS) method and M. Mehdi [9] proposed a new cooperative hybrid schemes combining GA and B&B. This work is a modified Tabu Search (TS) proposed in [6].

In our work, we build on the previous research results on the Q3AP and develop a new hybrid method for the Q3AP using automated design ideas. In particular, in this presentation we focus on an iterated local search algorithm that exploits a tabu search as the local search for solving the Q3AP (ILS-TS). As a first step, we extend the original tabu search algorithm to implement additional alternative choices when compared to a first prototype implementation. Next, we configure the ILS-TS algorithm problem using the irace[1] (Iterated Race for Automatic Algorithm Configuration) configuration tool and evaluate the performance improvements obtained. Experimental results show great promise for the automatic configuration approach. Based on these initial results, it is our intent to extend our approach to perform a proper evaluation of the state-of-the-art on heuristic methods for the Q3AP using systematically automatic configuration and design techniques for developing algorithms for the Q3AP.

References

[1] W. P. Pierskalla. *The multi-dimensional assignment problem.* Technical Memorandum No. 93, Operations Research Department, CASE Institute of Technology, 1967.

[2] P. M. Hahn, B-J. Kim, T. Stützle, S. Sebastian, W. L. Hightower, H. Samra, Z. Ding, M. Guignard. *The Quadratic Three-dimensional Assignment Problem : Exact and Approximate Solution Methods.* European Journal of Operational Research, 2008.

[3] F. Galea et al. *Bob++: a Framework for Exact Combinatorial Optimization Methods on Parallel Machines.* 21st European Conference on Modelling and Simulation, 2007.

[4] M. Mezmaz, M. Mehdi, P. Bouvry, N. Melab, E-G. Talbi, D.Tuyttens, *Solving the Three Dimensional Quadratic Assignment Problem on Grid Computing System.* Cluster Computing Journal, Volume 17 Issue 2, June 2014, Pages 205-217.

[5] Hans D. Mittelmann, Domenico Salvagnin, *On Solving a Hard Quadratic 3-Dimensional Assignment Problem.* The Mathematical Optimization Society, 2014.

[6] T. V. Luong, L. Loukil , N. Melab, E.-G Talbi, *A GPU-based Iterated Tabu Search for Solving the Quadratic 3-dimensional Assignment Problem.* IEEE AICCSA Conference, Workshop on Parallel Optimization in Emerging Computing Environments (POECE), Tunisia, 2010.

[7] L. Loukil, M. Mehdi, N. Melab, E.-G.Talbi, P.Bouvry, *Parallel Hybrid Genetic Algorithms for Solving Q3AP on Computational Grid.* International Journal of Foundations of Computer Science, Volume 23, Issue 02, February 2012.

[1] http://iridia.ulb.ac.be/irace/

[8] P. Lipinski,*A Hybrid Evolutionary Algorithm to Quadratic Three-Dimensional Assignment Problem with Local Search for Many-Core Graphics Processors.* Springer-Verlag Berlin Heidelberg, Volume 6283, 2010, pp344-351.

[9] M. Mehdi, J.-C. Charr, N. Melab, E.-G. Talbi, P. Bouvry. *A Cooperative Tree-based Hybrid GA-B&B Approach for Solving Challenging Permutationbased Problems*, GECCO 2011, Proceedings, Dublin, Ireland, July 12-16, 2011

Exact solution methods for the bi-objective {0, 1}-quadratic knapsack problem

J.P. Hubinont,
Université Libre de Bruxelles, Service CoDE, Département Math et gestion (SMG)
e-mail: jhubinon@ulb.ac.be

J. Rui Figueira
Universidade de Lisboa, CEG-IST, Instituto Superior Técnico de Lisboa
e-mail: figueira@tecnico.ulisboa.pt

Y. De Smet
Université Libre de Bruxelles, Service CoDE, Département Math et gestion (SMG)
e-mail: yves.de.smet@ulb.ac.be

The binary quadratic knapsack problem has been intensively studied over the last decades. Despite most of real-world problems are of a multi-dimensional nature there is not a considerable body of research for multi-objective binary quadratic knapsack problems, even in the case when only two objective functions are considered. The purpose of this paper is twofold: First, it aims at giving some insights on the bi-objective quadratic knapsack problem in order to solve it faster by building adequate heuristics and bounds for being used in approximations methods; and, second, to study a new linearization that involves less constraints than the existing ones. This paper addresses thus the bi-objective binary quadratic knapsack problem and proposes new algorithms, which allow for generating the whole set of exact nondominated points in the objective space and one corresponding efficient solution in the decision space. Moreover, this paper proposes a new linearization of the binary quadratic knapsack problem, which appears, on average, to be faster than the existing ones for solving the single and bi-objective quadratic knapsack problem. Computational results are presented for several types of instances, which differ according to the number of items or variables and the type of correlation between the two objective functions coefficients. The paper presents thus a comparison, in terms of CPU time, of four linearizations over several types of 30 instances of the same type and size.

Analyzing objective interaction in lexicographic optimization

Federico Mosquera Pieter Smet
Greet Vanden Berghe

KU Leuven, Department of Computer Science, CODeS & imec
Gebroeders De Smetstraat 1, 9000 Gent, Belgium
email: federico.mosquera@kuleuven.be

Despite the multitude of optimization methods proposed throughout operational research literature, such approaches can be difficult to grasp for decision-makers who lack a technical background. Solving a multi-objective optimization problem requires expertise from the decision-makers to manage complex objectives which may interact unintuitively. This study focuses on improving lexicographic optimization, a type of approach which has the advantage of being easily implemented by decision-makers given that its only requirement is the hierarchical arrangement of objectives in terms of their relative importance.
A pure lexicographic strategy optimizes each objective sequentially, in order of importance, without deteriorating previous objectives. Algorithmic performance is, however, often hindered when a pure lexicographic optimization strategy is employed. Typically, heuristic methods quickly converge to a local optimum which, depending on the specifics of the solution space, results in poor solutions. Employing integer programming solvers may also prove difficult as lexicographic optimization is not implemented natively and often requires solving the integer programming problem several times. This study overcomes these difficulties by first analyzing objective interactions and categorizing the various objectives accordingly. This information is subsequently used to develop a new methodology which seeks to improve search algorithms for solving lexicographic optimization problems.
A real-world case study concerning home care scheduling demonstrates how the proposed methodology improves pure lexicographic optimization. Home care scheduling not only involves the scheduling of caregivers, but also their assignment and routing, both of which are necessary to deliver the necessary services to different clients. Among the objectives requiring optimization are task frequency, travel time, caregiver/client preference satisfaction and the weekly spreading of tasks. Despite the complexity of these objectives, lexicographic optimization offers a user-friendly approach but, as demonstrated by way of a series of computational experiments, may result in poor quality solutions. Computational results demonstrate how the new methodology based on inter-objective relations results in better solutions.

Military manpower planning through a career path approach

Oussama Mazari Abdessameud Filip Van Utterbeeck
Johan Van Kerckhoven
Royal Military Academy, Departement of mathematics
e-mail: omazaria@vub.ac.be

Marie-Anne Guerry
Vrije Universiteit Brussel, Department of Business Technology and Operations

Military manpower planning aims to match the required and available staff. In order to tackle this challenge, we resort to military manpower planning, which involves two linked aspects, statutory logic and competence logic. The statutory logic aims to estimate the workforce evolution on the strategic level, whereas the competence logic targets the assignment of the right person to a suitable position. These two aspects are interdependent; therefore this article presents a technique to combine both logics in the same integrated model using career paths.

A military organization consists of a limited number of hierarchical job positions occupied by soldiers. Each job position requires a combination of occupant rank and skill/competence (usually gained by training). The military organization has a strict hierarchical structure, and can recruit only at the lowest rank. Because of these restricted possibilities in recruitment, military manpower planning is of major importance.

During his service at the military organization, each soldier is characterized by his system state: the combination of his job position, rank, acquired skills and competences. The evolution through time of this system state is called the soldier's career path. Having a limited number of job positions, ranks, skills and competences makes the number of possible career paths in the military organization limited. Moreover, not all feasible career paths have to be considered as certain career paths might be undesirable either from the point of view of the organization or the point of view of the soldier.

We propose a model based on a career path approach, in which we consider the assignment aspect as well as the strategic view on the manpower of the organization. In order to model the organization with a career path approach, the first step is the identification of possible career paths. Possible career paths adhere to human resources management policies of the organization, which are policies related not only to the competence logic but also to the statutory logic.

Representing the organization with a career path approach means that the manpower

in the organization is divided into subgroups assigned to the identified possible career paths. If we suppose that the organization has a known capacity of recruitment each year, the operational goal of the organization as well as the strategic one are translated into finding the optimal amounts to recruit in each career path. The recruited amounts in each career path are proportions of the recruitment capacity of the organization. In order to find the optimal proportions, we resort to mathematical programming. An important aspect to take in consideration is the fact that the available workforce is not always equal to the demand. Having this deviation between the required and the available workforce, goal programming is an appropriate approach to use.

To write the goal program, we consider three types of variables: recruitment distributions, initial manpower distributions and the deviations. The recruitment distribution gathers the amounts of recruited personnel to assign to each career path. The initial manpower distribution define the assignment of the initial manpower to different career paths. The deviation variables are used in the soft constraints to define the deviation from the targeted goal.

The developed model permits the military human resources managers to study the impact of the used policies on the strategic level as well as the operational level. Implemented in a military human resource management department, the model gives detailed plans for the future actions to be taken. The provided actions are linked to the job transfers, the yearly recruitment and recruitment distributions, and the retirement plan in a case of flexible retirement.

Optimization of Emergency Service Center Locations Using an Iterative Solution Approach

Mumtaz Karatas
National Defense University, Turkish Naval Academy, Industrial Engineering Department
e-mail: mkaratas@dho.edu.tr

Ertan Yakici
National Defense University, Turkish Naval Academy, Industrial Engineering Department
e-mail: eyakici@dho.edu.tr

Location problems seek to determine the best locations for facilities distributing services or goods. In particular, problems considering the locations of emergency service systems, i.e. police centers, medical centers, search and rescue resources, firefighting centers, have been widely studied by many scholars.

In this study, we consider the problem of determining the optimal locations of temporary emergency service centers for a regional natural gas distribution company in the Marmara Region of Turkey. These emergency centers will be activated for a temporary period of time after a possible earthquake or natural disaster. They will be used to respond cases like gas leakages, gas distribution problems, or other situations which may cause emergencies. In the region of interest, there are a total of 82,816 customers that are categorized as: (1) residential, e.g. houses, apartment blocks, dormitories, (2) commercial, e.g. automobile repair shops, convention centers, pharmacies, gas stations, hotels, markets, shopping malls, warehouses, restaurants, and (3) industrial, e.g. factories, power plants. Considering a discrete case, we seek to locate an unknown number of temporary emergency service centers to a total of 32 candidate locations. The company, currently, has not decided on the exact number of centers to be activated in case of an earthquake, but limited the maximum possible number to 15. Therefore, our problem includes determining both the locations and the number of facilities to be installed.

The effectiveness of such emergency service systems are measured in terms of a number of performance metrics. The most commonly-used metrics are: (1) minimization of the average waiting time of customers before being served, (2) maximization of the total number of customers that are served before a pre-determined time threshold, and (3) minimization of the maximum waiting time of customers before being served. These metrics are the objectives of the three well-known location models, p-median problem, maximal covering location problem and p-center problem, respectively. Most real-world location problems involve all three conflicting objectives and satisfying them simultaneously is a challenge for researchers.

In our problem, we incorporate the three objectives as to minimize the average and maximum transfer time between all customers and closest temporary emergency service centers, along with covering as many customers as possible within a critical range determined for each customer type separately. To generate a representative

Pareto optimal solution set for the abovementioned large-scale multi-objective location problem, we first employ a k-means clustering algorithm and decrease the number of customers from 82,816 to 274. Next, we adopt an iterative solution technique which is based on the combination of the branch and bound and iterative goal programming methods. This method, basically, solves each of the individual location model iteratively, while keeping a track of the quality of the objective function values of each model. Finally, it provides a Pareto optimal solution set as well as a compromise solution for the three objectives. We solve the problem for different numbers of centers varying between 1 and 15. Our results show that additional centers after activating 10 does not create a substantial contribution to any of the objectives. Hence, we suggest opening 10 centers and provide the decision-maker with a set of diverse Pareto-optimal solutions as well as a compromise solution.

Vertical supply chain integration - Is routing more important than inventory management?

Florian Arnold Johan Springael
Kenneth Sörensen
University of Antwerp, Departement of Engineering Management
ANT/OR - Operations Research Group
e-mail: florian.arnold@uantwerpen.be

Consider the following problem. A logistic service provider or retailer owns multiple depots that are located in different regions of a country. The company is responsible for delivering one or several products from the depots to fulfill the demand of customers. From historical data the company has information about the demand density per region. The company takes care of the inventory management in its depots as well as the distribution processes from the depots to the customers, and aims to minimize the total costs of both logistic activities (*vertical integration*).
The traditional approach has been to tackle these two problems sequentially and avoid any compounding complexity. In the inventory literature, the best inventory strategy - when to reorder and to which level - is usually determined for each depot separately, so that the demand of any depot is treated independently from the demand of other depots. The assignment of which customer is delivered by which depot is *fixed*. Multiple depots can collaborate by *pooling* their inventory. Pooling has been shown to be highly cost-efficient in many scenarios and is realized via *lateral transshipments*. Transshipments allow depots with low or missing inventory to request an order from another depot. If the other depot is nearby and has the demanded products on stock, the inventory can be restocked faster and at lower costs than with a regular order, which results in a higher flexibility to adapt to customer demand.
Reversely, in the routing literature it is generally assumed that any customer can be delivered by any depot, the distribution activities are carried out jointly. Thus, the assignment of which customer is delivered by which depot is *flexible*, which is a crucial assumption to solve multi-depot vehicle routing problems (MDVRP). With the above terminology, we could also speak of a pooling of distribution activities. However, routing problems largely ignore the inventory perspective, assuming a sufficient availability of inventory in all depots.
Consequently, the fields of inventory management and distribution take different perspectives and make different assumptions. In inventory management, constrained inventory can be pooled to supply customers with a fixed allocation in such a way that inventory costs are minimised. In distribution, unconstrained inventory is delivered flexibly to customers in such a way that routing costs are minimised. From the perspective of a vertical integration, there is thus the immediate question how to integrate the optimisation of the distribution and inventory activities. Should we determine an optimal inventory strategy first and derive the routing second, or

rather derive the inventory levels on the basis of the best distribution plans?
In this paper, we take a first step to answer these questions and develop an optimisation framework that enables the analysis of inventory strategies for multiple depots by taking the distribution into account. This approach requires us to tackle two problems: We need to compute inventory policies and we need to solve routing problems that consider inventory constraints. A certain strategy can then be evaluated by simulating the routing and inventory activities over a longer time horizon, an approach called *optimisation-simulation*.
With this optimisation framework we aim at answering the question: 'Is it useful to pool inventory and distribution, and if so, in which circumstances?'. We perform a series of experiments to answer this question and derive guidelines for a cost-efficient vertical integration of inventory and distribution decisions in a multi-depot context. As the main finding we demonstrate that the pooling of distribution is more cost-efficient than the pooling via lateral transshipments.

Stochastic solutions for the Inventory-Routing Problem with Transshipment

Wouter Lefever
Department Industrial Engineering Systems and Product Design, Ghent University,
Technologiepark 903, Zwijnaarde 9052, Belgium
e-mail: wouter.lefever@ugent.be

El-Houssaine Aghezzaf
Department Industrial Engineering Systems and Product Design, Ghent University,
Technologiepark 903, Zwijnaarde 9052, Belgium
e-mail: elhoussaine.aghezzaf@ugent.be

Khaled Hadj-Hamou
Univ. Lyon, INSA Lyon, DISP, 69621 Villeurbanne, France
e-mail: khaled.hadj-hamou@insa-lyon.fr

Introduction

Currently, one of the most investigated supply chain strategies is Vendor-Manged Inventory (VMI). In a VMI system the supplier manages the retailer's inventory. He determines the timing and size of his deliveries so that he can optimize the overall distribution costs.
The mathematical model that optimizes the decisions of the supplier in a VMI system is called the Inventory-Routing Problem (IRP). Recent attempts to introduce uncertainty on the demand, have shown that the basic IRP performs very poorly when the demand is not known a priori.
In this work, we study a more flexible variant of the IRP, the Inventory-Routing Problem with Transshipment (IRPT), and investigate its suitability in a stochastic context. We develop an sample average approximation (SAA) algorithm that effectively solves the stochastic IRPT (SIRPT).

Methodology

Chrysochoou and Ziliaskopoulos [1] demonstrate that a two-stage stochastic program can be formulated to solve the SIRPT. The master problem optimizes the routing and delivery decisions of the supplier. Then, after the exact demand is revealed, the second-stage decisions determine the timing and size of additional deliveries, the transshipments. The computational results show that only a small number of scenarios can be included due to the complexity of the problem.
Starting from this result, we reformulate the second-stage problem of the SIRPT. In our altered second-stage problem, we allow forbidden transshipments, but penalize these very hard. These changes make the first-stage solution always second-stage feasible. Our experiments show that the optimality cuts from our new second-stage

problem are far more effective than the feasibility cuts from the original second-stage problem. Since the forbidden transshipments are penalized, they do not appear in the final solution.
To further improve the quality of our stochastic solution, we apply an "upgrade" to the solution of the SIRPT, based on the work of Maggioni and Wallace [3]. The first-stage decision variables consist of the binary routing variables and the continuous delivery variables. When the binary routing variables are fixed, the resulting SIRPT can be solved as a linear problem in polynomial time. In order to fully utilize this particular structure of the problem, we decompose the optimization of the binary routing variables and the optimization of the continuous delivery variables: (1) we solve the SIRPT for a relatively small scenario set N so that it can be solved to optimality in reasonable time. (2) we fix the first-stage routing decisions based on the result of (1) and resolve the SIRPT for a much larger larger scenario set $N' > N$. Our computational experiments showed that this "upgraded" solution results in improvements up to 5%.
One drawback of this approach is that the skeleton of the solution is determined based on a small scenario set. If the scenarios are chosen unfortunately, the obtained skeleton may be very different from the skeleton of the optimal solution. To overcome this problem, we could add more scenarios to the initial scenario set N. However, the computational complexity of the SIRPT increases faster than linearly with respect to the number of scenarios in the scenario set N. Therefore, we choose to replicate our procedure multiple times. An additional advantage is that the gap between the obtained solution values and the "true" stochastic solution value can be estimated (Kleywegt et al. [2]).

Results and Conclusion

To validate our SAA algorithm, we set up an experiment using benchmark instances from literature. The results show that for instances with up to 20 retailers a good quality stochastic solution can be found within reasonable time.

References

[1] E. Chrysochoou and A. Ziliaskopoulos. Stochastic inventory routing problem with transshipment recourse action. *SMMSO 2015*, page 17, 2015.

[2] A. J. Kleywegt, A. Shapiro, and T. Homem-de Mello. The sample average approximation method for stochastic discrete optimization. *SIAM Journal on Optimization*, 12(2):479–502, 2002.

[3] F. Maggioni and S. W. Wallace. Analyzing the quality of the expected value solution in stochastic programming. *Annals of Operations Research*, 200(1):37–54, 2012.

Solving an integrated order picking-vehicle routing problem with record-to-record travel

Stef Moons, Kris Braekers, Katrien Ramaekers, An Caris
UHasselt - Hasselt University, Research group Logistics
e-mail: stef.moons@uhasselt.be

Yasemin Arda
QuantOM, HEC Liège - Management School of the University of Liège

In the last decade, European business-to-consumer (B2C) e-commerce sales have been increasing with an annual growth rate of approximately 17%. The growing popularity of e-commerce has a positive impact on the number of parcels that needs to be delivered with approximately 4.2 billion parcels in Europe in 2016 [1]. When customers purchase goods online, they expect a fast and accurate delivery at low cost. The combination of the increasing number of orders and the high expectations puts a pressure on the logistic activities of B2C e-commerce companies. In order to efficiently satisfy the customer expectations, companies have to optimize their picking and distribution operations. Implementing an integrated approach to obtain simultaneously order picking schedules and vehicle routes can lead to better results [2].

Goods purchased online at webshops need to be picked in a distribution centre (DC) before they can be delivered to the customer. As such, the order picking and delivery operations are related. Consequently, to determine better picking lists and vehicle routes, ideally, these two problems should be integrated into a single optimization problem. In an integrated order picking-vehicle routing problem (I-OP-VRP), the two problems are solved simultaneously by taking into account the requirements and constraints of both problems.

In the I-OP-VRP, order pickers work in parallel in a single zone in a DC. Additional temporary order pickers can be hired from a fixed pool of workers in case of high customer demand. The labour cost of the temporary pickers is slightly higher than this of regular pickers. Each order is picked in an individual tour throughout the DC without batching. The picking routes are considered to be known in advance. A homogeneous fleet of vehicles is used for the deliveries to the customers. During the purchasing process, each customer has selected a delivery time window. A cost is incurred for each kilometre travelled and for the working time of the driver. The objective of the I-OP-VRP is to minimize total cost, i.e., both order picking and distribution costs.

Since both an order picking problem and a vehicle routing problem are hard to solve themselves, a heuristic solution algorithm is needed to solve the I-OP-VRP. A record-to-record travel (RRT) algorithm with local search operators is developed to solve the integrated problem in a small amount of computational time. RRT is a deterministic variant of simulated annealing and is first developed by [3]. Each new

solution is accepted when its objective value is not worse than the record, i.e., best solution found, plus a deviation value, which is a certain percentage of the record.
An initial solution is generated by using a constructive heuristic which consists of two parts; one for each subproblem. The RRT solution algorithm is used to improve the quality of the initial solution and is applied for a maximum number of iterations. Within each iteration, five local search operators are executed. Three operators are related to the vehicle routing part of the problem: exchange, relocate, and, 2-Opt. Two operators work on the order picking schedules: exchange and relocate. The algorithm restarts with the best solution found so far when the RRT could not improve the solution for a number of consecutive iterations.
The algorithm is tested on randomly generated instances with 10 and 15 customer orders. The performance of the proposed heuristic is evaluated by comparing its results with the optimal solutions obtained by CPLEX. The heuristic is capable of obtaining high-quality solutions in small computational time. Solving an instance with 10 customer orders to optimality using CPLEX takes 15 minutes on average, while the RRT heuristic finds a solution within less than a second. Obtaining the optimal solution of an instance with 15 customer orders with CPLEX takes 57 hours on average, whereas the proposed heuristic is capable of finding a solution within 3 seconds.

References

[1] Ecommerce Europe (2016). *European B2C E-commerce Report 2016.*

[2] Moons, S., Ramaekers, K., Caris, A., & Arda, Y. (in press), Integration of order picking and vehicle routing in a B2C e-commerce context, *Flexible Services and Manufacturing Journal*, doi: 10.1007/s10696-017-9287-5.

[3] Dueck, G. (1993). New optimization heuristics: The great deluge algorithm and the record-to-record travel. *Journal of Computational physics*, 104(1), 86-92, doie: 10.1006/jcph.1993.1010.

The simultaneous vehicle routing and service network design problem in intermodal rail transportation

H. Heggen[(a,b)], A. Caris[(a)], K. Braekers[(a)]
(a) UHasselt, Research Group Logistics
(b) Maastricht University, Department of Quantitative Economics
e-mail: {hilde.heggen, an.caris, kris.braekers}@uhasselt.be

Within the aim of stimulating the use of intermodal rail transport, an important goal is the reduction of the total costs of the intermodal transport system. Costs related to the execution of first- and last-mile drayage operations are a critical aspect within this regard. In addition, in order to get a competitive advantage, customers may be offered the possibility to book orders close to the required pick-up deadline for transport.

In order to deal with these challenges, intermodal logistics service providers must make important decisions with respect to the planning of drayage and main-haul transport. Tactical decisions related to the service network design concern the selection of services and their characteristics as well as the routing of demand flows throughout the network [1]. Characteristics of regular long-haul services include the route, intermediary stops, frequency, vehicle and convoy type, capacity and speed. At the operational decision level, intermodal routing is concerned with the selection of itineraries for individual shipments through a given intermodal network in which the selected services are fixed [2]. The fact that both problems are interrelated is recently stressed by several authors (e.g. [3, 4]).

Current research includes detailed capacity requirements at the operational decision level when load units are already assigned to specific rail routes, such as the train load planning problem in which load units are assigned to specific locations on intermodal trains [5]. On the other hand, tactical problems generally consider a very high-level view on capacity (e.g. a number of load units). Moreover, traditionally, rail planning decisions on how shipments are routed throughout the long-haul rail network are fixed before truck routing decisions are made to combine pickups and deliveries in the service area of terminals.

In practice, at the tactical decision level, detailed operational aspects must already be considered. Slots to be purchased on externally managed trains and wagon lease agreements of own services should be determined based on expected demand patterns. In this research, we aim to integrate operational aspects into tactical decisions on the service network design from the perspective of an intermodal logistics service provider who has to route orders throughout its intermodal truck-rail transport network. Available loading meters and path weight constraints of own trains as well as the number of purchased slots, length types and weight limits on trains owned by external operators are considered. We study how these tactical decisions influence

total local drayage routing costs and patterns to pick up and deliver load units in each terminal service area by considering truck routing and rail planning decisions simultaneously instead of sequentially. In accordance with a more detailed view on capacity, a heterogeneous fleet of trucks is considered.
A real-life case study is investigated in which direct trains are available between a number of terminals in the Benelux and a number of terminals in Northern Italy. Trains are operated based on given periodical schedules with a cycle of one week, using both own trains and slots purchased with other operators. The aim is to simultaneously determine the routing of trucks within each service area and the services and their respective capacities offered by an intermodal operator. Decisions on local drayage routing in large-volume freight regions with multiple terminals on the one hand and intermodal long-haul routing throughout the service network on the other hand are merged into an integrated intermodal routing problem. A mathematical problem formulation is presented. Preliminary insights can be obtained based on results of small problems using CPLEX in C++.
In order to quantify the advantages of the integrated problem, its results are compared with the classical approach in which first the long-haul transportation routes are defined, and next truck pickups and deliveries are performed in the service region. Moreover, this integrated model can provide important insights to logistical service providers by analyzing the impact of tactical decisions on operational transport costs and current routing policies. Results of a number of realistic tactical decision scenarios are analyzed. For example, the influence can be quantified of accepting new customers, changing the number of terminals in each region and capacity levels and service frequencies of long-haul connections. Moreover, different demand levels can be analyzed to anticipate possible future demand scenarios. In future work, heuristics will be used to solve real-life instances and analyze the influence of tactical decisions.

References

[1] Crainic, T.G., Kim, K.H. (2007). Chapter 8: Intermodal transportation. In *Handbooks in Operations Research and Management Science*. Elsevier, 14, 467-537.

[2] Caris, A., Macharis, C., Janssens, G.K. (2013). Decision support in intermodal transport: A new research agenda. *Computers in Industry*, 64 (2), 105âĂŞ12.

[3] Van Riessen, R.R., Dekker, R., Lodewijks, G. (2015). Service network design for an intermodal container network with flexible transit times and the possibility of using subcontracted transport. *International Journal of Shipping and Transport Logistics*, 7(4), 457-478.

[4] Zhang, M., Pel, A.J. (2016). Synchromodal hinterland freight transport: Model study for the port of Rotterdam. *Journal of Transport Geography*, 52, 1-10.

[5] Heggen, H., Braekers, K., Caris, A. (2017). A multi-objective approach for intermodal train load planning. *OR Spectrum*, Special issue on rail terminal operations, accepted for publication (doi: 10.1007/s00291-017-0503-1).

Local evaluation techniques in bus line planning

Evert Vermeir
KU Leuven, Leuven Mobility Research Centre - CIB
e-mail: evert.vermeir@kuleuven.be

Pieter Vansteenwegen
KU Leuven, Leuven Mobility Research Centre - CIB
pieter.vansteenwegen@kuleuven.be

Public transport is essential to cope with the ever increasing demand for mobility. Hence, it is important to utilize all available resources as efficiently as possible. For a public transport system, this efficiency depends strongly on the decisions made during the planning process -[1]-. The planning process usually consists of a few different phases. One of these steps is line planning, where a public transport company decides where its vehicles will drive, which stops will be served and in which order. A main challenge of the bus line planning problem is finding a good solution for (very) large networks -[2, 3]-.
The problem is that the evaluation of a line plan is computationally quite costly. In an evaluation, the total travel time of all the passengers is calculated. To make such a calculation, the route each passenger takes has to be known. In line planning, this is nearly always done by finding the shortest path for all origin-destination pairs. Deciding which routes are used to get from each possible origin to each possible destination is the most time consuming part of the evaluation-[4]-.
Local evaluation techniques can be used to decrease the computational burden of each evaluation. The computational burden scales more than linearly with the size of the network -[3]-, limiting the calculations to the part of the network where a change has happened thus results in a significant decrease in computation time. Apart from being able to find a result for larger networks, this also allows the use of more complex and realistic techniques or the integration of some of the other planning steps into the line planning process.
This paper implements local evaluation techniques based on two core principles. The first idea is that the impact of a change decreases the further you move away from this change, and the second one is that in a good line plan passengers only make a low number of transfers to get to their destination. In stead of always evaluation the entire transit network, a smaller network is cut from the original using the two core principles. It is then assumed that all routes that enter and/or leave this cut, will still do so on the same place after the change. This means that the current route travelled by each passenger has to be saved. With this information, the demand of the cut can be constructed from the demand for the complete network. Note that routes that previously did not use any of the lines included in the cut are assumed not to change, nor will any routes change their place of entry nor exit in the cut-out network.
To test this framework, an iterated local search algorithm is developed to solve the

line planning problem. This algorithm constructs a line plan from scratch and then improves the current solution until a local optimum is reached. Then, a part of the solution is destroyed and a new search for a local optimum begins. This process continues until the stop criterion is reached.
While the local evaluation techniques are applied to an iterated local search, the two core concepts can be applied to many other algorithms and could be useful for all research in line planning or for other similar public transport optimization problems.

References

[1] Lusby, R. M., Larsen, J. and Bull, S., *A survey on robustness in railway planning*, European Journal of Operational Research, Volume 266, Issue 1, 2018, Pages 1-15.

[2] Guihaire, V. and Hao, J.-K., *Transit network design and scheduling: A global review*, Transportation Research Part A, Volume 42, Issue 10, 2008, Pages 1251-1273.

[3] Schöbel, A., *Line planning in public transportation: models and methods*, OR Spectrum, Volume 34, Issue 3, 2012, Pages 491-510.

[4] Schmidt, M. E., *Integrating Routing Decisions in Public Transportation Problems*, Springer Optimization and Its Applications, Volume 89, 2014, New York, Springer.

Integrating dial-a-ride services and public transport

Y. Molenbruch
Hasselt University, Research Group Logistics
e-mail: yves.molenbruch@uhasselt.be

K. Braekers
Hasselt University, Research Group Logistics
e-mail: kris.braekers@uhasselt.be

January 17, 2018

Dial-a-ride systems are applications of demand-responsive, collective passenger transport. They usually provide door-to-door services to people with reduced mobility, such as elderly or disabled. Contrary to general taxi services, different users may be grouped together in the same vehicle, as long as certain service level requirements (i.e. time windows, maximum user ride time, ...) are respected. Because of this, providers of dial-a-ride services face a complicated vehicle routing problem in their daily operational activities, which is referred to as the dial-a-ride problem (DARP).

In a traditional mobility policy, DAR services are exploited separate from the regular public transport (PT). To increase the overall efficiency and reduce the required subsidies, modern policy visions combine both systems into a hierarchical structure. The regular PT is only maintained on a core network of (sub)urban and interurban lines, whereas a DAR company provides on-demand transportation in rural areas, both to people with and without mobility restrictions. Trips of an ambulant person may involve a combination of DAR services and PT. This requires the operational activities of both systems to be integrated, such that the users in rural areas can submit one single request for their entire trip. The DAR provider needs to determine itineraries and potential transfer locations such that operational costs are minimized, given the timetables of the PT services. To the best of our knowledge, the academic literature does not provide any algorithm to efficiently solve this integrated problem variant.

The metaheuristic solution approach developed in this research is an extension of the large neighborhood search (LNS) algorithm presented in Molenbruch et al. [1]. It includes three well-known destroy operators (random, worst and related) and three well-known repair operators (random order, greedy, 2-regret) which iteratively remove and reinsert a certain percentage of the requests. Promising solutions are intensified using additional local search operators.

Within this LNS algorithm, the possibility of integration is considered during the insertion phase. For requests of mobile users, it is checked whether the total distance of the DAR vehicles can be reduced by covering part of their trips by PT. To fully

exploit the benefits of integration, the choice of the transfer locations is not fixed in advance, but determined as a function of the actual structure of the DAR routes. Specific criteria are proposed to determine which transfer nodes are more likely to be selected, e.g. based on the structure of the PT network.

Moreover, attention needs to be devoted to the time feasibility check. Because of the combination of maximum user ride time constraints and time windows, it is already hard to determine whether a feasible time schedule exists for a single DAR route. In the integrated problem variant, changing the schedule of one route may affect the feasibility of another route. These interdependencies make the feasibility check even more complicated. A scheduling procedure for the DARP with transfers, based on a representation as a simple temporal problem [2], is extended to the integrated problem variant.

Computational experiments are performed on an extended version of the data set introduced by Røpke et al. [3]. It consists of artificial instances which include up to 96 requests and 8 DAR vehicles. User locations are randomly and independently generated in a square region. Time windows, service durations, maximum user ride times, maximum route durations and vehicle capacities are taken into account. For the purpose of these experiments, an artificial (urban) PT network is added to all instances. Results show that for this data set, operational savings up to 24% can be achieved by the DAR provider. These savings strongly depend on the density of the PT network, whereas the frequency is less decisive. Service level decisions and demand characteristics also influence the savings.

Acknowledgements

We thank prof. dr. Patrick Hirsch and Marco Oberscheider (University of Natural Resources and Life Sciences, Vienna) for their contribution in this research project.

References

[1] Molenbruch, Y., Braekers, K. and Caris, A., 2017. Benefits of horizontal cooperation in dial-a-ride services. Transportation Research Part E 107, 97–119.

[2] Masson, R., Lehuédé, F. and Péton, O., 2014. The dial-a-ride problem with transfers. Computers & Operations Research, 41, 12–23.

[3] Røpke, S., Cordeau, J.-F. and Laporte, G., 2007. Models and branch-and-cut algorithms for pickup and delivery problems with time windows. Networks 49 (4), 258–272.

Benchmarks for the prisoner transportation problem

J. Christiaens T. Wauters
G. Vanden Berghe
KU Leuven, Department of Computer Science
e-mail: jan.christiaens@cs.kuleuven.be

This paper formally introduces a new problem to the operational research academic community: the Prisoner Transportation Problem (PTP). This problem essentially concerns transporting prisoners to and from services such as hospitals, court and family visits. Depots are located across a geographic region, each of which is associated with several special prisoner transportation trucks which themselves are subdivided into compartments prisoners may share with others. The PTP constitutes a variant of the commonplace vehicle routing problem, but more specifically the dial-a-ride problem. Two unique constraints make it a particularly interesting problem from a research context: (i) simultaneous service times and (ii) inter-prisoner conflicts. Simultaneous service times mean that prisoners may be processed simultaneously rather than sequentially, as is standard throughout most vehicle routing and delivery problems. Meanwhile, compartment compatibility means that each individual prisoner may be either permitted to share the same compartment with another prisoner, unable to share the same compartment, or they may be unable to share even the same vehicle. The inclusion of this additional constraint results in a very computationally challenging problem. The algorithm developed and employed to solve this problem does not represent this work's primary contribution. Rather the much more fundamental task of introducing and modelling this problem's properties and unqiue constraints constitutes the primary research interest. Given that the problem has never been formalised by the academic community, it continues to be solved primarily by manual planners who spend multiple hours working on a solution for the following day. These solutions are often infeasible or violate hard constraints. The ruin & recreate algorithm proposed by the current paper significantly improves upon these manually-constructed solvers and only require approximately ninety seconds of runtime. Computational results will be presented and analysed in detail. Furthermore, the academic instances which were generated and employed for experimentation will be made publicly available with a view towards encouraging further follow-up research.

Optimizing city-wide vehicle allocation for car sharing

T. I. Wickert[a,b], F. Mosquera[a], P. Smet[a], E. Thanos[a]
[a]KU Leuven, Department of Computer Science, CODeS & imec
[b]Federal University of Rio Grande do Sul (UFRGS), Department of Computer Science
e-mail: toniismael.wickert@kuleuven.be

Car sharing is a form of shared mobility in which different users share a fleet of vehicles rather than all using their own vehicle. Users make requests to reserve these shared vehicles and only pay for the time of use or distance driven. Car sharing stations, where the vehicles are parked when they are not used, must be strategically located within a city such that as many users as possible can be serviced by the available fleet of vehicles. Typically, these stations are also close to train or bus stations to facilitate combination with public transport.
This study proposes a decision support model for optimizing the allocation of car sharing stations within a city. The goal is to distribute the fleet of available vehicles among zones such that as many user requests as possible can be assigned to a specific vehicle. The optimization problem to be solved is thus a combination of two assignment problems: the assignment of vehicles to zones and the assignment of user requests to vehicles. A request may be feasibly assigned to a vehicle if that vehicle is located in the request's home zone or, less preferable, in an adjacent zone. Requests which overlap in time cannot be assigned to the same vehicle. The objective is to minimize the total cost incurred by leaving requests unassigned and by assigning requests to vehicles in adjacent zones.
Computational experiments on both real-world data and computer-generated problem instances showed that medium-sized and large problems cannot be solved using integer programming. To address such challenging problem instances, a heuristic decomposition approach is introduced. The proposed algorithm is inspired by the successful application of this technique on various optimization problems such as the traveling umpire problem, nurse rostering, project scheduling and capacitated vehicle routing. The heuristic decomposition algorithm starts by generating an initial feasible solution. Then, iteratively, a subset of decision variables is fixed to their current values, and the resulting subproblem is solved to optimality using a general purpose integer programming solver. Three types of decomposition are investigated: 1) fixing a subset of vehicles to zones, 2) fixing a subset of requests to vehicles and 3) fixing requests from a subset of days to vehicles.
A computational study showed that the proposed method outperforms both an integer programming solver and a simulated annealing metaheuristic. Details regarding the heuristic decomposition implementation and computational results will be presented at the conference.

On computing the distances to stability for matrices

Punit Sharma
Department of Mathematics and Operational Research, University of Mons
Belgium
e-mail: punit.sharma@umons.ac.be

Nicolas Gillis
Department of Mathematics and Operational Research, University of Mons
Belgium

The stability of a continuous linear time-invariant (LTI) system $\dot{x} = Ax + Bu$, where $A \in \mathbb{R}^{n,n}$, $B \in \mathbb{R}^{n,m}$, solely depends on the eigenvalues of the stable matrix A. Such a system is stable if all eigenvalues of A are in the closed left half of the complex plane and all eigenvalues on the imaginary axis are semisimple. It is important to know that when an unstable LTI system becomes stable, i.e. when it has all eigenvalues in the stability region, or how much it has to be perturb to be on this boundary. For control systems this *distance to stability* is well-understood. This is the converse problem of the *distance to instability*, where a stable matrix A is given and one looks for the smallest perturbation that moves an eigenvalue outside the stability region. In this talk, I will talk about the distance to stability problem for LTI control systems. Motivated by the structure of dissipative-Hamiltonian systems, we define the *DH matrix*: a matrix $A \in \mathbb{R}^{n,n}$ is said to be a DH matrix if $A = (J - R)Q$ for some matrices $J, R, Q \in \mathbb{R}^{n,n}$ such that J is skew-symmetric, R is symmetric positive semidefinite and Q is symmetric positive definite. We will show that a system is stable if and only if its state matrix is a DH matrix. This results in an equivalent optimization problem with a simple convex feasible set. We propose a new very efficient algorithm to solve this problem using a fast gradient method. We show the effectiveness of this method compared to other approaches such as the block coordinate descent method and to several state-of-the-art algorithms. For more details, this work has been published as [1].

References

[1] N. Gillis and P. Sharma. On computing the distance to stability for matrices using linear dissipative Hamiltonian systems. *Automatica*, 85, pp. 113-121, 2017.

Low-Rank Matrix Approximation in the Infinity Norm

Nicolas Gillis
Université de Mons, Department of Mathematics and Operational Research
e-mail: nicolas.gillis@umons.ac.be

Yaroslav Shitov
National Research University Higher School of Economics
e-mail: yaroslav-shitov@yandex.ru

Low-rank matrix approximations (LRA) are key problems in numerical linear algebra, and have become a central tool in data analysis and machine learning; see, e.g., [1]. A possible formulation for LRA is the following: given a matrix $M \in \mathbb{R}^{m\times n}$ and a factorization rank r, solve

$$\min_{X\in\Omega} ||M - X|| \quad \text{such that} \quad \text{rank}(X) \leq r, \tag{3}$$

for some given (pseudo) norm $||.||$. In this paper, we focus on unconstrained variants, that is, $\Omega = \mathbb{R}^{m\times n}$, although there exists many important variants of (3) that take constraints into account such as nonnegative matrix factorization [2].
When the norm $||.||$ is the Frobenius norm, that is, $||M - X||_F^2 = \sum_{i,j}(M_{ij} - X_{ij})^2$, the problem can be solved using the singular value decomposition and is closely related to principal component analysis (PCA). In practice, it is often required to use other norms, e.g., the ℓ_1 norm which is more robust to outliers [3], and weighted norms that can be used when data is missing [4]. However, as soon as the norm is not the Frobenius norm, (3) becomes difficult in general; in particular it was proved to be NP-hard for all the previously listed cases [6, 5, 3].
In this paper, we focus on the variant with the component-wise ℓ_∞ norm:

$$\min_{X\in\mathbb{R}^{m\times n}} ||M - X||_\infty \quad \text{such that} \quad \text{rank}(X) \leq r, \tag{4}$$

where $||M - X||_\infty = \max_{i,j} |M_{ij} - X_{ij}|$. We will refer to this problem as ℓ_∞ LRA. It should be used when the noise added to the low-rank matrix follows an *i.i.d. uniform distribution.*

Outline and contributions In this talk, we prove that the decision variant of the problem (4) for $r = 1$ is NP-complete using a reduction from the problem 'not all equal 3SAT'. We also analyze several cases when the problem can be solved in polynomial time, and propose a simple practical heuristic algorithm which we apply on the problem of the recovery of a quantized low-rank matrix.

References

[1] M. Udell, C. Horn, R. Zadeh, and S. Boyd, *Generalized low rank models*, Foundations and Trends in Machine Learning, 9 (2016), pp. 1–118.

[2] D.D. Lee and H.S. Seung, *Learning the parts of objects by nonnegative matrix factorization*, Nature, 401 (1999), pp. 788–791.

[3] Z. Song, D.P. Woodruff, and P. Zhong, *Low rank approximation with entrywise ℓ_1-norm error*, in 49th Annual ACM SIGACT Symposium on the Theory of Computing (STOC 2017), 2017.

[4] K.R. Gabriel and S. Zamir, *Lower rank approximation of matrices by least squares with any choice of weights*, Technometrics, 21 (1979), pp. 489–498.

[5] N. Gillis and F. Glineur, *Low-rank matrix approximation with weights or missing data is NP-hard*, SIAM Journal on Matrix Analysis and Applications, 32 (2011), pp. 1149–1165.

[6] N. Gillis and S.A. Vavasis, *On the complexity of robust PCA and ℓ_1-norm low-rank matrix approximation*, Accepted in Mathematics of Operations Research, (2017).

Benchmarking some iterative linear systems solvers for deformable 3D images registration

Justin Buhendwa Nyenyezi
Université de Namur, Mathematics department, NaXys
e-mail: justin.buhendwa@unamur.be

Annick Sartenaer
Université de Namur, Mathematics department, NaXys

In medical image analysis, it is common to analyse several images acquired at different times or by different devices. The image registration is a process of finding a spatial transformation that allows these images to be aligned in a common spatial domain and correspondences to be established between them. The non-rigid registration, that allows local deformations in the image, requires resolution of large linear systems. These systems are derived from physical principles, modelled as Partial Differential Equations (PDEs), whose resolution allows to obtain a smooth displacement field. On the one hand, such linear systems are generally ill-conditioned. This leads to low convergence rate of the algorithms used to solve the problem or to inaccurate solutions. On the other hand, many applications of registration algorithms require faster registration algorithms [1]. Thus, there is a need to provide efficient system solvers especially for deformable nonparametric registration applied to high-resolution $3D$ images. In such applications, the system resolution is indeed among the most expensive steps.

For designing efficient system solvers, one may analyze structures of general operators or matrices used by the algorithms. In our case, although these systems are often sparse and structured, they are very large and ill-conditioned. Thus, their solvers are time consuming and their complexity in operation counts is polynomial. As a consequence, fast and superfast direct methods reveal numerical instabilities and thus lead to breakdowns or inaccurate solutions. The most used alternative are iterative system solvers that enable a reduction of the number of expensive operations such as matrix-vector products and thus a speed up of the registration process. Although iterative solvers provide only an approximation of the solution, they are well suited for very large systems when cheap and well suited preconditioners are available. Preconditioners may be stationary (Jacobi, Gauss-Seidel or SOR) and non-stationary (polynonmial or low-rank tensors).

In this study, we are especially interested in benchmarking [2, 3] iterative system solvers on a large set of 3D medical images. These system solvers are based on the Conjugate Gradient method but different from the preconditioning techniques [5, 4] . The ongoing results confirm [3] that there is no single system solver that is the best for all the problems. However, the results indicate which solver has the highest probability of being the best within a factor $f \in [1, \infty[$ and considering a limited computational budget in terms of time and storage requirements.

References

[1] Oseledets, Ivan and Tyrtyshnikov, Eugene and Zamarashkin, Nickolai. Evaluation of optimization methods for nonrigid medical image registration using mutual information and B-splines. *Image Processing, IEEE Transactions on*, 16:12,2879–2890, 2007.

[2] Moré, Jorge J and Wild, Stefan M. Benchmarking derivative-free optimization algorithms. *SIAM Journal on Optimization*, 20:1,172–191, 2009.

[3] Gould, Nicholas and Scott, Jennifer. The State-of-the-Art of Preconditioners for Sparse Linear Least-Squares Problems. *ACM Transactions on Mathematical Software (TOMS)*, 43:4,36, 2017.

[4] Nocedal, Jorge and Wright, Stephen J. Numerical optimization. *Springer New York* , 2, 2006.

[5] SAAD,Yousef Iterative methods for sparse linear systems. *Society for Industrial and Applied Mathematics*, 2003.

Spectral Unmixing with Multiple Dictionaries

Jérémy E. Cohen and Nicolas Gillis
University of Mons

Summary Spectral unmixing aims at recovering the spectral signatures of materials, called endmembers, mixed in a hyperspectral or multispectral image, along with their abundances. A typical assumption is that the image contains one pure pixel per endmember, in which case spectral unmixing reduces to identifying these pixels. Many fully automated methods have been proposed in recent years, but little work has been done to allow users to select areas where pure pixels are present manually or using a segmentation algorithm. Additionally, in a non-blind approach, several spectral libraries may be available rather than a single one, with a fixed (or an upper or lower bound on) number on endmembers to chose from each. In this paper, we propose a multiple-dictionary constrained low-rank matrix approximation model that address these two problems. We propose an algorithm to compute this model, dubbed M2PALS, and its performance is discussed on both synthetic and real hyperspectral images.

Technical contents First, before introducing the new multiple dictionary matrix factorization model, let us recall the single dictionary formulation. Let M be a m-by-n data matrix. In this work, we assume the following:

1. The matrix M admits an approximate low-rank factorization $M \approx AB^T$ of size r, that is, A and B have r columns.

2. The noise is Gaussian. Missing data, stripes or impulsive noise are ignored although common in hyperspectral imaging. The factorization therefore can be written as $X = AB^T + N$ where N is the realization of a random variable following a white Gaussian distribution.

3. Columns of factor A are a subset of columns of the known m-by-d dictionary matrix D.

These assumptions leads to the following low-rank factorization model for M:

$$\begin{aligned} M &= AB^T + N, \\ A &= D\left(:, \mathcal{K}\right) = DS \ \text{ where } S \in \{0,1\}^{d \times r}, \\ \text{vec}(N) &\sim \mathcal{N}\left(0, \sigma^2 I_{nm}\right), \end{aligned} \tag{5}$$

for a given noise variance σ^2. In this model, $\mathcal{K}$ is a set of indices of atoms in D corresponding to the columns of the factor matrix A. The matrix S is a sparse selection matrix which has only 1 non-zero entry per column.

In the self-dictionary setting, a user may want some control over the regions of the HSI where the pure pixels are selected from. On the other hand, hand-picking pixels that seem pure may lead to poor results because the pure-pixel property is difficult to assess visually. Moreover, using a segmentation algorithm to compute homogeneous areas does not provide a good input for usual sparse coding method since the segmented regions would be bundled, and coherence lost. Similarly, when working with external libraries, it is reasonable to assume that only a few material from a specific library should be selected, for instance exactly, at least or at most two spectra related to vegetation among some 20 available vegetation spectra. Lumping the libraries together and running any methods described above would not guarantee such an assignment.
To tackle these issues, we suggest to drop the single dictionary constraint. Using notations from (5), the third working hypothesis is modified to:

3. Columns of matrix A belong to matrices D_i, with d_i the exact number of atoms to be picked in each D_i, $1 \leq i \leq p$.

This new multi-dictionary model becomes:

$$\begin{aligned} M &= AB^T + N, \\ A &= [D_1(:, \mathcal{K}_1), \ldots, D_p(:, \mathcal{K}_p)]\,\Pi, \\ \#\mathcal{K}_i &= d_i \text{ (or } \leq d_i) \text{ and } \sum_{i=1}^{p} \#\mathcal{K}i = r, \\ \mathrm{vec}(N) &\sim \mathcal{N}(0, \sigma I_{nm}), \end{aligned} \tag{6}$$

where $\mathcal{K}_i$ contains the indices of the d_i selected atoms in dictionary D_i for $1 \leq i \leq p$, and Π is the permutation matrix that matches factors to their corresponding atoms. Algorithm 1 is used to compute the proposed model. We will illustrate the effectiveness of this approach on several real-world hyperspectral images compared to state-of-the-art algorithms.

Algorithm 1 M2PALS

Input: Data matrix M, dictionaries $D_1, \ldots, D_p$, exact (or upper bound on the) number of atoms to pick in each library $d_1, \ldots, d_p$, initials factors A and B.
Output: Estimated factors A and B, selected atoms $\mathcal{K}_i$ $(1 \leq i \leq p)$.
while the convergence criterion is not met **do**
Least squares estimate of A: $A = \mathrm{argmin}_X ||M - XB^T||_F$
Solve the assignment problem: Find the $\mathcal{K}_i$'s to match A the best using the available dictionaries and update $A = [D_1(:, \mathcal{K}_1), \ldots, D_p(:, \mathcal{K}_p)]\Pi$
Least squares estimate of B: $B = \mathrm{argmin}_Y ||M - AY^T||_F$
end while

The maximum covering cycle problem

W. Vancroonenburg[a,b]
[a]Research Foundation Flanders (FWO-Vlaanderen)
[b]KU Leuven, Department of Computer Science, CODeS & imec
Technology Campus Ghent, Gebroeders De Smetstraat 1, 9000 Gent
e-mail: wim.vancroonenburg@kuleuven.be

A. Grosso[c], F. Salassa[d]
[c]Universitá degli Studi di Torino, Torino, Italy
[d]Politecnico di Torino, Torino, Italy

The *maximum covering cycle problem* (MCCP) [1] deals with the problem of finding a simple cycle C in an undirected graph $G(V, E)$ such that the number of vertices that are in C or adjacent to a vertex in C is maximal. In this contribution, the problem is formulated as an Integer Linear Programming (ILP) model and it is shown that the problem is NP-Hard. Next, an iterative constraint generation procedure (detailed in [1]) using a relaxed ILP formulation for the MCCP is presented, that allows to find the optimal solution for a given graph in a finite number of iterations. This procedure is subsequently enhanced by means of a set of heuristics that diversify and intensify the added cycle constraints in order to reduce the number of iterations required to find the optimal solution. We present computational results on a diverse set of instances that highlight the effectiveness of the added heuristic methods, and conclude with further research directions.

References

[1] Grosso, A., Salassa, F. & Vancroonenburg, W., *Searching for a cycle with maximum coverage in undirected graphs*, Optimization Letters (2016) 10(7): 1493–1504., https://doi.org/10.1007/s11590-015-0952-x

Linear and quadratic reformulation techniques for nonlinear 0–1 optimization problems

Elisabeth Rodríguez-Heck
QuantOM, HEC Liège - Management School of the University of Liège
e-mail: elisabeth.rodriguezheck@uliege.be

Yves Crama
QuantOM, HEC Liège - Management School of the University of Liège
e-mail: yves.crama@uliege.be

Nonlinear unconstrained optimization in binary variables is a very general class of problems that has many applications in classical operations research problems such as production planning or supply chain management, but also in engineering and computer science topics such as computer vision or machine learning [1, 2]. Nonlinear optimization problems are currently of great interest to the mathematical programming community; recent papers from different authors have focused on this type of problems and there exists a strong trend towards a better understanding of the nonlinear case.

We consider methodological aspects of the resolution of nonlinear optimization problems in binary variables. More precisely we examine resolution techniques based on linear and quadratic reformulations of nonlinear problems by introducing artificial variables. This approach attempts to draw benefit from the extensive literature on integer linear and quadratic programming and has been recently considered by different authors [3, 4, 5, 6].

In the framework of linear reformulations, we defined a class of inequalities that improve a well-known linearization technique. The use of these inequalities leads to very good computational results for several classes of problems, especially for a class of instances inspired from the image restoration problem in computer vision, for which we obtain up to ten times faster computing times with respect to a commercial solver when using our inequalities [7]. Concerning quadratic reformulations, we recently obtained a result that introduces a quadratic reformulation reducing the smallest number of required artificial variables from linear to logarithmic for monomials with a positive coefficient, which is a very simple but fundamental class of functions [8]. Indeed, the definition of a quadratic reformulation for monomials with a positive coefficient together with a reformulation for negative monomials allows us to reformulate any multilinear optimization problem in binary variables.

Finally, we present the results of some preliminary computational experiments comparing the performance of several linear and quadratic reformulation techniques. For both types of methods, it seems that reformulation techniques that take into account the structure of the original nonlinear problem are very effective computationally. More precisely, reformulation techniques that take into account the interaction between monomials seem to be particularly promising, especially for well-structured

problems. An interesting open question is to gain a better understanding of why these procedures seem to work better computationally on our classes of instances than reformulations that only focus on the monomials term by term, without taking into account monomial interactions.

References

[1] E. Boros and P. L. Hammer. Pseudo-Boolean optimization. *Discrete Applied Mathematics*, 123(1):155–225, 2002.

[2] Y. Crama and P. L. Hammer. *Boolean Functions: Theory, Algorithms, and Applications*. Cambridge University Press, New York, N. Y., 2011.

[3] M. Anthony, E. Boros, Y. Crama, and A. Gruber. Quadratic reformulations of nonlinear binary optimization problems. *Mathematical Programming*, 162(1-2):115–144, 2017.

[4] C. Buchheim and G. Rinaldi. Efficient reduction of polynomial zero-one optimization to the quadratic case. *SIAM Journal on Optimization*, 18(4):1398–1413, 2007.

[5] A. Del Pia and A. Khajavirad. A polyhedral study of binary polynomial programs. *Mathematics of Operations Research*, 42(2):389–410, 2017.

[6] A. Fischer, F. Fischer, and T. S. McCormick. Matroid optimisation problems with nested non-linear monomials in the objective function. *Mathematical Programming*, 2017. Published online.

[7] Y. Crama and E. Rodríguez-Heck. A class of valid inequalities for multilinear 0-1 optimization problems. *Discrete Optimization*, 25:28–47, 2017.

[8] E. Boros, Y. Crama, and E. Rodríguez-Heck. Quadratizations of symmetric pseudo-boolean functions: sub-linear bounds on the number of auxiliary variables. Presented at ISAIM 2018, Fort Lauderdale, USA, 2018.

Unit Commitment under Market Equilibrium Constraints

Jérôme De Boeck
Université libre de Bruxelles, Belgium
e-mail: jdeboeck@ulb.ac.be

Luce Brotcorne
INOCS, INRIA Lille Nord-Europe, France
e-mail: luce.brotcorne@inria.fr

Fabio D'Andreagiovanni
Université de Technologie de CompiÃÍgne, France
e-mail: f.dandreagiovanni@gmail.com

Bernard Fortz
Université libre de Bruxelles, Belgium
e-mail: bernard.fortz@ulb.ac.be

The classical Unit Commitment problem (UC) can be essentially described as the problem of establishing the energy output of a set of generation units over a time horizon, in order to satisfy a demand for energy, while minimizing the cost of generation and respecting technological restrictions of the units (e.g., minimum on/off times, ramp up/down constraints). The UC is typically modelled as a (large scale) mixed integer program and its deterministic version, namely the version not considering the presence of uncertain data, has been object of wide theoretical and applied studies over the years.
Traditional (deterministic) models for the UC assume that the net demand for each period is perfectly known in advance, or in more recent and more realistic approaches, that a set of possible demand scenarios is known (leading to stochastic or robust optimization problems).
However, in practice, the demand is dictated by the amounts that can be sold by the producer at given prices on the day-ahead market. One difficulty therefore arises if the optimal production dictated by the optimal solution to the UC problem cannot be sold at the producer's desired price on the market, leading to a possible loss. Another strategy could be to bid for additional quantities at a higher price to increase profit, but that could lead to infeasibilities in the production plan.
Our aim is to model and solve the UC problem with a second level of decisions ensuring that the produced quantities are cleared at market equilibrium. In their simplest form, market equilibrium constraints are equivalent to the first-order optimality conditions of a linear program. The UC in contrast is usually a mixed-integer nonlinear program (MINLP), that is linearized and solved with traditional Mixed Integer (linear) Programming (MIP) solvers. Taking a similar approach, we are faced to a bilevel optimization problem where the first level is a MIP and the second level linear.
In this talk, as a first approach to the problem, we assume that demand curves and offers of competitors in the market are known to the operator. This is a very

strong and unrealistic hypothesis, but necessary to develop a first model. Following the classical approach for these models, we present the transformation of the problem into a single-level program by rewriting and linearizing the first-order optimality conditions of the second level. Then we present some preliminary results on the performance of MIP solvers on this model. Our future research will focus on strengthening the model using its structure to include new valid inequalities or to propose alternative extended formulations, and then study a stochastic version of the problem where demand curves are uncertain.

A decade of academic research on warehouse order picking: trends and challenges

B. Aerts, T. Cornelissens, K. Sörensen
University of Antwerp, Dept Engineering Mgmt, ANT/OR Operations Research Group
e-mail: babiche.aerts@uantwerpen.be

Warehouses play a key role in supply chain operations. Due to the recent trends in, e.g., e-commerce, the warehouse operational performance is exposed to new challenges such as the need for faster and reliable delivery of small orders. Order picking, defined as the process of retrieving stock keeping units from inventory to fulfil a specific customer request, is seen as the most labor-intensive activity in a warehouse and is therefore considered to be an interesting area of improvement in order to deal with the aforementioned challenges. A literature review by de Koster et al. offers an overview of order picking methods documented in academic literature up until 2007. We take this publication as a starting point and review developments in order picking systems that have been researched in the past ten years. The aim of our presentation is to give an overview of the current state of the art models and to identify trends and promising research directions in the field of order picking.

Iterated local search algorithm for solving operational workload imbalances in order picking

S. Vanheusden, T. van Gils,
K. Ramaekers, A. Caris and K. Braekers
Hasselt University, research group logistics
e-mail: sarah.vanheusden@uhasselt.be, teun.vangils@uhasselt.be,
katrien.ramaekers@uhasselt.be, an.caris@uhasselt.be,
kris.braekers@uhasselt.be

Warehouses play an important role in the supply chain. For the fulfilment of orders, warehouse operations need to satisfy basic requirements such as receiving, storing, picking and shipping stock keeping units (SKU). Order picking, where goods are retrieved from storage or buffer areas to fulfil incoming customer orders, is considered to be the most costly activity. Up to 50% of the total warehouse operating costs can be attributed to this process [2].

Trends such as the upcoming e-commerce market, shortened product life cycles, and greater product variety, expose order picking activities to new challenges. Order pickers need to fulfil many small orders for a large variety of SKUs due to a changed order behaviour of customers. Furthermore, warehouses accept late orders from customers to stay competitive and preserve high service levels. This results in extra difficulties for order picking operations, as more orders need to be picked and sorted in shorter and more flexible time windows [1]. Warehouse managers therefore experience difficulties in balancing the workload of the order pickers on a daily basis, resulting in peaks of workload during the day [4].

The problem of peaks in daily workload, that is examined in this study, originates from a large international B2B warehouse located in Belgium. The warehouse is responsible for the storage and distribution of automotive spare parts. Spare part warehouses are characterized by customer orders that can be grouped based on their destination. An order set refers to a group of order lines with a common destination that is picked in a single zone. All order lines of a common destination from all pick zones are referred to as the order lines from a shipping truck. Deadlines of order lines are determined by the shipping truck and resulting schedule of shipping trucks (i.e. shipping schedule). Each shipping truck can consist of multiple order sets (i.e., a single order set for each order picking zone). The assignment of order sets to shipping trucks as well as the shipping schedule are assumed to be fixed at the operational level. This fixed shipping schedule often results in workload peaks during the day, as order patterns vary across customers (e.g., varying number of orders and customers, varying order time and resulting available time to pick orders) [4].

Balancing the workload in an order picking system can be addressed from several perspectives. Instead of putting the focus on strategic or tactical solutions, the emphasis of this study is on the operational level, in order to avoid peaks in the number of picked order lines every hour of the day. The objective function aims

to minimise the variance of planned order lines over all time slots, for each pick zone [4]. Solving instances of realistic size to optimality in a reasonable amount of computation time does not seem feasible due to the complex nature of the operational workload imbalance problem (OWIP). The objective of this study is the development of an iterated local search (ILS) algorithm to solve OWIP in a parallel zoned manual order picking system.
The developed model can be used by warehouse managers and supervisors as a simulation tool to plan order sets more accurately during the day, in this way, avoiding peaks in workload. The proposed ILS is able to provide fast and accurate results so that it can be used as a supporting tool in practice. This tool can advise warehouse managers in negotiations with customers on changes in cut-off times for order entry and shipping schedules to further reduce workload imbalances. Additionally, the balancing of workloads results in a more stable order picking process and overall productivity improvements for the total warehouse operations.

References

[1] De Koster, R., Le-Duc, T., Roodbergen, K.J., 2007. Design and control of warehouse order picking: A literature review. European Journal of Operational Research. 182, 481-501.

[2] Davarzani, H., and Norrman, A., 2015. Toward a relevant agenda for warehousing research: literature review and practitioners' input. Logistics Research. 8, 1-18.

[3] Van Gils, T., Ramaekers, K., Caris, A., De Koster, R., 2018. Designing efficient order picking systems by combining planning problems: State-of-the-art classification and review. European Journal of Operational Research (in press). 000, 1-15.

[4] Vanheusden, S., van Gils, T., Ramaekers, K., Caris, A., 2017. Reducing workload imbalance in parallel zone order picking systems. Proceedings of the International Conference on Harbour, Maritime & Multimodal Logistics Modelling and Simulation. 18-20

Scheduling container transportation with capacitated vehicles through conflict-free trajectories: A local search approach

Emmanouil Thanos
KU Leuven, Department of Computer Science, CODeS & imec
e-mail: emmanouil.thanos@kuleuven.be

Tony Wauters
KU Leuven, Department of Computer Science, CODeS & imec
e-mail: tony.wauters@kuleuven.be

Greet Vanden Berghe
KU Leuven, Department of Computer Science, CODeS & imec
e-mail: greet.vanden.berghe@kuleuven.be

Internal transportation procedures constitute a critical link in container terminal operations. Space limitations lead to complex traffic networks where operating vehicles block each other's trajectories. Capacity, precedence and stacking constraints impose further difficulties in scheduling containers' pickups and dispatches. In settings involving large numbers of dynamic requests, the additional requirement of making instantaneous decisions constitutes a significant scientific challenge for decision support systems.
This study proposes an integrated model and a Local Search (LS) algorithm for real-time scheduling of transportation requests in a container terminal. Terminal's detailed layout is modeled as a network graph. The model includes stacks of containers with different physical properties. Capacitated vehicles are integrated alongside their loading, unloading and transfer operations. Various real-world movement restrictions and conflict rules are incorporated in the network's edges and paths, in addition to requests' precedence and containers stacking constraints.
In the developed LS framework each LS node indirectly defines a routing schedule by gradually routing all assigned vehicles. Waiting times are imposed to resolve conflicts and maintain safety distances, while satisfying all capacity, precedence and piling restrictions. LS neighborhoods permit numerous possible decisions for modifying an existing schedule, including execution order, vehicles assignments to requests, merging or splitting distinct operations, dispatching locations and path selections. Feasible schedules are evaluated in terms of their total makespan. The integrated approach is applied to the real-world container yard and transport requests of a Belgian terminal. Produced solutions are compared against those currently employed at the terminal and experimental evaluation shows that the proposed algorithm significantly reduces the operations' execution time for instances of various input sizes.

The Robust Machine Availability Problem

Guopeng Song
KU Leuven, ORSTAT
e-mail: guopeng.song@kuleuven.be

Daniel Kowalczyk
KU Leuven, ORSTAT
e-mail: daniel.kowalczyk@kuleuven.be

Roel Leus
KU Leuven, ORSTAT
e-mail: roel.leus@kuleuven.be

We define and solve the robust machine availability problem in a parallel-machine environment, which aims to minimize the required number of identical machines that will still allow to complete all the jobs before a given due date. The deterministic variant of the proposed problem is essentially equivalent to the bin packing problem. A robust formulation is presented, which preserves a user-defined robustness level regarding possible deviations in the job durations. For better computational performance, a branch-and-price procedure is proposed based on a set covering formulation, with the robust counterpart formulated for the pricing problem. Zero-suppressed binary decision diagrams (ZDDs) are introduced for solving the pricing problem, in order to tackle the difficulty entailed by the robustness considerations as well as by extra constraints imposed by branching decisions. Computational results are reported, showing the effectiveness of a pricing solver with ZDDs compared with a MIP solver.

Optimizing line feeding under consideration of variable space constraints

Nico André Schmid
Ghent University, Department of Business Informatics and Operations Management
e-mail: nico.schmid@ugent.be

Veronique Limère
Ghent University, Department of Business Informatics and Operations Management
Flanders Make, Lommel, Belgium
e-mail: veronique.limere@ugent.be

Nowadays, assembly systems are used for the assembly of a number of models, often being mass-customized, which increases the number of parts required for assembly. Due to a rise of part numbers, space scarcity at the line is a problem occouring frequently in practice. Therefore, parts must not only be fed to the line in a cost efficient manner, but space constraints need to be taken into account as well [3]. The assembly line feeding problem (ALFP) deals with the assignment of parts to line feeding policies in order to reduce costs and obtain feasible solutions [1].
Within this paper, we examine all known distinct line feeding policies, namely line stocking, kanban, sequencing and kitting (stationary and traveling kits). There is, to the best of our knowledge, no research conducted, including more than three line feeding policies in a single model [4]. However, although using different line feeding policies might solve the problem of space scarcity, it might also be an expensive solution and more cost effective alternatives might exist. Therefore, in our approach we model the available storing space at every single station of the line as a variable, while the overall space available for the complete line stays constrained [2]. This way, we can investigate if space sharing between stations leads to cheaper solutions. In our research, we model this problem as a MILP problem including a representation of costs and constraints caused by the necessary line feeding processes, being storage, preparation, transportation, line side presentation and usage. By incorporating variable space constraints for every station, we provide a decision model minimizing the overall costs for line feeding in assembly systems, since rigid space constraints at the BoL usually lead to more expensive line feeding policies. In contrast, variable space constraints enable balancing of unequal space requirements at different stations, which allows cheaper line feeding policies to be selected. The proposed model can be solved by standard solvers, such as CPLEX or Gurobi.
Furthermore, we propose a data generation algorithm using case study data being capable of creating data sets with multiple products, although in the case study one model was produced on the line. This paves the way to examine the effect of different products on space borrowing and line feeding policy selection.
Finally, we use the decision model and multiple generated data sets to run experiments and analyze the results for decision patterns for line feeding policy and space

borrowing decisions. First results, such as cost effects of space borrowing or the influence of demand or part volume on line feeding policy decisions, will be shown.

References

[1] Bozer, Y. A. and McGinnis L. F. (1992). Kitting versus line stocking: A conceptual framework and a descriptive model. *International Journal of Production Economics*, 28(1):1–19.

[2] Hua, S. Y. and Johnson, D. J. (2010). Research issues on factors influencing the choice of kitting versus line stocking *International Journal of Production Research*, 48(3):779–800.

[3] Limere, V., van Landeghem, H., and Goetschalckx, M. and McGinnis L. F. (2015). A decision model for kitting and line stocking with variable operator walking distances. *Assembly Automation*, 35(1):47–56.

[4] Sali, M. and Sahin E. (2016). Line feeding optimization for just in time assembly lines: An application to the automotive industry. *International Journal of Production Economics*, 174:54–67.

Scheduling Markovian PERT networks to maximize the net present value: New results

Ben Hermans, Roel Leus
KU Leuven, ORSTAT, Faculty of Economics and Business
e-mail: ben.hermans@kuleuven.be

We study the problem of scheduling a project so as to maximize its expected net present value when task durations are exponentially distributed. Based on the structural properties of an optimal solution we show that, even if preemption is allowed, it is not necessary to do so. Next to its managerial importance, this result also allows for a new algorithm which improves on the current state of the art [1] with several orders of magnitude.

Both the algorithm of Creemers et al. [1] and our new approach are based on the Markov chain of Kulkarni & Adlakha [2] and use a stochastic dynamic program to identify an optimal solution. The key difference between the two algorithms is that the derived structural properties allow to reduce the dynamic program's state space significantly. Table 1 compares the computation time in seconds[1] for the two approaches, where n equals the number of tasks in the project and the order strength OS is a measure for the number of precedence constraints.

	Creemers et al. [1]			New algorithm		
n	OS=0.80	OS=0.60	OS=0.40	OS=0.80	OS=0.60	OS=0.40
20	0.00	0.01	0.27	0.00	0.00	0.00
40	0.02	3.84	1,213.43	0.00	0.01	0.31
60	0.42	1,288.93		0.01	0.29	58.96
80	6.06	25,628.35		0.03	5.72	
100	98.09			0.12	151.44	
120	2,495.16			1.18	1,473.39	

Table 1: Average CPU time in seconds for solved instances.

(The research was funded by a PhD Fellowship of the Research Foundation – Flanders.)

References

[1] S. Creemers, R. Leus, and M. Lambrecht. Scheduling Markovian PERT networks to maximize the net present value. *Operations Research Letters*, 38(1): 51–56, 2010.

[1] Using an Intel Core i7-4790 3.60 GHz processor and an upper bound of 1 GB of RAM.

[2] V. G. Kulkarni and V. G. Adlakha. Markov and Markov-regenerative PERT networks. *Operations Research*, 34(5):769–781, 1986.

Log-determinant Non-Negative Matrix Factorization via Successive Trace Approximation

Andersen Ang
Department of Mathematics and Operational Research
Faculté Polytechnique, Université de Mons
Rue de Houdain 9, 7000 Mons, Belgium
manshun.ang@umons.ac.be

Nicolas Gillis
Department of Mathematics and Operational Research
Faculté Polytechnique, Université de Mons
Rue de Houdain 9, 7000 Mons, Belgium
nicolas.gillis@umons.ac.be

Non-negative matrix factorization (NMF) is the problem of approximating a non-negative matrix X as the product of two smaller nonnegative matrices W and H so that X = WH. In this talk, we consider a regularized variant of NMF, with a log-determinant (logdet) term on the Gramian of the matrix W. This term acts as a volume regularizer: the minimization problem aims at finding a solution matrix W with low fitting error and such that the convex hull spanned by the columns of W has minimum volume. The logdet of the Gramian of W makes the columns of W interact in the optimization problem, making such logdet regularized NMF problem difficult to solve. We propose a method called successive trace approximation (STA). Based on a logdet-trace inequality, STA replaces the logdet regularizer by a parametric trace functional that decouples the columns on W. This allows us to transform the problem into a vector-wise non-negative quadratic program that can be solved effectively with dedicated methods. We show on synthetic and real data sets that STA outperforms state-of-the-art algorithms.

Audio Source Separation Using Nonnegative Matrix Factorization

Valentin Leplat
UMons, Department of Mathematics and Operational Research
e-mail: valentin.leplat@umons.ac.be

Nicolas Gillis
UMons, Department of Mathematics and Operational Research
e-mail: nicolas.gillis@umons.ac.be

Xavier Siebert
UMons, Department of Mathematics and Operational Research
e-mail: xavier.siebert@umons.ac.be

The audio source separation concerns the techniques used to extract unknown signals called sources from a mixed signal. In this talk, we assume that the audio signal is recorded with a single microphone. Considering a mixed signal composed of various audio sources, the blind audio source separation consists in isolating and extracting each of the source on the basis of the single recording. Usually, the only known information is the number of estimated sources present in the mixed signal.

The blind source separation problem is said to be underdetermined as there are fewer sensors (only one in our case) than sources. It then appears necessary to find additional information to make the problem well posed. The most common technique used for this kind of problem is to get some form of redundancy in the mixed signal in order to make it overdetermined. This is typically done by representing the signal in the time and frequency domains simultaneously (splitting the signals into overlapping time frames). The time-frequency representation of a signal highlights two fundamental properties: the sparsity and the redundancy (low-rank) of the data. These two fundamental properties led sound source separation techniques to integrate algorithms such as nonnegative matrix factorization (NMF).

The talk is organized as follows. First, several mathematical notions, such as the short time Fourier transform, are briefly presented in order to provide the background to understand the use of NMF applied to audio sources. Second, the most common separation models and their algorithms are presented; in particular the use of the β-divergence for the NMF objective function.
Then, we present our three contributions:

1. The implementation of a convex cost function integrating the low-rank and sparse properties of the data.
2. A new variant of β-NMF integrating a penalty term in the cost function in order to promote activation matrices to have higher sparsity.

3. The integration of an additional penalty term in the optimization problem in order to make the evolutions of activations smoother, namely, the use of smooth Itakura-Saito NMF.

These methods are tested using real audio signals. Their performance, in terms of the quality of the estimated signals, are compared in order to highlight their advantages and to identify their weaknesses. Further work will try to tackle these weaknesses by improving our models, in particular by incorporating the use of artificial neural networks.

Orthogonal Joint Sparse NMF for 3D-Microarray Data Analysis

Flavia Esposito
University of Bari Aldo Moro, Department of Mathematics
e-mail: flavia.esposito@uniba.it

Nicolas Gillis Université de Mons
Nicoletta Del Buono University of Bari Aldo Moro

Data generated by microarray experiments consist from thousands to millions of biological variables and they pose many challenges in their analysis such as discovering and capturing valuable knowledge to understand biological processes and human disease mechanisms. The challenging increases when the time evolution of the observed features is introduced. The 3D microarray, generally known as gene-sample-time microarray, joint information collected by the most famous 2D microarrays measuring gene expression levels among different samples on different time points. Its data analysis is useful in several biomedical applications, like monitor drug treatment responses of patient over time in pharmacogenomics studies. Many statistical and data analysis tools could be applied to bring out useful information from so large datasets. Nonnegative Matrix Factorization (NMF) among all, for its natural nonnegative constraints, is demonstrated to be able to extract from 2D microarrays relevant information on specific genes involved in the particular biological process. Nevertheless, although multilinear algebra decompositions were adopted to tackle 3D microarrays to reduce their dimensionality and extract relevant features, we prefer to focus on a new NMF problem for its tested capabilities. In this work, we propose a new NMF model, namely Orthogonal Joint Sparse NMF (OJSNMF), to extract relevant information from 3D microarrays containing the time evolution of a 2D microarray, that adding additional constraints could be able to enforce some biological proprieties being an useful tool to further biological analysis. We developed some updates rules, tested and compared our approach on both synthetic and real dataset.

This is a joint work with Nicoletta Del Buono, Department of Mathematics, University of Bari Aldo Moro; Angelina Boccarelli, Angelo Vacca and Maria Antonia Frassanito, Department of Biomedical Science and Human Oncology University of Bari Medical School and Mauro Coluccia, Department of Pharmacology, University of Bari Aldo Moro.

Intermodal Rail-Road Terminal Operations

Michiel Van Lancker Greet Vanden Berghe
Tony Wauters
KU Leuven, Department of Computer Science, CODeS
e-mail: michiel.vanlancker@cs.kuleuven.be

The intermodal rail-road terminal (IRRT) is a logistics facility where freight units such as containers, swap bodies and semi-trailers are transshipped between trains and trucks. IRRTs typically consist of a set of parallel rail tracks for trains to park, a gate for truck registration, several lanes parallel to the tracks for truck serving and a container yard along the tracks for temporary container storage. Multiple rail-mounted gantry cranes (RMG) arch over the tracks, lanes and yard. Throughout the day, trains enter the terminal and are served by the RMGs. When no trains are available, RMGs can reshuffle containers already located in the yard with the objective of improving future storage access times. In parallel with train serving and yard reshuffling, trucks arrive at the terminal and need to be served by the RMGs within a certain time window. Typically, train arrival times are known in advance, but truck arrival times are uncertain. The operational performance of an IRRT may be measured by the number of (unique) container transshipments within a given time interval. One of the primary bottlenecks is that many transshipments involving trucks cannot be executed between trucks and trains directly, but instead take place between trucks and temporary storage facilities, thus introducing an extra transshipment.

Several factors complicate IRRT operations. First of all, intermodal transport accommodates different freight unit types, each with their own specific needs regarding storage and transport. Examples include electricity requirements for refrigerated containers, or the fact that swap bodies cannot be stacked. Secondly, while containers can be placed on a wagon in multiple ways, each placement of containers requires a set of pins on the wagon to be manually configured accordingly. Train loading is further constrained by weight limitations for both wagons and locomotives. All the aforementioned complexities give rise to the Train Load Planning Problem (TLPP). In the TLPP, as described in [2], freight units require assignment to wagons and wagon configurations must be selected, while maximizing the utilization of the train and minimizing terminal costs. Besides the TLPP, a second problem arises, related to scheduling the RMGs. Given that all RMGs are mounted on a single rail, they cannot cross each other, thus introducing non-crossing constraints. This is often solved by assigning cranes to fixed operational zones. Such a strategy might, however, result in sub-optimal crane scheduling policies. In the temporary storage yard, containers are stored in rows and columns of stacks, resulting in a three-dimensional block structure which is only accessible from the top down. Thus, if a container stored below an other container requires transshipment, an extra transshipment must be executed to first remove the top container. In addition to determining

which container should be put where and on which train (TLPP) and how crane operations should be executed, several other decisions must be made. Not only should both trains and trucks be assigned parking positions with serving time in mind, but the management of the temporary storage yard significantly influences future train serving times. Finally, in the examined use case, IRRT operations contain a degree of uncertainty related to train structure (in what order do the wagons of the train arrive?) and truck arrival (will a truck be able to respect its time slot?). While such uncertainty does not lead to additional optimization problems, it does introduce a probabilistic factor in most of the operational decisions. The IRRT may therefore be characterized as a complex operational system combining multiple challenging optimization problems. For an extensive survey on IRRT operations, [1] can be referred to.

A problem relating to the IRRT is studied, in which containers are transshipped between multiple trains by means of two RMGs. The goal is to minimize the time needed to transship all containers to trains with the appropriate destination. Trains arrive in groups and are already assigned to tracks, but not to destinations. The cranes have fixed, overlapping operational zones. Temporary storage is simplified to a quay with infinite capacity. Different load configurations of wagons are possible, but no weight restrictions or costs relating to configuration changes are present. In [3], this problem is solved by decomposing it into subproblems and solving each subproblem separately. Preliminary results are presented. Additionally, an extensive problem description of IRRT operations is formulated.

References

[1] Boysen, N., Fliedner, M., Jaehn, F., and Pesch, E. (2013). A survey on container processing in railway yards. *Transportation Science*, 47(3):312–329.

[2] Bruns, F. and Knust, S. (2012). Optimized load planning of trains in intermodal transportation. *OR Spectrum*, 34(3):511–533.

[3] Souffriau, W., Vansteenwegen, P., Vanden Berghe, G., and Van Oudheusden, D. (2009). *Variable Neighbourhood Descent for Planning Crane Operations in a Train Terminal*, pages 83–98. Springer Berlin Heidelberg, Berlin, Heidelberg.

Reducing Picker Blocking in a Real-life Narrow-Aisle Spare Parts Warehouse

T. van Gils [(a)], K. Ramaekers[(a)], and A. Caris[(a)]
(a) Hasselt University
e-mail: teun.vangils@uhasselt.be, katrien.ramaekers@uhasselt.be, an.caris@uhasselt.be

New market developments, such as e-commerce, globalization, increased customer expectations and new regulations, have intensified competition among warehouses and force warehouses to handle a large number of small orders within tight time windows [1]. In order to differentiate from competitors with respect to customer service, warehouses aim at providing quick deliveries to customers. Consequently the remaining time to handle orders is reduced. Moreover, the order behaviour of e-commerce customers is characterised by many orders consisting of only a limited number of order lines [2].

In order to fulfil customer orders, order pickers should retrieve the ordered products from storage locations (i.e. order picking). In this paper, two of the main operational planning problems are studied in a narrow-aisle order picking system: storage location assignment (i.e. determining the physical location at which incoming products are stored) and order picker routing (i.e. determining the sequence of storage locations to visit to compose customer orders) [1].

Order picking management has been identified as an important and complex planning operation as a consequence of the existing relations among planning problems and the existing trade-off among decisions [3]. Narrow-aisle picking systems are designed to increase the storage capacity, but multiple order pickers may require to enter the same aisle which results in blocking of order pickers. Moreover, most storage location assignment policies aim to increase the pick density by assigning fast moving stock keeping units (SKUs) to storage locations closely located to the depot in order to reduce the order picker travel time. High pick densities in certain order picking areas increase the probability of picker blocking [4]. A wide range of routing methods (e.g., traversal, return, largest gap) have been evaluated in literature in a system with a single order picker, focusing on reducing either travel time or travel distance. In practice, multiple order pickers are working in the same order picking area to retrieve items. Efficient methods have been proposed to dynamically change order picking routes during the pick tour for multiple order pickers [5]. These complex methods require innovative automation technologies to implement the dynamic order picker routing methods in practice to minimise travel time and picker blocking time simultaneously. Due to this complexity, straightforward routing methods are still widely used in practice [3].

The objective of this study is to analyse the joint effect of the two main operational order picking planning problems, storage location assignment and order picker routing, on order picking time, including travel time and wait time due to picker blocking.

Multiple combinations of storage and straightforward routing policies are simulated in a real-life narrow-aisle order picking system with the aim of reducing order picking time. Order picking systems in previous research are subject to a large number of assumptions to simplify operations, such as ignoring picker blocking [6, 3] and low-level storage locations [4, 6, 3]. Our study narrows this gap between practice and academic research by simulating a real-life business-to-business (B2B) warehouse storing automobile spare parts in a narrow-aisle high-level order picking system, which is a convenient system to store spare parts.
To the best of our knowledge, we are the first to analyse the joint effect of storage and routing policies on the trade-off between travel time and picker blocking time in high-level order picking systems. The main contributions of this paper are managerial insights into the trade-off between reducing travel time and picker blocking by varying storage location assignment and routing policies in a real-life narrow-aisle picking system. As the simulation experiments have focused on operational order picking planning problems only, the proposed combinations are rather easy to implement and result in substantial performance benefits (-6.5% in pick time for the case). Results of this study can be used by warehouse managers to increase order picking efficiency in order to face the new market developments.

References

[1] Van Gils, T., Ramaekers, K., Caris, A., De Koster, R., 2018. Designing efficient order picking systems by combining planning problems: State-of-the-art classification and review. European Journal of Operational Research (in press). 000, 1-15.

[2] De Koster, R., Le-Duc, T., Roodbergen, K.J., 2007. Design and control of warehouse order picking: A literature review. European Journal of Operational Research. 182, 481-501.

[3] Van Gils, T., Ramaekers, K., Braekers, K., Depaire, B., Caris, A., 2018. Increasing Order Picking Efficiency by Integrating Storage, Batching, Zone Picking, and Routing Policy Decisions. International Journal of Production Economics (in press). 000, 1-21.

[4] Pan, J. C.-H., Wu, M.-H., 2012. Throughput analysis for order picking system with multiple pickers and aisle congestion considerations. Computers & Operations Research. 39, 1661-1672.

[5] Chen, F., Wang, H., Xie, Y., Qi, C., 2016. An ACO-based online routing method for multiple order pickers with congestion consideration in warehouse. Journal of Intelligent Manufacturing. 27, 389-408.

[6] Petersen, C.G., Aase, G., 2004. A comparison of picking, storage, and routing policies in manual order picking. International Journal of Production Economics. 92, 11-19.

An analysis on the destroy and repair process in large neighbourhood search applied on the vehicle routing problem with time windows

J. Corstjens, A. Caris, and B. Depaire
UHasselt, Research Group Logistics
e-mail: jeroen.corstjens@uhasselt.be

For many years experimenters have successfully resorted to heuristic algorithms when solving complex optimisation problems, ranging from construction methods to high-level metaheuristic frameworks. Whenever such a method is presented, the added value it delivers is shown in a competitive evaluation context. This means that the algorithm is applied to solve the instances of one or several well-known benchmark problem sets after which its performance on these instances is compared to the performance of other algorithms that solved the same instances. The aim is to show that the presented algorithm performs better, in the sense of a better solution quality, computation time or trade-off of both performance measures. One rarely sees a thorough investigation of how an algorithm establishes its performance. What elements of the algorithm contribute the most to performance? How does this contribution depend on the problem characteristics? What can we learn about any performance difference observed between distinct parameter configurations? Such insights are necessary to truly understand algorithmic behaviour [5].

In recent years, there has been increased focus on identifying the algorithm elements that are most relevant to performance ([2, 3, 4]). We want to go a step further and understand why these elements work well or not, enabling us to explain why it is that two parameter configurations or two distinct algorithms differ in the performance they obtain. That is the type of question we aim to address when experimenting with heuristic algorithms. In a case study we analyse a large neighbourhood search (LNS) algorithm applied on instances of the vehicle routing problem with time windows (VRPTW). A first exploratory analysis exposed several patterns raising questions that are to be answered in new consecutive experiments [1]. One observed pattern concerns certain combinations of destroy and repair operators. More specifically, a first experimental analysis showed that removing customers at random from a solution each iteration leads to a better performance than removing geographical clusters of customers if these customers are to be reinserted using a regret measure that takes into account which customers are more difficult to insert and should be prioritised. In this follow-up research we wish to understand why there is a performance difference.

Searching for explanations, the focus shifts from a complete large neighbourhood search framework to a single destroy and repair iteration. A data set of 10,000 single iteration observations is generated, consisting of 200 VRPTW instances and

50 parameter settings per problem instance. A parameter setting is interpreted in this experiment as a combination of a single destroy operator with a single repair operator. Solutions are destroyed using either random or related removal, while they are repaired using a regret-k operator with k equal to 2,3 or 4.

The observed performance difference between the destroy operators in the LNS experiment is also observed in the new experiment. Given this confirmation, the search for explanations started by decomposing the destroy and repair process. The analysis showed that the removal of a cluster of customers that are geographically nearby is detrimental for the repair phase as it reduces the number of alternative routes available. The latter is crucial for a regret operator if it wants to make the best insert decisions. Even more, some customer no longer have any feasible alternative in one of the existing routes and are considered to be isolated cases. It is found that prioritising these customers by inserting them in individual routes reduces the majority of the performance gap between random and related removal. The creation of individual routes early on in the repair phase benefits all customers that are to be inserted as it increases the number of available alternatives. Hence, a regret operator will make a better estimation of customer difficulty and consequently a better prioritisation if each individual customer has existing routes nearby in which it can be feasibly inserted.

We still observe a small significant performance difference between random and related removal, but before looking for further explanations we will perform an experiment on a complete large neighbourhood search algorithm. In this experiment it will be tested whether the findings for the one iteration experiment also hold when performing a normal LNS experiment that runs many more iterations.

References

[1] Corstjens, J., Depaire, B., Caris, A., & Sörensen, K. (2016). Analysing metaheuristic algorithms for the vehicle routing problem with time windows. In Verolog 2016 proceedings, 89.

[2] Fawcett, C., & Hoos, H. (2015). Analysing differences between algorithm configurations through ablation. Journal of Heuristics, 22(4), 431-458.

[3] Hutter, F., Hoos, H., & Leyton-Brown, K. (2015). Identifying key algorithm parameters and instance features using forward selection. Lecture Notes in Computer Science, 7997, 364-381.

[4] Hutter, F., Hoos, H., & Leyton-Brown, K. (2015). An efficient approach for assessing hyperparameter importance. International Conference on Machine Learning, 754-762.

[5] Rardin, R. L., & Uzsoy, R. (2001). Experimental evaluation of heuristic optimization algorithms: A tutorial. Journal of Heuristics, 7(3), 261-304.

Multi-tool hole drilling with sequence dependent drilling times

Reginald Dewil
KU Leuven, Department of Mechanical Engineering
e-mail: reginald.dewil@kuleuven.be

İlker Küçükoğlu
Uludag University
e-mail: ikucukoglu@uludag.edu.tr

Corrinne Luteyn
KU Leuven, Department of Mechanical Engineering
e-mail: corrinne.luteyn@kuleuven.be

Dirk Cattrysse
KU Leuven, Department of Mechanical Engineering
e-mail: dirk.cattrysse@kuleuven.be

Hole drilling is one of the fundamental machining processes used in the manufacturing of durable goods. Drilling a set of holes on a simple two dimensional work piece using a single tool can be modeled as a Traveling Salesman Problem (TSP). However, hole drilling typically involves work pieces that require holes to be drilled of different diameters. In many cases, a single tool is assigned beforehand to every hole. This so-called multi-tool hole problem also reduces to the TSP if only a single tool is used to complete a hole. Many effective algorithms for TSPs exist and it is surprising to the authors that still many papers are being published on this basic hole drilling application. However, as pointed out by Dewil et al. [1], there remain quite some research questions on multi-tool hole drilling with precedence constraints and on multi-tool hole drilling with sequence dependent drilling times. Multi-tool hole drilling with precedence constraints can be modeled as the precedence constrained TSP and fast optimization approaches suited for the problem sizes typically encountered in hole drilling have recently been proposed [2].

However, multi-tool hole drilling application with sequence dependent drilling times as first presented by Kolahan and Liang [3] is much more complex. Only a limited number of papers tackled this problem and do not exploit the problem structure at all [4][5] and no linear MIP formulation has been proposed.

We present two exact linear MIP formulations based on modeling the problem as a Precedence Constrained Generalized Traveling Salesman Problem. In addition, we identify several quick-win improvements in the original tabu search algorithm presented by Kolahan & Liang. Firstly, two preprocessing rules can greatly reduce the problem size in some instances. Secondly, Kolahan & Liang apparently used a very computationally expensive implementation of a tabu list. We implement a

more efficient approach based on a hash function. Thirdly, for a problem with n holes and m tools, the original approach required an O(n) operation to evaluate the objective function value of a single neighbor. By using a slightly different solution representation, this can be reduced to an O(m) evaluation.

References

[1] Dewil, R., Vansteenwegen, P., Cattrysse, D., A review of cutting path algorithms for laser cutters, The International Journal of Advanced Manufacturing Technology, 87(5-8):1865-1884 (2016)

[2] Montemanni R., Smith D.H., Gambardella L.M., A heuristic manipulation technique for the sequential ordering problem, Computers & Operations Research, 35(12):3931-3944 (2008)

[3] Kolahan, F., Liang, M., Optimization of hole-making operations: a tabu-search approach, International Journal of Machine Tools & Manufacture, 40:1735-1753 (2000)

[4] Dalavi A.M., Pawar P.J., Singh T.P., Tool path planning of hole-making operations in ejector plate of injection mould using modified shuffled frog leaping algorithm, Journal of Computational Design and Engineering, 3:266-273 (2016)

[5] Dalavi A.M., Optimal sequence of hole-making operations using particle swarm optimization and shuffled frog leaping algorithm, Engineering Review, 36(2):187-196 (2016)

Giving Priority Access to Freight Vehicles at Signalized Intersections

Joris Walraevens
Ghent University - UGent, Department of Telecommunications and Information Processing
e-mail: Joris.Walraevens@UGent.be

Tom Maertens
Ghent University - UGent, Department of Telecommunications and Information Processing
e-mail: Tom.Maertens@UGent.be

Sabine Wittevrongel
Ghent University - UGent, Department of Telecommunications and Information Processing
e-mail: Sabine.Wittevrongel@UGent.be

Deceleration and acceleration of vehicles have a large impact on pollution and on waiting times. This is even more outspoken for freight vehicles, as they have much longer acceleration times and higher pollution levels than regular vehicles. Therefore, it makes sense to avoid, as much as possible, that freight vehicles have to stop at intersections. Note that this can also be beneficial for regular vehicles, as they have a lesser chance of getting stuck behind a freight vehicle at an intersection.
We concentrate on signalized intersections in this abstract. The aim is to give freight vehicles a high probability of green-light passage. This can be done by extending standard actuated control signals. First, freight vehicles should be detected. This can be done through sensing traffic and distinguishing freight vehicles from regular vehicles (by means of height, weight, ...). A future and promising alternative is active identification of freight vehicles by means of V2I (Vehicle-to-Infrastructure) communication. Second, an algorithm is required to decide whether the freight vehicle can be given a green light or not. This algorithm should be based on efficiency of passage, fairness for all intersecting traffic streams and vehicle types, safety, ...
We analyze the effect of increasing the green-light probability for freight vehicles on the mean waiting times of freight and regular vehicles. We concentrate on a fairly simple algorithm and intersection configuration. We assume that within a complete green/red cycle of the intersection, one of the green periods can be extended for a fixed period if freight vehicles are approaching. Obviously, the concurrent red periods are extended as well. We model separately each stream passing the intersection, where we assume that all vehicles that use the same lane make up one stream. So for a regular four-way intersection with separate turn lanes, we have 12 streams (4 origins and 3 destinations per origin). We further assume that a stream that is given a green light can pass the intersection without interference of other streams.
To model the streams, we group them in two categories: streams that benefit (category 1) and streams that suffer (category 2) from the extended green time for freight vehicles. As a first model, we assume independent Poisson arrivals for regular vehi-

cles and freight vehicles in each stream. A generic stream of category 1 is modeled as follows: freight vehicles arrive at rate λ_f, regular vehicles at rate λ_r. Without the extension of the green period, the stream sees cycles of free flow (green light; duration t_g seconds) and blocked access (red or yellow light; duration t_r seconds). The green light is extended for a period of t_{eg} seconds if at least one freight vehicle of the streams of category 1 arrives during this period. A generic stream of category 2 has similar parameters; the main difference is that the red period of these streams is extended when the green period of the streams of category 1 is extended.
We analyze the mean waiting times of regular vehicles and freight vehicles in generic streams of both categories (i.e., 4 different analyses and end formulas). Since we expect that the algorithm will have the biggest effect in a light-traffic scenario, we assume that no waiting time is added by the intersection when a vehicle arrives during a green period. When arriving in a red period, the waiting time of a vehicle consists of the time until the traffic sign turns green and the time until the vehicle passes the traffic light after it turns green. The latter depends on the number of vehicles in front of the vehicle when the traffic light turns green and the speed with which the vehicle can leave the intersection. Since we assume slower discharge speeds for freight vehicles than for regular vehicles, the latter, in turn, depends on whether the vehicle itself is a freight vehicle and/or whether at least one freight vehicle is in front of the vehicle.
The analyses exist of conditioning and calculating conditioned mean waiting times, but is in essence rather straightforward. We are in particular interested in the difference between mean waiting times in this scenario with extended green periods and mean waiting times when traffic control is static. Although the latter could be analyzed more accurately (for instance by using results from polling models), we propose to calculate the latter simply by substituting the green period extension t_{eg} by 0 in the expressions of the former scenario. The advantage of this approach is that both results suffer from the same approximation error.
The obtained formulas are explicit in the parameters of the model and can therefore easily be used to assess the impact of the extension length t_{eg} (and other parameters) on the mean waiting times of different streams. Furthermore, they can also be used in an optimization function to find the optimal t_{eg}.
In future, we will concentrate on extending the analysis to also deal with overflow during a red/green cycle. Relaxing the independent Poisson arrival processes to dependent platooning arrival processes is another interesting research direction.

Creating a dynamic impact zone for conflict prevention in real-time railway traffic management

Sofie Van Thielen
KU Leuven, Leuven Mobility Research Centre
e-mail: sofie.vanthielen@kuleuven.be

Pieter Vansteenwegen
KU Leuven, Leuven Mobility Research Centre
e-mail: pieter.vansteenwegen@kuleuven.be

One way to deal with the growing need for reliable public transport, is to improve the accuracy of the trains in real-time. The accuracy starts with a robust timetable, capable of accounting for possible delays. However, in real-time, unexpected events, e.g., mechanical failures, can lead to primary and secondary delays. Once trains start deviating from their scheduled times, conflicts can occur. A conflict takes place when two trains want to reserve the same part of the infrastructure at the same time. These conflicts need to be resolved quickly in a way that the entire network is as least as possible disturbed. Therefore, the impact on the network should be taken into account when trying to prevent conflicts. In order to prevent conflicts, the Belgian railway infrastructure manager Infrabel has recently implemented a new Traffic Management System (TMS) that is capable of predicting train movements and detecting conflicts in real-time. However, a good conflict prevention module is not present for practical use yet ([1]; [2]). This paper introduces an approach to deliver good, feasible solutions for preventing conflicts in real-time. This approach tries to complete a Decision Support System (DSS) supporting dispatchers in making good decisions.

Every time a conflict is detected by TMS, it is immediately sent to the Conflict Prevention Strategy (CPS). If the conflict takes place in a station area, a solution based on rerouting is looked for first. The optimization module is based on a flexible job shop and is limited in calculation time in Cplex to 30 seconds. If the rerouting solution does not solve the conflict or if the conflict does not take place in a station area, a solution based on retiming/reordering one of the trains is given. Choosing which train to delay, implies considering what the impact on the rest of the network will be in the near future. On the one hand, all trains that could be impacted by the solution of the current conflict need to be taken into account when deciding on the current conflict. On the other hand, the computation time of the CPS needs to remain as low as possible for its usage in practice. Therefore, it is important to only consider the most relevant trains for the current conflict.

Offline calculations are carried out beforehand to examine which conflicts are *most likely* to occur in practice. These *most likely conflicts* are considered in our Dynamic Impact Zone (DIZ) heuristic. In this manner, a *dynamic impact zone* can be created by considering conflicts where one of the trains is in the current conflict, i.e. a

first order conflict, and *most likely conflicts* with at least one of the trains in a *first order conflict*. Whenever a conflict needs to be prevented by delaying one of the trains, the progress of further movements is examined in both cases (delaying either one of the trains). During this progress, only trains relevant to the current conflict are considered. Relevant trains are trains in the *dynamic impact zone*. For more explanation on the methods used, we refer to [3].
The DIZ heuristic is tested on a large network: Brugge-Gent-Denderleeuw in Belgium. The simulation considers trains between 7 and 8 a.m., and includes in total 152 trains. A delay scenario assumes 60 % of all trains enter the study area with a delay randomly taken from an exponential distribution with an average of 3 minutes and maximum 15 minutes. The DIZ heuristic is compared to different dispatching strategies. The first strategy is First Come, First Served, resembling an unexperienced dispatcher. The second strategy considers all first-order conflicts. Table 2 shows results for 20 runs. Our DIZ heuristic clearly outperforms both FCFS and the first-order strategy, and at the same time has a rather low computation time. Whenever a conflict is sent to the CPS, it renders a feasible solution in 1 second on average. In 95 % of the cases the computation time of the CPS remains under the 2 seconds, which makes it very suitable for usage in practice.

Strategy	Total secondary delay	Average computation time
FCFS	774 min	0.02 s
First-order + rerouting	436 min (- 44 %)	0.9 s
DIZ + rerouting	252 min (- 67 %)	1 s

Table 2: Comparison of conflict prevention techniques based on total secondary delay and computation time.

References

[1] Borndörfer, R., Klug, T., Lamorgese, L., Mannino, C., Reuther, M. and Schlechte, T., *Recent succes stories on integrated optimization of railway systems.* Transportation Research Part C, 74 (2017), pp. 196–211.

[2] Corman, F., and Quaglietta, E., *Closing the loop in real-time railway control: Framework design and impacts on operations*, Transportation Research Part C, 54 (2015), pp. 15–39.

[3] Van Thielen, S., Corman, F. and Vansteenwegen, P., *Considering a dynamic impact zone for real-time railway traffic management.* Transportation Research Part B (2018) review submitted.

The inventory/capacity trade-off with respect to the quality processes in a Guaranteed Service Vaccine Supply Chain

S. Lemmens
KU Leuven, Research Center for Operations Management
e-mail: stef.lemmens@kuleuven.be

C. J. Decouttere
KU Leuven, Research Center for Operations Management

N. J. Vandaele
KU Leuven, Research Center for Operations Management

M. Bernuzzi
KU Leuven, Research Center for Operations Management

A. Reichman
GlaxoSmithKline Vaccines

Vaccine manufacturing supply chains are characterized by long lead times which may exceed 300 days. Such long lead times make it challenging to guarantee on-time deliveries. The long lead times are determined by the high-volume manufacturing processes (formulation, filling and packaging), but they are especially the result of the stringent quality control and quality assurance procedures. For vaccine manufacturing, the touch time is only 5 to 10% of the total lead time which indicates that it is worthwhile to study the impact of capacity on the lead times of the quality processes into more detail. For such supply chain stages, these long lead times are partly due to limited capacity in terms of laboratory equipment as well as skilled people.
The Quality Control (QC) and Quality Assurance (QA) department guarantees persistence of quality during these processes by various tests. To continue the manufacturing processes (formulation, filling and packaging), some critical tests (e.g. sterility tests) need to be performed immediately after the formulation and filling stages to identify whether the vaccines are allowed to proceed to the next stage. When the results of these critical tests are positive, the vaccines are put in a restricted status and may proceed to the next manufacturing stage. The other quality requirements, the QC and QA, are performed in parallel with the subsequent manufacturing stages as the lead times of these quality stages can be long. Each QC/QA stage (for formulation, filling and packaging) can be decomposed into four independent processes:

1. Actual quality testing of the product
2. Investigation of deviations in the actual testing (e.g. due to the calibration of

laboratory instruments)

3. Document review of the production activities

4. Request for Process Change: post-approval regulatory process to change the manufacturing process according to customers' requirements. Such changes may include changes to the vaccine composition, quality control, equipment, facilities or product labelling information after a vaccine has been approved

Queuing networks and the Guaranteed Service Approach (GSA) are two well established modeling methodologies. However, as both models include nonlinear effects, [2] mention that the direct impact of the capacity on the GSA's resulting stock levels is particularly cumbersome to derive with a resulting closed-form expression. To integrate both models, [2] demonstrate the following three steps to embed capacity into the Guaranteed Service Approach with Variable lead times (GSA-VAR):

1. Calculate the average and variance of the lead time using batch queuing networks

2. Characterize the entire lead time distribution

3. Extend the lead time information obtained in previous steps into the GSA-VAR ([1])

In this work, step 1 will be replaced with an approximate queuing methodology to calculate the average and variance of the lead time of two capacity-dependent quality process nodes: the investigation of deviations for formulation and

References

[1] Humair, S., Ruark, J., Tomlin, B., and Willems, S. P. (2013). Incorporating stochastic lead times into the guaranteed service model of safety stock optimization. Interfaces, 43(5):421-434.

[2] Lemmens, S., Decouttere, C., Vandaele, N., Bernuzzi, M., and Reichman, A. (2016). Performance measurement of a rotavirus vaccine supply chain design by the integration of production capacity into the guaranteed service approach. KULeuven FEB Research Report KBI_1621.

Optimization tool for the drug manufacturing in the pharmaceutical industry

Charlotte Tannier
Head of Solutions Implementation, **N-SIDE**
e-mail: cta@n-side.com

The pharmaceutical industry lacks an efficient system to manage the forecast at the drug substance and drug product level for their R&D activities. This requires to smartly pool the demand of multiple clinical trials and technical demands while managing the complex aspects of drug expiries. Based on this statement, N-SIDE jointly developed with a Top 10 Pharmaceutical company a new optimization software, called CT-PRO. CT-PRO is a cutting-edge tool based on optimization algorithms which fulfills this need by addressing both risk management and cost minimization of all steps from drug substance manufacturing to the finished drug dispensed to patients.

Last year, the 10 largest pharmaceutical companies invested more than 70 billions USD on their R&D activities, running clinical trials to assess the safety and efficiency of their new drug candidates. About 10% of these costs are related to their end-to-end supply chain, from manufacturing of the drug substance to the final dispensing of the finished product (i.e. drug kit) to the patients at the different hospitals participating on these clinical trials.

The drug waste related to these activities is very high (i.e. usually larger than the amount of drug that will be dispensed to the patients), due to several complexity factors (i.e. uncertainty on the demand at each stage in the supply chain) associated to the R&D process.

Indeed, a clinical trial will be recruiting patients during a few months at different places all over the world, creating a sudden peak in the drug demand, that will only last for a few months before the end of the trial. The uncertainty in the drug patient demand is then huge, both in term of timing and location (e.g. will we find more patients in Russia or in Brazil?). Still, the golden rule is to ensure 100% of the demand so that no patient will suffer from treatment discontinuation and the final clinical trial results may be investigated to -if successful- lead to the drug approval commercialization. Last complexity factor, related to the pharmaceutical industry and the drug management; the very long lead times. Indeed, it will take weeks or even months to move the drug from one country/continent to another, as it may take months to manufacture drug and/or transform it (e.g. from drug substance to drug product, drug kit packaging or labelling, etc.). This may be due to strong regulations in place and/or to quality management aspects.

Having highly uncertain demand that should be perfectly fulfilled with huge lead time (and therefore the obligation to anticipate many key decisions) in the supply chain is therefore a great situation for optimization. The optimization provided by N-SIDE is therefore provided by two interconnected systems combining some different mathematical techniques.

The first system, called CT-FAST, is estimating drug demand for different clinical trials using stochastic modelling (using standard Poisson distribution to represent patient recruitment and other uncertain events), Monte-Carlo simulations and machine learning techniques to leverage real-time data. CT-FAST is also optimizing the distribution strategy as well as the finished drug supply plan.

The second system, called CT-PRO, is optimizing the end-to-end supply chain for all clinical trials sharing the same ingredients (i.e. called clinical program), from drug substance manufacturing to the final dispensing of the finished product to the patients. To do so, CT-PRO software is composed of mixed integer programming (MIP) algorithms in the context of complex supply chain management points, including expiring products and specific flow constraints.

New optimization levers enabled by CT-PRO, and by its integration with CT-FAST, result in significant savings - from 5%to 30% of manufacturing cost - while keeping the very high service level for patients. The scope of CT-PRO covers all production and transformation stages, drug storage & shipments, expiry management aspects (with time-dependent shelf-life), different kind of demands, internal & external suppliers, etc.

CT-PRO's outputs will be composed of an optimized global production planning (including the manufacturing resource selection, drug substance lot size, timing, etc.), the drug/lot allocation to the different sources of demand (e.g. to the multiple clinical trials sharing the same drug substance).

The presentation will focus on the unanswered questions and challenges within the R&D drug manufacturing and how various mathematical techniques may be mixed together to bring risk-based optimization results - and therefore huge waste and cost reduction to the pharmaceutical industry!

Staff Scheduling at a Medical Emergency Service: a case study at Instituto Nacional de Emergência Médica

H. Vermuyten
KU Leuven Campus Brussels,
Department of Information Management, Modeling and Simulation
e-mail: hendrik.vermuyten@kuleuven.be

J. Namorado Rosa
University of Lisbon, Instituto Superior Técnico,
Centre for Management Studies
joana.namorado.rosa@gmail.com

I. Marques
University of Lisbon, Instituto Superior Técnico,
Centre for Management Studies
ines.marques.p@tecnico.ulisboa.pt

J. Beliën
KU Leuven Campus Brussels,
Department of Information Management, Modeling and Simulation
jeroen.belien@kuleuven.be

A. Barbosa-Póvoa
University of Lisbon, Instituto Superior Técnico,
Centre for Management Studies
apovoa@tecnico.ulisboa.pt

This presentation addresses a real-life personnel scheduling problem at a medical emergency service. In a first step, the problem is formulated as an integer program. However, the problem turns out to be too complex for a commercial IP solver for realistic problem instances. For this reason, two heuristics are developed, namely a diving heuristic and a variable neighbourhood decomposition search (VNDS) heuristic. The diving heuristic reformulates the problem by decomposing on the staff members. The LP relaxation is then solved using column generation. After the column generation phase has finished, integer solutions are found by heuristically branching on all variables with a value above a certain threshold until the schedule has been fixed for each staff member. Additionally, two different column generation schemes are compared. The VNDS heuristic consists of a local search phase and a shake phase. The local search phase uses the principles of fix-and-optimise in combination with different neighbourhoods to iteratively improve the current solution. After

a fixed number of iterations without improvement, the search moves to the shake phase in which the solution is randomly changed to allow the heuristic to escape local optima and to diversify the search.
The performance of both heuristics is tested on a real-life case study at Instituto Nacional de Emergência Médica (INEM) as well as nineteen additional problem instances of varying dimensions. Four of these instances are derived from the INEM dataset by changing one of the problem dimensions, while an instance generator was built to construct the fifteen other instances. Results show that the VNDS heuristic clearly outperforms both diving heuristic implementations and a state-of-the-art commercial IP solver. In only hour of computation time, it finds good solutions with an average optimality gap of only 4.76 percent over all instances. The solutions proposed by the VNDS are then compared with the actual schedule implemented by INEM. To facilitate this comparison, the heuristic is implemented in a scheduling tool with an attractive graphical user interface. For six choices of objective function weights, which put different emphases on the relative importance of the various soft constraints, schedules are constructed by the VNDS. Each of the six schedules is compared with the actual schedule implemented by INEM and evaluated on eight key performance indicators (KPIs). All proposed schedules outperform the real schedule. Finally, two what-if scenarios are analysed to illustrate the impact of managerial decisions on the quality of the schedules and the resulting employee satisfaction.

Home chemotherapy: optimizing the production and administration of drugs

V. François
QuantOM, HEC Liège - Management School of the University of Liège
e-mail: veronique.francois@uliege.be

Y. Arda
QuantOM, HEC Liège - Management School of the University of Liège

D. Cattaruzza
Centrale Lille, INRIA, INOCS

M. Ogier
Centrale Lille, INRIA, INOCS

We introduce an integrated production scheduling and vehicle routing problem originally motivated by the rising trend of home chemotherapy. Oncology departments of Belgian hospitals have a limited capacity to administer chemotherapy treatments on site. Also, patients in a stable condition often prefer to get their treatment at home to minimize its impact on their personal and professional life. The optimization of underlying logistical processes is complex. At the operational level, chemotherapy drugs must be produced in the hospital pharmacy and then administered to patients at home by qualified nurses before the drugs expire. The lifetime of the drug, i.e., the maximum period of time from the production start until the administration completion, is sometimes very short, depending on the treatment type. The problem we consider calls for the determination, for each patient, of the drug production start time and administration start time, as well as the assignment of these two tasks to a pharmacist and a nurse respectively.

Drugs may be produced by any of the available pharmacists which are considered to be homogeneous. The working duration of a pharmacist is the time elapsed between the production start time of the first drug and the completion time of the last drug. The allowed working duration of each pharmacist is limited but the start time of each working shift is a decision variable implicitly defined by the production schedule.

Drugs are administered by nurses that have a limited working duration. The administration of each drug must be scheduled within a time window that depends on the concerned patient and before the drug expires. This expiry aspect is crucial since it calls for the integration of the scheduling and the routing aspects of the problem. Multiple trips are allowed, i.e., each nurse may come back at the hospital during its journey to take drugs recently produced.

The objective of the problem is to minimize the total working time of pharmacists and nurses.

We address this problem heuristically using destroy-and-repair mechanisms and allowing infeasible solutions to be visited. The main challenge is that a small change in the production schedule or the administration schedule affects other decision variables in a cyclic fashion. Thus, the change in the objective function associated to

a given move is difficult to evaluate efficiently. To circumvent this issue, we propose to iterate between the production and the routing subproblems, by imposing constraints on one of the subproblems depending on the incumbent solution of the other. Moves on the solution of one subproblem are thus evaluated seemingly independently from the other one. Then, occasionally, an exact procedure is called to determine the optimum schedule of the integrated incumbent solution given the task sequence of each pharmacist and of each nurse.

An analysis of short term objectives for dynamic patient admission scheduling

Yi-Hang Zhu
KU Leuven, Department of Computer Science, CODeS & imec - Belgium
e-mail: yihang.zhu@cs.kuleuven.be

Túlio A. M. Toffolo
Department of Computing, Federal University of Ouro Preto, Ouro Preto, Brazil

Wim Vancroonenburg
KU Leuven, Department of Computer Science, CODeS & imec - Belgium
Research Foundation Flanders - FWO Vlaanderen

Greet Vanden Berghe
KU Leuven, Department of Computer Science, CODeS & imec - Belgium

Every day, new patients registering in hospitals for treatment may need surgery and rooms for recovery. The Dynamic Patient Admission Scheduling Problem (DPAS), introduced by [3], captures the dynamic nature of this patient admission procedure and concerns providing an admission day, room and surgery day (if necessary) to each patient while several hard and soft constraints are respected.

Hard constraints must be respected. They include a maximum admission delay, room capacity, maximum overtime in operating rooms (ORs), and room equipment and specialisms required by patients. The soft constraints may be violated with penalties considered in the objective function and include room properties required by patients, room gender policies, sudden room transfers, overtime in ORs and room overcrowd risks.

The DPAS is addressed by a periodic re-optimization method [1] with the length of each period as one day. Each day's problem, denoted as a short term problem, contains all the latest patient information of this day. After all short term problems are solved, a long term solution is obtained which contains a schedule for each patient.

There is, however, an issue associated with this method. When simply using the long term objective function for solving short term problems, patient requests may be delayed. Indeed, more preferable resources may be available in the future whereas no information is available on future requests. This may subsequently result in a short term problem of a later period containing a large number of delayed requests from early periods and also the requests from that day. The available resource capacity may be insufficient to accommodate all the requests and, eventually, low quality or even infeasible solutions are generated. Therefore a well designed objective for each day's short term problem is important in terms of obtaining a good quality long term solution.

The DPAS and its variants have been investigated in several papers [4, 2, 6, 7, 5, 3]. These studies focused on developing better algorithms for solving short term problems. Currently, to the best of our knowledge, no studies have focused on designing a better short term objective while targeting long term solution quality for the DPAS.
This work investigates *how different short term strategies impact long term solutions*. Three integer programming formulations are developed, two of which are designed by incorporating different short term strategies. Experiments are performed to investigate the performance of the proposed approaches on solving the instances from [3]. The obtained solutions are subsequently evaluated with newly calculated lower bounds and suggest that the proposed approach provides better solutions than the best-known for the problem.

References

[1] Angelelli, E., Bianchessi, N., Mansini, R. and Speranza, M.G., 2009. Short term strategies for a dynamic multi-period routing problem. Transportation Research Part C: Emerging Technologies, 17(2), pp.106-119.

[2] Ceschia, S. and Schaerf, A., 2012. Modeling and solving the dynamic patient admission scheduling problem under uncertainty. Artificial intelligence in medicine, 56(3), pp.199-205.

[3] Ceschia, S. and Schaerf, A., 2016. Dynamic patient admission scheduling with operating room constraints, flexible horizons, and patient delays. Journal of Scheduling, 19(4), pp.377-389.

[4] Demeester, P., Souffriau, W., De Causmaecker, P. and Berghe, G.V., 2010. A hybrid tabu search algorithm for automatically assigning patients to beds. Artificial Intelligence in Medicine, 48(1), pp.61-70.

[5] Lusby, R.M., Schwierz, M., Range, T.M. and Larsen, J., 2016. An adaptive large neighborhood search procedure applied to the dynamic patient admission scheduling problem. Artificial intelligence in medicine, 74, pp.21-31.

[6] Range, T.M., Lusby, R.M. and Larsen, J., 2014. A column generation approach for solving the patient admission scheduling problem. European Journal of Operational Research, 235(1), pp.252-264.

[7] Vancroonenburg, W., De Causmaecker, P. and Berghe, G.V., 2016. A study of decision support models for online patient-to-room assignment planning. Annals of Operations Research, 239(1), pp.253-271.

Robust Kidney Exchange Programs

B. Smeulders
QuantOM, HEC Liège - Management School of the University of Liège
e-mail: bart.smeulders@uliege.be

Y. Crama
QuantOM, HEC Liège - Management School of the University of Liège

F.C.R. Spieksma
TU Eindhoven, Department of Mathematics and Computer Science

Kidney transplants are the preferred treatment for end-stage renal disease. Many patients have a donor, often a friend or family member, who has volunteered to donate one of their kidneys for a transplant. However, the donor must be compatible with the patient for a transplant to be possible. Kidney exchanges can solve this problem. Patients who are incompatible with their own donor are matched to other incompatible donor-patient pairs. The donor gives his kidney to a different patient, in return the patient receives a kidney from a different donor. Such cycles consist of 2 or more pairs. Kidney exchange programs may also include non-directed donors. These altruists are willing to donate to any compatible patient in the pool. This may be the start a chain of donations, as the donor associated with that patient in turn donates to another patient with whom he is compatible. Typically, both cycles and chains are limited in length. Matching donor-patient pairs and non-directed donors to one another so as to maximize the number of transplants (or the value of weighted transplants) is an NP-Complete optimization problem.

One complication in kidney exchanges is that compatibilities are rarely certain. Preliminary tests may indicate that a given donor can donate to a given patient, but more thorough testing may show that this is impossible. Since more thorough tests are more expensive and time-consuming, matchings are usually computed based on preliminary tests. Afterwards, tests are performed to make sure the identified donations are possible. In this stage, some donations may turn out to be impossible, and the associated cycles or chains are broken.

We are interested in the kidney exchange problem, taking into account such post-testing failures. Dickerson et al. [1] compute the probability that a given cycle is successful, and using this information maximizes the expected number of transplants. Other approaches include the option for recourse, a limited re-matching if transplants included in the initial solution turn out to be impossible after further testing. Pedroso [2] studied an internal-recourse scheme. In this scheme, if a donation in a cycle can not be performed, the patient-donor pairs within the cycle can be rearranged to form a new (shorter) cycle. Klimentova et al. [3] proposes a subset-recourse scheme. In this approach, subsets of patient-donor pairs are chosen. Within these subsets, all possible donations are tested. The subsets to be tested are

chosen so as to maximize the expected number of transplants.

In essence, the recourse schemes have two stages. In a first stage, a number of potential transplants are selected to test. The potential transplants for which these tests are successful are then used in a second stage, to maximize the number of transplants. Pedroso and Klimentova et al. limit their choice of potential transplants to test, by linking the tests to inclusion in a longer cycle or in a subset.

In this talk, we look at a generalization of these recourse schemes, where the only constraint on the testing is an upper bound on the number of tests that can be performed. We discuss what the potential gains of this generalization could be, in terms of extra transplants, but also the drawbacks regarding organization and computational difficulties.

References

[1] J. P. Dickerson, D. F. Manlove, B. Plaut, T. Sandholm, and J. Trimble, *Position-indexed formulations for kidney exchange*, in Proceedings of the 2016 ACM Conference on Economics and Computation, 2016, pp. 25-42.

[2] J. P. Pedroso, *Maximizing expectation on vertex-disjoint cycle packing*, in International Conference on Computational Science and Its Applications, 2014, pp. 32-46.

[3] X. Klimentova, J. P. Pedroso, and A. Viana, *Maximising expectation of the number of transplants in kidney exchange programmes*, Computers & Operations Research, vol. 73, pp. 1-11, 2016.

Considering complex routes in the Express Shipment Service Network Design problem

José Miguel Quesada
Université catholique de Louvain, Louvain School of Management
e-mail: jose.quesada@uclouvain.be

Jean-Charles Lange

Jean-Sébastien Tancrez
Université catholique de Louvain, Louvain School of Management
e-mail: js.tancrez@uclouvain.be

In the transportation industry, the express integrators are the providers that offer the fastest and more reliable door-to-door delivery services worldwide. With the increasing demand of customers for faster responses, the service proposition of the express integrators plays a key role in the logistics industry to make supply and demand meet. The premium service of the express carriers is the overnight delivery of packages within regions as large as the US or Europe. To achieve such services, the express integrators are highly dependent on improving the efficiency of their network operations. Defining a schedule of flights that enables the delivery of this service at the best cost is known as the Express Shipment Service Network Design (ESSND) problem.

Due to the difficulty to solve real size instances of the ESSND problem, the works from the literature follow a two-stage process to model it. In the first stage, a set of feasible routes that the express integrator is able to perform is computed by enumeration. This process is called *route generation*. Then, the generated routes become the input of a mixed integer programming (MIP) model, which is solved to design a transportation network at minimal cost. The advantage of this two-stage process is that the route generation computes the paths, schedules and costs of the routes, which no longer need to be modeled explicitly in the second stage. Despite this two-stage process, authors have paid attention to generate enough route types to obtain useful solutions, but not too many so as to keep models tractable. The common route types in these models are the one-leg, multi-leg and ferry routes. We call them *standard routes*. A few works include some non-standard route types, the transload, direct and inter-hub routes, but usually by predefining their loads and/or by fixing them as part of the solution, and without studying their specific contribution. Therefore, the benefit of the non-standard routes when solving the ESSND problem is an open question of practical importance. We refer to non-standard routes as *complex routes*.

In this research, our goal is to assess the contribution of five types of complex routes that are not or are rarely explored in the literature: two-hub routes that connect gateways with two hubs with a single aircraft; transload routes that transfer packages

between aircrafts; direct routes that move packages from their origins to their destinations without visiting hubs; inter-hub routes that move packages between hubs; and early and late routes with relaxed release and due times. We assess the contribution of these routes as follows. First, we propose an MIP model in which we do not enforce them in the solution nor predefine their loads. To improve the tractability of the model, we enhance it with three families of valid inequalities. Second, we solve the model for an extensive set of experiments, first on limited size instances (21 gateways and 6 equipment types), and then on realistic instances (77 gateways and 7 equipment types). The inter-hub, direct and early and late routes have the best performance, with average savings of 3.42%, 1.05% and 3.89% respectively for large instances .

Acknowledgments

This project was funded by the Brussels-Capital Region and Innoviris.

Assessing Collaboration in Supply Chains

Thomas Hacardiaux, Jean-Sébastien Tancrez
Université catholique de Louvain, Louvain School of Management, CORE
e-mail: thomas.hacardiaux@uclouvain.be, js.tancrez@uclouvain.be

Nowadays, companies have to find the right balance between a cost-effective supply chain and a competitive service to their customers. Delivery frequencies are increased to reduce inventory costs and improve customer service, but lead to incomplete loading of trucks, while the oil price and road taxes are persistently increasing [6, 9]. In order to improve cost-effectiveness as well as customer service, some companies turn towards horizontal cooperation; an active cooperation between two or more firms that operate on the same level of the supply chain and perform a comparable logistics function on the landside [1].

The literature on horizontal cooperation discusses several real-life cases and simulation studies. Authors generally acknowledge that horizontal collaborations generate savings. Recent researches have however shown that these savings may have various magnitudes from 5% to 30% [4, 3]. To explain these mixed results, partners' and markets' particularities must be taken into account [1]. Transportation costs [5], inventory costs [2], distances [7], CO_2 emissions [8] are some of the parameters that may be integrated. The horizontal cooperation brings many advantages but also some risks and challenges. Indeed, the failure rate for horizontal cooperation projects is rather high, ranging between 50 and 70% [10].

This work analyzes the benefits for companies to use a joint supply chain network (facilities and vehicles), and investigates the markets' and partners' characteristics that influence these benefits to understand when horizontal cooperation is particularly profitable. For this, we propose a location-inventory model, formulated as a conic quadratic mixed integer program. The model integrates the main logistical costs such as facility opening, transportation, cycle inventory, ordering and safety stock costs. The transportation cost is accounted per vehicle and shipment size decisions are included so that an important benefit of collaboration, i.e. the improvement of the vehicle loading rate, is accurately assessed.

We perform about 30,000 experiments varying the parameters values (i.e. vehicle capacity, facility opening cost, inventory holding cost, ordering cost and demand variability). From these experiments, we infer valuable managerial insights to help companies assessing the potential benefits they can get when cooperating, depending on their characteristics. The synergy value ranges from 15% to 30% with an average of 22.4%. The benefits come from, in order of importance, the cycle inventory at retailers (increased delivery frequency), the transportation (improved loading rate and reduced distances), the opening of joint facilities and the safety stock costs at retailers (reduced lead times). On the opposite, the order and the safety stock costs at DCs increase. Our results also pinpoint the characteristics that should motivate companies to collaborate, and for which they should look in their partners.

In particular, the horizontal cooperation is more profitable for companies with high facility opening costs, large vehicle capacity, low order costs, carrying small products (low unit holding cost) on a market with a low demand variability. Moreover, we observe that horizontal cooperation also improves the service, increasing the service level from 97.5% to 98.2% (or further reducing the costs for a fixed service level) and increasing the delivery frequency for a specific product.

Finally, we run additional experiments which show that the synergy value keeps increasing with the number of partners but that the marginal benefit is smaller with each new entrant. Companies have to balance the additional gains from adding a partner and the higher complexity to manage the partnership. Furthermore, a cooperation without location decisions opens several DCs in close locations and deteriorates their spread. We investigate when it is necessary to completely relocate DCs or when closing some existing DCs, requiring a lower commitment in term of investment and risk, is sufficient. Lastly, we analyze the impact of the demand regional distribution on the synergy value, and on the costs of the cooperation case.

References

[1] Cruijssen, F., 2006. Horizontal cooperation in transport and logistics. Ph.D. Thesis. Tilburg University, Center for Economic Research.

[2] Ergun, O., Kuyzu, G., Savelsbergh, M., 2007. Reducing truckload transportation costs through collaboration. Transportation Science, 41(2) : 206-221.

[3] Frisk, M., Göthe-Lundgren, M., Jörnsten, K., Rönnqvist, M., 2006. Cost allocation in forestry transportation. Proceedings of the ILS Conference, Lyon.

[4] Hageback, C., Segerstedt, A., 2004. The need for co-distribution in rural areas - a study of Pajala in Sweden. In: International Journal of Production Economics, 89(2) : 153-163.

[5] Juan, A., Faulin, J., Pérez-Bernabeu, E., Jozefowiez, N., 2014. Horizontal cooperation in vehicle routing problems with backhauling and environmental criteria. Procedia - Social and Behavioral Sciences, 111 : 1133-1141.

[6] Lozano, S., Moreno, P., Adenso-Diaz, B., and Algaba, E., 2013. Cooperative game theory approach to allocating benefits of horizontal cooperation. European Journal of Operational Research, 229(2): 444-452.

[7] Mason, R., Lalwani, C., Boughton, R., 2007. Combining vertical and horizontal collaboration for transport optimization. Supply Chain Management: An International Journal, 12 (3) : 187-199.

[8] Pan, S., Ballot, E., Fontane, F., 2013. The reduction of greenhouse gas emissions from freight transport by pooling supply chains. In: International Journal of Production Economics, 143(1) : 86-94.

[9] Pomponi, F., Fratocchi, L., Rossi Tafuri, S., 2015. Trust development and horizontal collaboration in logistics: a theory based evolutionary framework. Supply Chain Management: An International Journal, 20(1): 83-97.

[10] Schmoltzi, C. and Wallenburg, C.M., 2011. Horizontal cooperations between logistics service providers: motives, structure, performance. International Journal of Physical Distribution and Logistics Management, 41(6): 552-575.

Planning of feeding station installment and battery sizing for an electric urban bus network

Virginie Lurkin
Ecole polytechnique fédérale de Lausanne (EPFL),
Transport and Mobility Laboratory (Transp-OR)
e-mail: virginie.lurkin@epfl.ch

with A. Zanarini, Y. Maknoon, S.S. Azadeh and M. Bierlaire.

Motivation

During the last few decades, environmental impact of the fossil fuel-based transportation infrastructure has led to renewed interest in electric transportation infrastructure, especially in the urban public transportation sector. The deployment of battery-powered electric bus systems within the public transportation sector plays an important role to increase energy efficiency and to abate emissions. Rising attention is given to bus systems using fast charging technology. The MyTOSA project which is jointly undertaken by the ABB company and the TRANSP-OR Lab at EPFL aims to provide industrial solutions to implement an electric urban bus network. An efficient feeding stations installation and an appropriate dimensioning of battery capacity are crucial to minimize the total cost of ownership for the citywide bus transportation network and to enable an energetically feasible bus operation.

Problem description

In the frame of the industrial project MyTOSA, catenary-free buses have been developed and equipped with fast charging technology in order to allow buses to charge at bus stops while in service. The idea is that the buses are equipped with an arm on top of their roof that can unfold in order to connect to a feeding station installed at a bus station (see Figure 1). This way the buses can quickly charge while passengers are entering or leaving the bus.

The aims of this project are to decide at which stations we should install a feeding station, which type of feeding station should be installed at these stations and with which battery we should equip the buses in order to minimize the total cost of ownership for implementing the bus network. These costs can be divided into two main categories:

1. Capital Expenditures (CAPEX): costs that occur once in the project lifetime. This includes acquisition and digging costs.

2. Operational Expenditures (OPEX): costs that occur yearly. This includes the maintenance and operating costs of buses and feeding stations.

Figure 1: A TOSA bus charging at a bus stop.

A set of constraints exists regarding the usage of the battery, the charging and dwell times, as well as the compatibilities between batteries and feeding stations and batteries and frequency of charging.

Implementation and preliminary results

A mixed-integer linear optimization model is developed to determine the optimal feeding stations installation for a bus network as well as the adequate battery capacity for each bus line of the network. The implementation has been carried out using the programming language *C++* on *VisualStudio15*$^{\hat{A}\acute{o}}$ and the optimization problem has been tackled with the 12.7 version of the *CPLEX*$^{\hat{A}\acute{o}}$ optimizer. The model has been tested on several instances and provides encouraging results.

A matheuristic for the problem of pre-positioning relief supplies

R. Turkeš
University of Antwerp Operations Research Group ANT/OR
e-mail: renata.turkes@uantwerpen.be

K. Sörensen
University of Antwerp Operations Research Group ANT/OR
e-mail: kenneth.sorensen@uantwerpen.be

D. Palhazi Cuervo
SINTEF Mathematics and Cybernetics Department
e-mail: daniel.palhazicuervo@sintef.no

Every year, natural and man-made disasters have devastating effects on millions of people around the world. An estimated 400 000 people lost their lives and more than 4 000 000 were affected in only two recent major natural disasters, the Haiti earthquake and the Indian Ocean earthquake and tsunami [1]. What is worse, both natural and man-made disasters are expected to increase another five-fold over the next 40 years due to environmental degradation, rapid urbanization and the spread of HIV/AIDS in the developing world [2].

No country is immune from the risk of disasters, but much human loss can be avoided by preparing to better deal with these emergencies. One mechanism to increase preparedness is advance procurement and pre-positioning of emergency supplies (such as water, food or medicine) at strategic locations. This allows to additionally speed up emergency assistance and save more lives by reaching areas that could be otherwise inaccessible. To develop a pre-positioning strategy is to determine:

(1) number, location and category of storage facilities to open, represented by binary variables $\mathbf{x} = [x_{iq}]$ that indicate whether a facility of category q is open at vertex i,

(2) amount $\mathbf{y} = [y_i^k]$ of commodity k to pre-position at a facility open at vertex i, and

(3) aid distribution strategy, represented by binary variables $\mathbf{z} = [z_{ij}^s]$ that indicate whether a facility open at vertex i serves the demands of vertex j in disaster scenario s,

under uncertainty about the demands, survival of pre-positioned supplies and transportation network availability.

Despite a growing scientific attention for the pre-positioning problem [3], there still does not exist a set of benchmark instances what hinders thorough hypotheses testing, sensitivity analyses, model validation or solution procedure evaluation. Researchers therefore have to invest a lot of time and effort to process raw historical data from several databases to generate a single case study that is nonetheless hardly sufficient to draw any meaningful conclusions. By carefully manipulating some of the instance parameters, we generated 30 diverse case studies that were inspired from 4 instances collected from the literature. In addition, we developed a tool to algorithmically generate arbitrarily many random instances of any size and with diverse characteristics. The case studies and the random instance generator will be made publicly available to hopefully foster further research on the problem of pre-positioning relief supplies.

For any of the instances of reasonable size, the NP-hard pre-positioning problem becomes intractable for commercial solvers such as CPLEX and thus calls for a suitable heuristic solution procedure. If possible, it seems natural to resort to an exact solver to make optimal choices for one of the three sets of decisions $\mathbf{x}, \mathbf{y}$ or $\mathbf{z}$, and optimize the remaining two sets of decision variables heuristically. Since the inventory decisions $\mathbf{y}$ are continuous and consequently difficult to tackle heuristically, it might seem reasonable to use CPLEX to solve the inventory sub-problem. However, a matheuristic that employs iterated local search techniques to look for good location and inventory configurations $\mathbf{x}$ and $\mathbf{y}$ and uses CPLEX to find the optimal solution of the aid distribution sub-problem $\mathbf{z}$ proves to be a better strategy. Numerical experiments on the set of instances we generated show that the matheuristic outperforms CPLEX for any given computation time, especially for large instances.

References

[1] EM-DAT: The Emergency Events Database - Université catholique de Louvain (UCL) - CRED, D. Guha-Sapir - www.emdat.be, Brussels, Belgium. `http://www.emdat.be/`

[2] Thomas, Anisya S. and Kopczak, Laura Rock. From logistics to supply chain management: the path forward in the humanitarian sector. *Fritz Institute* 15, 1-15 (2005).

[3] Burcu Balcik, Cem Deniz Caglar Bozkir, and O Erhun Kundakcioglu. A literature review on inventory management in humanitarian supply chains. *Surveys in Operations Research and Management Science* 21(2) 101-116 (2016).

Easily Building Complex Neighbourhoods With the Cross-Product Combinator

Fabian Germeau Yoann Guyot
Gustavo Ospina Renaud De Landtsheer
Christophe Ponsard
CETIC Research Center, {rdl,yg,go,fg,cp}@cetic.be

OscaR.cbls supports a domain specific language for modelling local search procedures called *combinators*[1, 2]. In this language, local search neighbourhoods are objects and they can be combined into other neighbourhoods. Such mechanisms provide great improvement in expressiveness and development time, as is targeted by the OscaR.cbls framework. This presentation focuses on one combinator that is the cross product of neighbourhoods. Let's consider C the cross-product of neighbourhoods A and B. It is built in source code as follows:

```
val C = A andThen B
```

The moves of C are all the chaining between moves of A and B. Practically, A and B are not aware of each other, so that their implementation does not need to be adapted to this setting. When neighbourhood C is explored, it explores neighbourhood A and gives it an instrumented objective function that triggers an exploration of B every time it is evaluated by neighbourhood A. This gives rise to search trees like the one illustrated in Figure 2 where x is the current state where C is explored, the edges labelled a_1 and a_2 represent moves of A, $x[a_1]$ represent the state x after applying the move a_1, and $x[a_1, b_1]$ represent the state $x[a_1]$ after applying the move b_1.

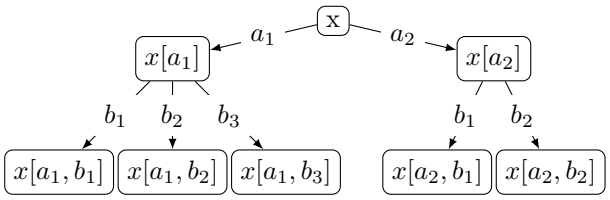


Figure 2: Exploration tree of neighbourhood C

An additional combinator was introduced to prune such search trees. Typically, when taking the cross-product of two neighbourhoods, the moves to be considered by the second one can be restricted according to the move currently explored by the first neighbourhood. Let's consider a pick-up & delivery problem (PDP), and A and B are neighbourhoods that insert pick-up points and delivery points, respectively. We may typically wish to restrict the points to insert by B to the delivery point that is related to the pick-up point that A is trying to insert, and consider only position

that occur after the position where A is trying to insert. To this end, a *dynAndThen* combinator was introduced. It is instantiated in source code as follows, where a is the type of move explored by A:

```
val D = A dynAndThen(a => B)
```

On the right-hand side of this combinator, the user specifies a function that inputs a move from A and returns a neighbourhood B to be explored once. This function is called by the combinator to generate a neighbourhood B that is specific to each move of A. This function is the opportunity for the operational research developer to specify two things: First, the neighbourhood B can be focused on relevant neighbours, as suggested above for the PDP. This can be specified to B through parameters passed when instantiating it. Second, some pruning can be performed at this stage, such as checking the violation of some strong constraint. Consider a PDP with time window constraint (so a *PDPTW*), we can check in this function that inserting the pick-up point does not violate any time window. If the time window constraint is violated at this stage, the search tree does not need to be explored further, so that the function can return null and the current move of A is discarded. An example of insert neighbourhood for PDPTW built using the `dynAndThen` combinator and featuring some pruning is given below.

```
val InsertPDP =
  insertPoint(nonRoutedPickupPoints, ...)
  dynAndThen((currentMove:InsertPointMove) =>
    if(timeWindowConstraint.violation == 0){
      insertPoint(relatedDelivery(currentMove.insertedPoint),
                  nodesAfter(currentMove.positionOfInsert),
                  ...)
    }else null)
```

Another very powerful combinator is the *Mu* that roughly is the repetitive cross-product of a neighbourhood with itself with a maximal number of cross products:

```
Mu(A,d) = A andThen A andThen A andThen ...          //"d" times
```

More elaborated version of this combinator are available and make it possible to share information between explorations of A like the `dynAndThen`. The Lin-Kernighan neighbourhood for instance can be built as a `Mu(TwoOpt)` with some additional pruning.

Acknowledgements

This research was conducted under the SAMOBI CWALITY research project from the Walloon Region of Belgium (grant nr. 1610019).

References

[1] OscaR Team. OscaR: Operational research in Scala, 2012. Available under the LGPL licence from `https://bitbucket.org/oscarlib/oscar`.

[2] Renaud De Landtsheer, Yoann Guyot, Gustavo Ospina, and Christophe Ponsard. *Recent developments of metaheuristics*, chapter Combining Neighborhoods into Local Search Strategies, pages 43–57. Springer, 2018.

Generic Support for Global Routing Constraint in Constraint-Based Local Search Frameworks

Renaud De Landtsheer
CETIC Research Center, rdl@cetic.be

Quentin Meurisse
Université de Mons, e-mail: quentin.meurisse@student.umons.ac.be

The OscaR.cbls framework supports a declarative constraint-based local search language for declaring optimization problems [1]. It features a new variable of type *sequence of integers* [2].
This variable was designed to embed global constraints into the framework. A classic example is a global constraint that maintains the length of the route in routing optimization. It inputs a distance matrix specifying the distance between each pair of nodes of the routing problem and the current route. When flipping a portion of route (a,b,c,d become d,c,b,a), and if the distance matrix is symmetric, smart algorithms are able to update the route length in $O(1)$-time [3].
The sequence variable of OscaR.cbls communicate their update in a symbolic way, even for complex structured moves such as flips or moving some sub-sequences. The sequence variable also features a checkpoint mechanism through which global constraints are notified that a neighbourhood exploration will start exploring around the current value of the variable. The proper time for global constraint to perform some pre-computation is when they receive a message notifying them about the definition of such a checkpoint. The goal of the presented contribution is to make the development of global constraint easier.

A Stereotype of Global Routing Constraints

We noticed that lots of global constraints on sequences rely on pre-computation and actually define a mathematical structure $(T, +, -)$ where T is a type, $+$ and $-$ are opposite of one another. In mathematical terms this structure is a non-Abelian group with neutral element, inverse values (for each x there exists a $-x$). We do not assume that $+$ is commutative, that's why it is non-Abelian. Each node has an associated value of type T, represented by the function:

$$\text{value(node)} : T$$

The classical pre-computation performed by global invariant is to associate to each node a value precomputed(node) defined as follows:

$$\text{precomputed(node)} : T = \sum_{n \in \text{PathBetween(vehicleStart,node)}} \text{value}(n)$$

All pre-computed values can be generated in a single pass over the current route, this in $O(n)$-time by iterating over the sequence and using an accumulator. When

considering a fraction of path represented by a pair of nodes (a, b), they rely on this formula to extract in $O(1)$ the sum of all *value* of nodes between a and b:

$$\sum_{n \in \text{PathBetween}(a,b)} \text{value}(n) = \text{precomputed}(b) - \text{precomputed}(a) + \text{value}(a)$$

The global gain is that pre-computations are performed once per neighbourhood exploration and cost $O(n)$-time, and can be queries in $O(1)$-time for each explored neighbour.
Consider the routing distance presented here above for the asymmetric case. We can perform some pre-computation to ensure that the overall route length can be updated in $O(1)$-time in case a portion of the route is flipped. To do this, we associate with each node a couple of integer, the first one is the length of the incoming hop, and the other one is the length of the incoming hop supposing it is taken in reverse direction. This defines T as (Int, Int) and the value function. We also define $+$ and $-$ as the pairwise sum and difference on the couple of integers.

A Generic Support For Global Constraints

Our contribution is an abstract invariant class. To instantiate it, the user must define the T, $+$, $-$ and value. The provided mechanics manages the execution of the pre-computation and the queries to the pre-computation. The user is also required to add some gluing code at the end to decode the resulting T values computed by the provided mechanics.
Pre-computation being an expensive operation, we designed our framework so that it can exploit pre-computation even if the sequence was modified since it was performed. This is achieved by keeping track of how the sequence was modified since the pre-computation was performed, and being able to translate queries on the actual sequence into queries performed on the latest available pre-computation.

Acknowledgements

This research results from an internship of UMons, Belgium, at CETIC Research Centre, Belgium, and supervised under the SAMOBI CWALITY research project from the Walloon Region of Belgium (grant nr. 1610019).

References

[1] OscaR Team. OscaR: Operational research in Scala, 2012. Available under the LGPL licence from `https://bitbucket.org/oscarlib/oscar`.

[2] Renaud De Landtsheer, Yoann Guyot, Gustavo Ospina, and Christophe Ponsard. Adding a sequence variable to the oscar.cbls engine. In *Proc. ORBEL'31: conference of the Belgian Operational Research Society*, pages 42 – 43, 2017.

[3] F.W. Glover and G.A. Kochenberger. *Handbook of Metaheuristics*. International Series in Operations Research & Management Science. Springer US, 2003.

Formalize neighbourhoods for local search using predicate logic

San Tu Pham Jo Devriendt
Patrick De Causmaecker
KU Leuven, Department of Computer Science
e-mail: san.pham, jo.devriendt, patrick.decausmaecker@kuleuven.com

Local search heuristics are powerful approaches to solve difficult combinatorial optimization problems and are particularly suitable for large size problems. However, local search approaches are often expressed in low-level concepts, which are not suitably formal to allow for modularity and reusability [1]. Related to this issue, Swan et. al. [2] point out the need of a pure-functional description of local search and metaheuristics. Lastly, Hooker opined in [3] that "Since modelling is the master and computation the servant, no computational method should presume to have its own solver. [...] Exact methods should evolve gracefully into inexact and heuristic methods as the problem scales up." The above issues inspire our work on formalizing local search heuristics, which should lead to modular, reusable and machine-interpretable formalizations of local search heuristics.

A local search heuristic consists of 3 layers [4]: search strategy, neighbourhood moves and evaluation machinery. While the first and last layers are most likely problem-independent, neighbourhood moves are usually problem-specific. Given a set of neighbourhoods moves, several local search heuristics can be created for a single problem. Small changes in problem requirements are also easy to be taken into account without affecting most of the neighbourhoods and the general algorithms. Therefore, formalizing neighbourhood moves is the first step towards formalizing local search heuristics.

In this work, we formally describe neighbourhood moves and their corresponding move evaluations in *FO(.)* [5], a rich extension of first-order logic, in which a problem's constraints and objective function can also be formalized. We also extend *IDP*, an automated reasoning system , with built-in local search heuristics, e.g. first improvement search, best improvement search and local search, to obtain a framework that employs the modular neighbourhood formalizations to simulate a local search algorithm.

This framework is demonstrated by reusing the formalizations of three different optimization problems: the Travelling Salesman Problem, the assignment problem and the colouring violations problems. The numerical results show the feasibility of our work.

References

[1] Hentenryck, Pascal Van, and Laurent Michel. Constraint-based local search. The MIT press, 2009.

[2] Swan, Jerry, et al. "A research agenda for metaheuristic standardization." Proceedings of the XI Metaheuristics International Conference. 2015.

[3] Hooker, John N. "Good and bad futures for constraint programming (and operations research)." Constraint Programming Letters 1 (2007): 21.

[4] Estellon, Bertrand, Frédéric Gardi, and Karim Nouioua. "High-Performance Local Search for Task Scheduling with Human Resource Allocation." SLS. 2009.

[5] De Cat, Broes, et al. "Predicate logic as a modelling language: The IDP system." arXiv preprint arXiv:1401.6312 (2014).
`http://www-cs-faculty.stanford.edu/~uno/abcde.html`

What is the impact of a solution representation on metaheuristic performance?

Pieter Leyman and Patrick De Causmaecker
CODeS, Department of Computer Science & imec-ITEC, KU Leuven KULAK
e-mail: pieter.leyman@kuleuven.be, patrick.decausmaecker@kuleuven.be

In metaheuristics, diversification and intensification are typically the driving forces of the search process. Diversification allows for exploration of different regions in the solution space, whereas intensification focuses on promising regions. Metaheuristics, however, often operate on a representation of the solutions, rather than on solutions themselves [1]. Hence, it is worth investigating the impact of the selected representation on algorithm performance. Furthermore, the decoding procedures, which translate a solution representation to an actual solution, should be examined as well. In this research, we analyze the impact of different solution representations, and their respective decoding schemes, in a project scheduling context, namely for the discrete time/cost trade-off problem with net present value optimization (DTCTP-NPV) with different payment models. We compare the following three representations:

- Activity list (AL): this type of list containts all activity numbers in the order in which they should be scheduled, i.e. the priority of different activities is taken into account. Care has to be taken such that activities do not accur in the list before any preceeding activities. A decoding procedure is also needed to translate the AL to a schedule with actual activity finish times.
- Finish time list (FTL): as the name implies, the FTL contains the finish times for all activities. As a result, no explicit decoding procedure is needed, but repair methods are required to ensure that no precedence restrictions are violated, and that the project deadline is not exceeded.
- Slack list (SL): the slack list was first proposed by [2], and contains a value $r_i \in [0; 1[$ for each activity i. This value determines both the activity priority and the usage of available activity slack. A decoding procedure translates the r_i value to an activity finish time. The resulting schedule is always feasible.

Table 3 provides a summary of the charactersitics of the three solution representations, namely the inclusion of schedule (i.e. finish times) and priority information, the feasibility, and the intricacy of the decoding scheme.

	Schedule	Priority	Feasible	Intricacy
AL		X	?	High
FTL	X		?	Low
SL	X	X	Y	Medium

Table 3: Characteristics per representation.

To compare the three representations, we employ the iterated local search (ILS) displayed in Algorithm 2. Only the perturbation and decode solution steps differ between the representations.

Algorithm 2 Iterated local search

ILS()
PREPROCESSING()
$x \leftarrow$ GENERATE-INITIAL-SOLUTIONS($\#Init$)
$x^* \leftarrow x$; $f(x^*) \leftarrow f(x)$
$T \leftarrow T_0$
repeat
  $x' \leftarrow$ PERTURBATE-MODES(x)
  $x'' \leftarrow$ PERTURBATE-ACTIVITIES(x')
  $f(x'') \leftarrow$ DECODE-SOLUTION(x'')
  if $f(x'') > f(x)$
    $x \leftarrow x''$; $f(x) \leftarrow f(x'')$
    if $f(x'') > f(x^*)$
      $x^* \leftarrow x''$; $f(x^*) \leftarrow f(x'')$
  else
    $x \leftarrow x''$; $f(x) \leftarrow f(x'')$ with probability $exp(\frac{f(x'')-f(x)}{T})$
  $T \leftarrow T * \Delta T$
until stopping criterion met
return x^*, $f(x^*)$

In Table 4, a summary of the algorithm's results, in terms of average project NPV ($AvNPV$), is provided. The values between brackets are p-values of two paired samples t-tests. We can conclude that the SL significantly outperforms both the AL and FTL for the DTCTP-NPV.

		SL	**AL**	**FTL**
Act	25	100.09	94.69	92.89
	50	157.14	152.24	145.20
	75	195.70	189.05	180.19
	100	228.00	220.71	210.77
Overall		170.23	164.17	157.26
			(<0.0001)	(<0.0001)

Table 4: Summary of results ($AvNPV$).

References

[1] Eiben, A.E. & Smith, J.E. (2015). *Introduction to Evolutionary Computing - Second Edition.* Springer.

[2] Leyman, P. (2016). *Heuristic algorithms for payment models in project scheduling.* PhD dissertation (Ghent University).

Construction of an Automated Examination Timetabling System for École Polytechnique de Louvain

Adrien Brogniet and Charles Ninane
UCL

Optimal timetables for temporarily unavailable tracks

Sander Van Aken
KU Leuven

Analysis of a Problem in Product Pricing

Fränk Plein
ULB

Dynamic programming algorithms for energy constrained lot sizing problems

Ayse Akbalik
LCOMS, Université de Lorraine, Technopole de Metz, FRANCE
e-mail: ayse.akbalik@univ-lorraine.fr

Christophe Rapine
LCOMS, Université de Lorraine, Technopole de Metz, FRANCE
e-mail: christophe.rapine@univ-lorraine.fr

Guillaume Goisque
Université de Lorraine, Technopole de Metz, FRANCE
e-mail: guillaume.goisque@univ-lorraine.fr

In this paper, we consider a single-item lot sizing problem under an energy constraint. We have to satisfy a demand known over a time horizon of T periods in a production system composed of M parallel, identical and capacitated machines. In each period, we have a limited amount E of available energy. Each machine consumes a certain amount of energy when being turned on, when producing units and simply when being on, whatever it produces or not. We call this last consumption the *running energy consumption*. Depending on the application, the start-up and running consumptions may be very large (for instance for ovens in metallurgical processes) or on the contrary, relatively negligeable to the production energy consumption. For that reason, we do not make any particular assumptions on these consumptions. The objective is to find a feasible production planning, that is, satisfying the demand and respecting the amount of available energy in each period, at minimal cost. In this work, in addition to the classical lot sizing costs (joint setup cost, unit production cost and unit inventory cost), we also consider a start-up cost, incurred each time a machine is turned on, and a running cost, incurred by each machine ready to produce, whatever it actually produces units or not. All these cost parameters are allowed to be time-dependent in our model. We call this problem *energy*-LSP.

Most of the papers published in the domain of energy-aware production planning focus on energy-efficient machine scheduling problems (see [1]). To the best of our knowledge, there are only a few studies in the literature coupling energy issues with discrete lot sizing problem: In [3] and [2], the authors consider respectively flow-shop and job-shop systems where they integrate some energy cost or constraints and propose heuristics to solve them. In [4], the problem studied in this work was first introduced. The authors propose an efficient $O(T \log T)$ time algorithm, but assuming stationary start-up costs, and considering that only one activity (start-up or production) consumes energy.

Notice that in a planning, we have to decide the quantity to produce in each period, but also the number of machines to turn on/off. On one hand, turning on machines

increases the capacity of the production systems in the subsequent periods, which can be necessary to tackle with a peak of demand, and potentially allows to reduce inventory levels. On the other hand, it consumes energy and incurs costs, at the period where the machine is turned on, and also at the subsequent periods due to the running energy consumption and running cost. Hence, in an optimal planning, some machines may have to be turned off, typically to save energy in a low demand period. One difficulty in this combinatorial problem is to arbitrate between the different activities consuming energy : In each period, we have to decide how the available amount E of energy is shared among the start-up of the machines and the production of units. We show this problem to be NP-hard if some energy parameters are time-dependent, even if almost all the cost parameters are null :

Theorem 1 *If the number M of machines is part of the instance, problem energy-LSP is NP-hard even with null production cost, null holding cost, null set-up cost and null running cost.*

On the contrary, we show that the problem is polynomially solvable if all the energy consumption parameters are stationary :

Theorem 2 *Problem energy-LSP can be solved in polynomial time in $O(M^6T^6)$ if all energy consumption parameters are stationary.*

Our result is based on dynamic programming. We show that, in a dominant solution, each production period with a positive entering stock either saturates the available capacity (all the machines that are not turned off produce at full capacity), or entirely consumes the available amount of energy (for production and/or for starting machines). Very classically, we decompose the problem into subplans, and evaluate the optimal cost of each subplan. The optimal cost can be obtained as a shortest path problem. The tricky part is the evaluation of the optimal cost of each subplan, where we have to fix the number of machines turned on at the beginning and at the end of the subplan.

References

[1] Biel, K. and Glock, C.H. (2016). Systematic literature review of decision support models for energy-efficient production planning. *Computers and Industrial Engineering*, 101, 243-259.

[2] Giglio, D., Paolucci, M. and Roshani, A. (2017). Integrated lot sizing and energy-efficient job shop scheduling problem in manufacturing/remanufacturing systems *Journal of Cleaner Production*, 148, 624-641.

[3] Masmoudi, O., Yalaoui, A., Ouazene, Y. and Chehade, H. (2017). Lot-sizing in a multi-stage flow line production system with energy consideration. *International Journal of Production Research*, 55 (6), 1640-1663.

[4] Rapine, C., Penz, B., Gicquel, C. and Akbalik, A. (2018). Capacity acquisition for the single-item lot sizing problem under energy constraints. To appear in Omega, DOI:10.1016/j.omega.2017.10.004.

Minimizing makespan on a single machine with release date and inventory constraints

M. Davari
KU Leuven, Research Center for Operations Management (Campus Brussels)
KU Leuven, Department of Computer Science, CODeS & imec-ITEC
e-mail: morteza.davari@kuleuven.be

M. Ranjbar
Ferdowsi University of Mashhad, Department of industrial Engineering
e-mail: m_ranjbar@um.ac.ir

R. Leus
KU Leuven, Research Group for Operations Research and Business Statistics
e-mail: roel.leus@kuleuven.be

P. De Causmaecker
KU Leuven, Department of Computer Science, CODeS & imec-ITEC
e-mail: patrick.decausmaecker@kuleuven.be

We consider a problem that has some interesting applications in truck scheduling. In a transshipment terminal, trucks arrive in a loading/unloading dock either to deliver (unload) or to pick up (load) a certain number of final products. The terminal is equipped with an inventory that is used to store these final products. The goal is to sequence all trucks in such a way that the inventory constraints are satisfied while the makespan (the completion time of all loading/unloading tasks) is minimized.
This problem can be looked upon as a single machine scheduling problem in which a set of jobs (which advocate trucks) must be processed. In this problem, each job is characterized with a processing time, a release date and an inventory modification. Jobs with positive inventory modification imply unloading operations and jobs with negative inventory modification indicate loading operations. We assume that a loading job can be processed only if the central inventory level is adequately high.
Inventory constraints have been considered in a number of papers in the field of scheduling. However, there are a very limited number of papers considering machine scheduling subject to inventory constraints. We cite Briskorn and Leung [1] who show that a single machine scheduling problem with inventory constraints to minimize the maximum lateness is NP-hard. They also develop a set of branch-and-bound algorithms that solve instances of the mentioned problem. Unlike Briskorn and Leung [1] who consider single machine scheduling problems, Bazgosha et al. [2] consider a transshipment terminal with a set of parallel stations for loading and unloading operations. They consider this problem as a parallel machine scheduling problem with release dates and inventory constraints.
Recently, Ghorbanzadeh et al. [3] proposed a branch-and-bound algorithm and dy-

Method	n				
	10	20	30	40	50
DP	0.00(0)	0.11(0)	373.24(0)	-(96)	-(96)
BB	0.00(0)	1.94(0)	-(17)	-(26)	-(34)

Table 5: Average CPU times (in seconds) and number of unsolved instances within the time limit (out of 96) for different values of $n = 10, 20, 30, 40$ and 50. Note that average CPU times are reported only if all instances are solved for the associated setting.

namic programming algorithm to solve the problem studied in this paper. Alternatively, we propose a novel scheduling terminology and a solution approach, namely a block-based branch-and-bound algorithm, to solve the instances of our problem efficiently. The contributions of this paper are fourfold:

1. we prove that the introduced problem is NP-hard in the strong sense by a reduction from 3-Partition;

2. we introduce the concept of block-based scheduling with which we propose a novel problem formulation;

3. we provide some theoretical results for instances with certain structure, based on which we classify instances into two groups of 'easy' and 'hard' instances; and

4. we propose two MILP formulations and a block-based branch-and-bound algorithm to solve the problem until optimality.

We compare our proposed branch-and-bound (BB) algorithm with the dynamic programming approach (DP) (from [3]) and report some preliminary results in Table 1. These results suggest that although DP performs better for small instances, BB is better at solving large instances within the time limit (which is 1000 seconds) than DP.

This research is supported by COMEX (Project P7/36) and BELSPO/IAP Programme.

References

[1] Briskorn, D. & Leung, J. Y.-T. Minimizing maximum lateness of jobs in inventory constrained scheduling. *Journal of the Operational Research Society*, 2013, 64, 1851-1864

[2] Bazgosha, A., Ranjbar, M. & Jamili, N. Scheduling of loading and unloading operations in a multi stations transshipment terminal with release date and inventory constraints. *Computers & Industrial Engineering*, 2017, 106, 20-31

[3] Ghorbanzadeh, M., Ranjbar, M. & Jamili, N. Transshipment scheduling at a single station with release date and inventory constraints. Ferdowsi University, 2017

Stocking and Expediting in Two-Echelon Spare Parts Inventory Systems under System Availability Constraints

Melvin Drent
University of Luxembourg,
Luxembourg Centre for Logistics and Supply Chain Management
e-mail: melvin.drent@uni.lu

Joachim Arts
University of Luxembourg,
Luxembourg Centre for Logistics and Supply Chain Management
e-mail: joachim.arts@uni.lu

We consider a two-echelon spare parts inventory system consisting of one central warehouse and multiple local warehouses. Each warehouse keeps multiple types of repairable parts to maintain several types of capital goods. The local warehouses face Poisson demand and are replenished by the central warehouse. We assume that unsatisfied demand is backordered at all warehouses. Furthermore, we assume deterministic lead times for the replenishments of the local warehouses. The repair shop at the central warehouse has two repair options for each repairable part: a regular repair option and an expedited repair option. Both repair options have stochastic lead times. Irrespective of the repair option, each repairable part uses a certain resource for its repair. Assuming a dual-index policy at the central warehouse and base stock control at the local warehouses, an exact and efficient evaluation procedure for a given control policy is formulated. To find an optimal control policy, we look at the minimization of total investment costs under constraints on both the aggregate mean number of backorders per capital good type and the aggregate mean fraction of repairs that are expedited per repair resource. For this non-linear non-convex integer programming problem, we develop a greedy heuristic and an algorithm based on decomposition and column generation. Both solution approaches perform very well with average optimality gaps of 1.56 and 0.23 percent, respectively, across a large test bed of industrial size.

Drone delivery from trucks: Drone scheduling for given truck routes

D. Briskorn
University of Wuppertal, Departement of Production and Logistics
e-mail: briskorn@uni-wuppertal.de

N. Boysen
University of Jena, Departement of Operations Management
e-mail: nils.boysen@uni-jena.de

S. Fedtke
University of Jena, Departement of Operations Management
e-mail: stefan.fedtke@uni-jena.de

S. Schwerdfeger
University of Jena, Departement of Management Science
e-mail: stefan.schwerdfeger@uni-jena.de

Last mile deliveries with unmanned aerial vehicles (also denoted as drones) are seen as one promising idea to reduce the negative impact of excessive freight traffic. To overcome the difficulties caused by the comparatively short operating ranges of drones, an innovative concept suggests to apply trucks as mobile landing and take-off platforms. In this context, we consider the problem to schedule the delivery to customers by drones for given truck routes. Given a fixed sequence of stops constituting a truck route and a set of customers to be supplied, we aim at a drone schedule (i.e., a set of trips each defining a drone's take-off and landing stop and the customer serviced), such that all customers are supplied and the total duration of the delivery tour is minimized. We differentiate whether multiple drones or just a single one are placed on a truck and whether or not take-off and landing stops have to be identical. We provide an analysis of computational complexity for each resulting subproblem, introduce mixed-integer programs, and provide a computational study evaluating them.

Addressing Uncertainty in Meter Reading for Utility Companies using RFID Technology: Simulation Experiments

Christof Defryn
Maastricht University, School of Business and Economics
e-mail: c.defryn@maastrichtuniversity.nl

Debdatta Sinha Roy, Bruce Golden
University of Maryland, Robert H. Smith School of Business
email: debsroy@rhsmith.umd.edu, bgolden@rhsmith.umd.edu

Introduction

Utility companies read the electric, gas, and water meters of their residential and commercial customers on a regular basis. A short distance wireless technology called the radio-frequency identification (RFID) technology is used for automatic meter reading (AMR). An AMR system has two parts: an RFID tag and a vehicle-mounted reading device. The RFID tag is connected to a physical meter and encodes the identification number of the meter and its current reading into a digital signal. The vehicle-mounted reading device collects the data automatically when it approaches an RFID tag within a certain distance. Utility companies would like to design routes for the reading vehicles such that all customers (meters) in the service area are covered and the total length (cost) of the routes is minimized. The problem to be solved is a *close-enough vehicle routing problem* (CEVRP) over a street network. There are issues with RFID technology that gives the meter reading problem a great amount of inherent uncertainty. The signal transmitted by an RFID tag occurs at regular time intervals and not continuously, for extending the battery life of the RFID transmitters. This leads to the possibility of a missed capture of a signal if the truck with the receiver is within the range of the meter only for a short time. Also, the signal range of a meter can vary from the distance specified by the manufacturer of the radio frequency transmitters and receivers due to weather conditions, surrounding obstacles, and decreasing battery life of the transmitters. These unplanned missed reads can lead to increased costs for a utility company, because another vehicle has to be sent at a later time to read the missed meters.

Research Goal

Published papers have not considered the issues with RFID technology and thereby have not taken into account the inherent uncertainty in the meter reading problem. The main contribution of our research is to address the issues of RFID technology

by generating robust routes for the CEVRP that minimize the number of missed reads and to demonstrate these ideas by simulation experiments using real-world data from utility companies. We want to design routes that are shorter in length and are better at capturing the uncertain signals from meters.

Integer Programming Formulation

We formulate the meter reading problem with RFID technology as a two-stage integer program (IP). The Stage 1 IP finds the street segments that are to be traversed for reading each meter with a pre-specified chance of being read from the full route. The Stage 2 IP solves a mixed rural postman problem that adds deadhead segments to the solution of the Stage 1 IP to obtain the full route. The deadhead segments added in the Stage 2 IP increase the likelihood of reading the meters.
Let E be the set of the edges and A be the set of the arcs in a street network. We denote the mixed graph by $G = (V, E \cup A)$, where V is the set of nodes. Let $c_j \geq 0$ be the cost (length) of street segment j. Let I be the set of the meters. Let p_{ij} be the probability that meter i is read at least once from street segment j. Let $L_i \in [0, 1]$ be the specified likelihood of reading meter i from the full route. We define x_j to be the decision variable denoting whether or not street segment j should be traversed in the full route. We consider a single meter reading vehicle. The Stage 1 IP formulation is given by the following.

$$\min \quad \sum_{j \in E \cup A} x_j \tag{7}$$

$$\min \quad \sum_{j \in E \cup A} c_j x_j \tag{8}$$

$$\text{s.t.} \quad \prod_{j \in E \cup A} (1 - p_{ij})^{x_j} \leq (1 - L_i) \quad \forall i \in I \tag{9}$$

$$x_j \in \{0, 1\} \quad \forall j \in E \cup A \tag{10}$$

This is a bi-objective formulation. Objective functions (7) and (8) minimizes the total number of street segments and the total cost (length) respectively. Constraint (9) selects the values of the decision variables (x_j) accordingly so that the probability of reading meter i from the full route is at least L_i. Constraint (10) restricts the decision variables to 0 and 1. Note that constraint (9) can be linearized in the decision variables $\sum_{j \in E \cup A} x_j \times \log(1 - p_{ij}) \leq \log(1 - L_i)$ for every meter $i \in I$ yielding a linear Stage 1 IP.

Regression and Bayesian Updating

In order to solve the Stage 1 IP, we need to estimate the values of the p_{ij}'s. We use a regression model. The data are in the form of 1 and 0, where 1 indicates that meter i is read from street segment j and 0 indicates that meter i is not read from street segment j. The predicted values of the dependent variable in the regression

model have to be between 0 and 1, which will denote the probability p_{ij}. Based on the type of the data we have and our requirements on the dependent variable, logit and probit models are considered.
Every time the meter reading vehicle collects readings, it adds more data to the previous readings. The more data we have, the better will be the estimates of the p_{ij}'s. Therefore, the routes generated by the two-stage IP will be of a higher quality. They will be better at capturing the uncertain signals thereby reducing the number of missed reads.
There are some serious issues if we use regression to update the estimates of the p_{ij}'s at every stage with the new data. Suppose in time period 1 we observe the first set of i.i.d. data (denoted by y_1). We run the regression on y_1. In time period 2, we observe a second set of i.i.d. data (denoted by y_2), independent of the first set. We run the regression on y_1 and y_2 together as a single data set, and so on. We are regressing on the older data sets repeatedly which makes this process of updating inefficient. Data sets from different time periods are given equal weights in the regression which should not be the case in practice. If we update the estimates of the p_{ij}'s at every stage when new data comes in using concepts from Bayesian statistics, then we can avoid the two drawbacks faced while updating using regression. Bayesian updating to estimate the p_{ij}'s can be done for both logit and probit models. A more complex method for estimating the p_{ij}'s is to perform Bayesian updating for hierarchical probit models. The estimates of the p_{ij}'s from hierarchical probit models are more accurate but difficult to calculate as compared to the logit and probit models. Hierarchical probit models account for the uncertain behavior of each meter separately while also accounting for the similarity between meters.

Simulation Experiments and Conclusions

Bayesian updating helps us to solve the two-stage IP at every time period when new data comes in, thereby potentially helping to produce more robust routes by updating the estimates of the p_{ij}'s. The main contribution of our research is combining vehicle routing with Bayesian statistics and data analytics to address the uncertainty in the meter reading problem. As mentioned, we will demonstrate these ideas and compare the performance of the different Bayesian updating models using real-world data and simulation experiments.

A UAV Location and Routing Problem with Synchronization Constraints

Ertan Yakıcı
National Defense University, Turkish Naval Academy,
Departement of Industrial Engineering
e-mail: eyakici@dho.edu.tr

Mumtaz Karatas
National Defense University, Turkish Naval Academy,
Departement of Industrial Engineering
e-mail: mkaratas@dho.edu.tr

Oktay Yılmaz
National Defense University, Turkish Naval Academy,
Departement of Industrial Engineering
e-mail: oktay.yilmaz.p337@gmail.com

An Unmanned Aerial Vehicle (UAV) is a vehicle which can fly with no pilot on board. These aircraft are often computerized and fully or partially autonomous. They can perform tasks that would otherwise prove difficult and risky for manned aircraft, e.g. attacking nuclear, biological and chemical infrastructures and well defended combatant platforms on land or at sea. Today, UAVs are employed for many military missions. Their superior flight capabilities in terms of range, altitude, and endurance, improve military power. UAVs can execute a number of roles in military operations including use of UAVs for intelligence, surveillance/reconnaissance missions which would take advantage of sustaining long flight times, positioning close to potential targets, and relative difficulty of being detected.

However, these capabilities and advantages come at the cost of increased complexity in efficient planning of the resources. The complications with UAV resource planning especially arise in littoral and confined geographies such as the coast of Mediterranean Sea, where small ships can station at specific points and conduct intelligence or surveillance/reconnaissance missions. These tasks bring about the problem of developing optimal UAV stationing and routing plans.

In this study, a problem which aims to optimize location and routing of a homogeneous UAV fleet is introduced. The problem allocates the available location capacity (ships) to the potential locations while it sustains the feasibility defined by synchronization constraints which include time windows at visited points and capacity monitoring in the stations.

First, we develop a Mixed Integer Linear Programming (MILP) formulation for defining the explained problem formally and providing a basis to convey the problem to commercial solvers. However, state-of-the-art commercial MILP solvers fail in solving relatively large-size operational combinatorial problem instances which combines

location and routing in reasonable times. For this reason, we have implemented an Ant Colony Optimization (ACO) meta-heuristic tailored to our problem for obtaining good solutions within short times. We have utilized a methodology inspired by an ACO approach which is called MAX-MIN Ant System (MMAS). The algorithm is tailored to handle the assignment of ships, multiple sorties of UAVs and several practical assumptions.
While our specific motivation comes primarily from naval intelligence or surveillance/reconnaissance missions, most aspects of our approach can be applicable to other combinatorial UAV LRPs observed in both military and civilian fields, including data collection and security, attacking stationary or moving targets.

Extensions of PROMETHEE to multicriteria clustering: recent developments

Y. De Smet, J. Rosenfeld and D. Van Assche
Université libre de Bruxelles, Ecole polytechnique de Bruxelles,
Computer and Decision Engineering departement, SMG research unit
e-mail: jean.rosenfeld@ulb.ac.be

PROMETHEE[1] belongs to the family of so-called *outranking* multicriteria methods. They have been initiated and developed by Prof. Jean-Pierre Brans since the 80*s*. For the last thirty years, a number of researchers have contributed to its methodological developments and applications to real problems. In 2010, Behzadian et al. already reported more than 200 applications published in 100 journals. These cover finance, health care, environmental management, logistics and transportation, education, sports, etc. The successful application of PROMETHEE is certainly due to its simplicity and the existence of user-friendly software such as PROMCALC, Decision Lab 2000, Visual PROMETHEE and D-SIGHT.
PROMETHEE has initially been developed for (partial or complete) ranking problems. Later, additional tools have been proposed like for instance *GAIA* [2] for the descriptive problematic or *PROMETHEE V* for portfolio selection problems. Following this trend, J. Figueira, Y. De Smet and J.P. Brans proposed, in 2004, an extension of PROMETHEE for sorting and clustering problems. Unfortunately, these first extensions presented some limits and therefore were never published in a scientific journal (but remained accessible as a technical report of the SMG research unit). Today, this work has been cited 78 times (according to Google Scholar) and has led to the development of different algorithms for sorting and clustering (such as for instance FlowSort or P2CLUST).
The aim of this presentation is to present recent advances in multicriteria clustering methods based on the PROMETHEE methodology. After a brief summary of existing procedures we will focus on related research questions such as the development of partition quality indicators or the integration of clusters size constraints.

[1] Preference Ranking Method for Enrichment Analysis Method
[2] Graphical Analysis for Interactive Assistance

More Robust and Scalable Sparse Subspace Clustering Based on Multi-Layered Graphs and Randomized Hierarchical Clustering

Maryam Abdolali
Amirkabir University of Technology, Department of Computer Engineering
e-mail: mabdolali@aut.ac.ir

Nicolas Gillis
Université de Mons, Faculté Polytechnique
e-mail: nicolas.gillis@umons.ac.be

For many decades, it was assumed that high-dimensional data were drawn from a *single* low-dimensional subspace/manifold and have fewer degree of freedom compared to the high ambient dimension. This assumption led to development of several dimension reduction methods such as principal component analysis (PCA). However, in many machine learning and computer vision applications, the data sets are mixtures of hybrid data with different intrinsic structures and are better approximated with *multiple* low-dimensional subspaces. This leads to the general problem of subspace clustering which is defined as partitioning the data into multiple clusters based on their intrinsic subspaces and identifying the parameters of each low-dimensional subspace [1].

Among the recently developed methods for partitioning the data from union of subspaces, the sparse subspace clustering (SSC) method using ℓ_1 regularization which is based on the sparse self-expressiveness property of the data is one of the most effective one [2]. The property states that each data point can be represented as a sparse linear combination of the points on the same subspace. This is combined with classic spectral clustering applied on the undirected graph which is obtained from the sparse coefficients. However, despite theoretical guarantees and empirical success of SSC, it is not practical for large-scale datasets with over $n \sim 10^4$ data points because n linear programs in n variables have to be solved.

To overcome the scalability issue of SSC, we propose a multi-layered graph framework that efficiently selects multiple sets of few common representative anchor points using a fast randomized hierarchical clustering technique [3]. Screening out large number of data points drastically reduces time complexity and memory requirements of SSC. A multilayered graph structure is utilized to merge the results of different sets of anchor points. This structure searches for a summary representation which is close to subspace representation of each graph on the Grassmann manifold while keeping the shared connectivity among different layers of graphs. We investigate the properties of our approach in challenging cases of noisy data and close subspaces. Numerical results on both synthetic data and real-world data show the effectiveness of our proposed framework.

References

[1] R. Vidal, *Subspace clustering*, IEEE Signal Processing Magazine 28(2), 52-68, 2011.

[2] E. Elhamifar and R. Vidal, *Sparse subspace clustering: Algorithm, theory, and application*, IEEE Transactions on Pattern Analysis and Machine Intelligence 35(11), 2765-2781, 2013.

[3] X. Dong, P. Frossard, P. Vandergheynst and N. Nefedov, *Clustering on multi-layer graphs via subspace analysis on Grassmann manifold*, IEEE Transactions on Signal Processing 62(4), 905-918, 2014.

Design and implementation of a modular distributed and parallel clustering algorithm

Landelin Delcoucq
University of Mons, Faculty of Engineering, Mons, Belgium,
e-mail: landelin.delcoucq@umons.ac.be

Pierre Manneback
University of Mons, Faculty of Engineering, Mons, Belgium,
e-mail: pierre.manneback@umons.ac.be

The number of data generated per year will reach more than 44.000 billions of gigaoctets in 2020, ten times more than in 2013 and this is likely to continue according to an EMC/IDC survey [1]. This means more than 10.000 gigaoctets per person and per year generated by the daily life. Nowadays, very large heterogeneous datasets are collected. The analysis of those data to be able to extract relevant information without getting lost in the vastness of the data represents a major challenge of the coming years. The raise of the amount of data due to the storage capacity big bang implies architectural modifications in the data storage and in the data management. Data mining methods have to adapt to those changes. We evolved from a single storage site to many distributed sites but the need of a single centralized data mining process is remained.

The most known and used solution developed to solve this problem is the Big Data. Big Data is a concept proposed and detailed by the Association for Computing Machinery in 1997. It represents a set of tools to deal with data and to reach a triple goal, the triple V (Volume, Variety and Velocity). The Big Data has to deal with very large datasets and those datasets are very heterogeneous because they are coming from different locations, with different structures or can be unstructured. In addition, the process has to reach a level of velocity and accuracy that involves its superiority against classical methods and tools. Therefore, the purpose of my work is to design and implement an algorithm to deal with large datasets while satisfying the triple V. The core of this algorithm being a clustering method, the K-Means algorithm.

The purpose of my work is to propose an implementation of K-Means using the pattern MapReduce. This pattern has been developed to deal with very large distributed datasets. This is an approach in which the K-Means clustering algorithm is a module which could be replaced by others clustering algorithms. That allows us to combine the advantages and to reduce the drawbacks. Clustering algorithms can't provide an unique solution and need many experimentations to reach the final

[1] The Digital Universe of Opportunities : rich data and the Increasing Value of the internet of things

solution. Those experimentations can be computationally expensive and require a lot of memory access then the optimization of the distribution and of the parallelization has to be highly considered. The distributed and parallel features of this algorithm will allow us to use it on data coming from multiple locations and the modular feature will allow us to use in on data heterogeneously structured.

MapReduce is a programming model for performing calculations on the data. It composed of two basics components : mapping functions and reducing functions. A MapReduce job can be divided into map tasks and reduce tasks, both that run parallel with each other. The map task converts a set of data into a set of individual elements constituted of tuples key-values. This key is the central component of the pattern. Values with the same key will be grouped in the next step. The reduce task takes outputs from the previous task and combines those tuples into another set of tuples.

The main feature of this algorithm is to use only the mapper to perform the K-Means algorithm itself and after that the algorithm will use multiple reducers to perform mergers between the clusters previously generated. This algorithm will process data in a parallel way because it will compute all the clusters in the same time at the first level of the algorithm. At the followings levels, the mergers will also be processed in a parallel way because the clusters will be separated into groups and the clusters into those groups will be compared and merged simultaneously, if they are similar, until there is only one group.

The algorithm will also process data in a distributed way because the Hadoop Distributed File System will allocate randomly the points to different nodes (and thus different processors). The distributed feature of this algorithm is managed by the pattern and the algorithm doesn't have an impact on it. The right configuration of the network is nonetheless a critical factor to the successful completion of the algorithm. For example, it has to be able to identify the nodes.

The mapper is completely independent of the reducer because the input of the reducers is clusters. It allows us to switch the algorithm used by the mapper. A first mapper can use a K-Means algorithm, a second can use a KNN algorithm and a third a DBSCAN algorithm, it will have no effect on the reducers. First of all, the mapper will be also used like a translator. The MapReduce pattern is designed on a "schema-on-read", the algorithm must adapt to the data. Thus, our algorithm will read the input, identify the corresponding pattern of the data and extract a common structure which will be used throughout the process.

The personalization of the mapper is a key point of the algorithm. The mapper is able to identify the node on which it is working and thus an application using this algorithm will be able to identify the node on which it is working, a node can be a datacenter representing a specific type of data. The application will be able to

modify its pattern to fit with the data.

(This work was realized in collaboration with the University College of Dublin as part of the Erasmus programme.)

On facility location problems and penalty-based aggregation

Raúl Pérez-Fernández and Bernard De Baets
KERMIT, Department of Data Analysis and Mathematical Modelling, Ghent University
e-mail: {raul.perezfernandez,bernard.debaets}@ugent.be

Facility location problems constitute a common study subject in operations research [9]. Formally, the aim is to – given a set of demand points, a set of potential locations for a facility and the transportation costs between each demand point and each location – find the location that minimizes the 'overall' transportation cost. When said 'overall' transportation cost is considered to be the sum (or, equivalently, the average) of the transportation costs between the facility location and each of the given demand points, one talks about the minimum facility location problem. Another classical facility location problem is the minmax facility location problem in which the 'overall' transportation cost is calculated as the maximum among the transportation costs between the facility location and each of the given demand points.

Closely related, aggregation problems consist in combining several elements into a single one [1, 4]. We pay special attention to penalty-based aggregation problems, which amount to selecting the consensus element that minimizes a penalty for disagreeing with the elements to be aggregated. Although the theoretical framework for penalty-based aggregation is mostly developed for the aggregation of real numbers [2, 3], one could see that this same rationale is also considered while dealing with other types of structures. For instance, the method of Kemeny [6] for the aggregation of rankings selects the ranking that minimizes the sum of the Kendall distances [7] to the rankings to be aggregated, or the closest string problem [8] for the aggregation of strings aims at finding the string that minimizes the maximum Hamming distance [5] to the strings to be aggregated.

In the case of real numbers, the penalty usually carries some desirable properties such as (quasi)convexity or (lower-semi)continuity, and, thus, the computation of the minimizers of the penalty (given a list of real numbers) oftentimes turns out to be straightforward. In other structures, exhaustive analyses of the computational complexity of finding the minimizer(s) of the penalty have been addressed. For instance, for a discussion on the computational complexity of computing medians of binary relations (minimizers of the penalty defined as the sum of the symmetric difference distances to the binary relations to be aggregated), we refer to [10]. In a more general setting in which the properties of the considered structure cannot be exploited, both abovementioned facility location problems might be considered. For instance, the Kemeny ranking and the median relation could be computed by solving a minimum facility location problem, and the closest string could be computed by

solving a minmax facility location problem. Overall, the minimum facility location problem might be used for computing minimizers of a penalty defined by the sum and the minmax facility location problem might be used for computing minimizers of a penalty defined by the maximum.

References

[1] G. Beliakov, A. Pradera, T. Calvo, Aggregation Functions: A Guide for Practitioners, vol. 221 of Studies in Fuzziness and Soft Computing, Springer, Berlin, Heidelberg, 2007.

[2] H. Bustince, G. Beliakov, G. P. Dimuro, B. Bedregal, R. Mesiar, On the definition of penalty functions in data aggregation, Fuzzy Sets and Systems 323 (2017) 1–18.

[3] T. Calvo, G. Beliakov, Aggregation functions based on penalties, Fuzzy Sets and Systems 161 (2010) 1420–1436.

[4] M. Grabisch, J.-L. Marichal, R. Mesiar, E. Pap, Aggregation Functions, Cambridge University Press, Cambridge, 2009.

[5] R. W. Hamming, Error detecting and error correcting codes, The Bell System Technical Journal 29 (2) (1950) 147–160.

[6] J. G. Kemeny, Mathematics without numbers, Daedalus 88 (4) (1959) 577–591.

[7] M. G. Kendall, A new measure of rank correlation, Biometrika 30 (1938) 81–93.

[8] J. K. Lanctot, M. Li, B. Ma, S. Wang, L. Zhang, Distinguishing string selection problems, Information and Computation 185 (2003) 41–55.

[9] S. H. Owen, M. S. Daskin, Strategic facility location: A review, European Journal of Operational Research 111 (1998) 423–447.

[10] Y. Wakabayashi, The complexity of computing medians of relations, Resenhas 3 (3) (1998) 323–349.

Improving the quality of the assessment of food samples by combining absolute and relative information

Bernard De Baets, Raúl Pérez-Fernández and Marc Sader
KERMIT, Department of Data Analysis and Mathematical Modelling, Ghent University
e-mail: {bernard.debaets,raul.perezfernandez,marc.sader}@ugent.be

A classical problem in food science concerns the assessment of the quality of food samples. Typically, several panellists are trained extensively on how to assess the quality of a given food sample on an ordinal scale regarding the perceived degree of spoilage [1, 2], the personal degree of liking [3, 4], or just the appearence [5, 7]. This training usually aims at teaching how to identify different quality indicators and how to make proper use of the considered ordinal scale. Several training sessions usually need to be scheduled, thus resulting in a time-consuming training that also carries big expenses.

For this very reason, it is common to search for an additional source of information by invoking untrained panellists [6]. Obviously, untrained panellists do not provide information as reliable as that provided by trained panellists, and, most of the times, they are simply unable to make use of the considered ordinal scale in an accurate manner. However, they are indeed able to compare different food samples with each other. Thus, untrained panellists are often required to rank several food samples rather than assessing the quality of each food sample individually.

We thus normally dispose of two types of information: absolute information gathered from trained panellists, in the form of labels on an ordinal scale, and relative information gathered from untrained panellists, in the form of rankings. In this presentation, we will explain how we can combine both types of information in order to improve the quality of the assessment of food samples.

References

[1] A. A. Argyri, A. I. Doulgeraki, V. A. Blana, E. Z. Panagou, G.-J. E. Nychas, Potential of a simple HPLC-based approach for the identification of the spoilage status of minced beef stored at various temperatures and packaging systems, International Journal of Food Microbiology 150 (2011) 25–33.

[2] A. A. Argyri, E. Z. Panagou, P. A. Tarantilis, M. Polysiou, G.-J. E. Nychas, Rapid qualitative and quantitative detection of beef fillets spoilage based on Fourier transform infrared spectroscopy data and artificial neural networks, Sensors and Actuators B: Chemical 145 (2010) 146–154.

[3] L. V. Jones, D. R. Peryam, L. L. Thurstone, Development of a scale for measuring soldiers' food preferences, Food Research 20 (1955) 512–520.

[4] D. R. Peryam, F. J. Pilgrim, Hedonic scale method of measuring food preference, Food Technology 11 (1957) 9–14.

[5] H. B. Rogers, J. C. Brooks, J. N. Martin, A. Tittor, M. F. Miller, M. M. Brashears, The impact of packaging system and temperature abuse on the shelf life characteristics of ground beef, Meat Science 97 (1) (2014) 1–10.

[6] M. Sader, R. Pérez-Fernández, L. Kuuliala, F. Devlieghere, B. De Baets, A combined scoring and ranking approach for determining overall food quality, submitted.

[7] M. Smolander, E. Hurme, K. Latva-Kala, T. Luoma, H.-L. Alakomi, R. Ahvenainen, Myoglobin-based indicators for the evaluation of freshness of unmarinated broiler cuts, Innovative Food Science and Emerging Technologies 3 (2002) 279–288.

Coordination and threshold problems in combinatorial auctions

Bart Vangerven
Bergische Universität Wuppertal, Schumpeter School of Business and Economics
e-mail: vangerven@uni-wuppertal.de
KU Leuven, Faculty of Economics and Business
e-mail: bart.vangerven@kuleuven.be

Dries R. Goossens
Ghent University, Faculty of Economics and Business Administration
e-mail: dries.goossens@ugent.be

Frits C.R. Spieksma
KU Leuven, Faculty of Economics Business
e-mail: frits.spieksma@kuleuven.be

Combinatorial auctions are allocation mechanisms that enable selling and buying multiple items simultaneously. In fact, combinatorial auctions allow bidders to bid on packages of items and the auctioneer can allocate any package only in its entirety to the corresponding bidder. Combinatorial auctions have established themselves as a viable allocation mechanism in settings where (i) market prices are not readily available (otherwise there is no need for an auction), and (ii) bidders have sub- or super-additive valuations (otherwise there is no need for a combinatorial auction, since sequential single item auctions would suffice). Combinatorial auctions offer the possibility for a *coalition* of bids on small packages to jointly outbid a single bidder's claim on the complete set of items. However, two hurdles need to be overcome before a coalition can become winning.

(i) The ***coordination problem***. As each item can be allocated at most once, bidders need to coordinate their bids and bid on complementary (i.e. non-overlapping) sets of items. The coordination challenge lies in bidders having to discover such a set of individually profitable and collectively complementary packages, given that the number of possible packages rises exponentially with the number of items. This is complicated by the existence of cognitive limits on the number of packages people can concentrate on during the auction. For instance, experimental research by [4] has shown that bidders only bid on six to ten different packages, independent of the auction format, although they had a multitude of packages with positive valuations to choose from. [3] also find that bidders only bid on a small number of packages, and that the majority of placed bids are on the myopically profitable packages. Furthermore, coordination is hindered by the assumption that a bidder only knows his/her private valuation for these packages, and not the preferences of other bidders. In fact, in order to mitigate collusion, it makes sense to restrict communication between bidders (see e.g. [2]).

(ii) The ***threshold problem***. Even if the coordination problem is overcome and a set of disjoint packages for which the combined valuation exceeds the currently winning bid is somehow identified, the task of determining appropriate bid prices to displace the currently winning bid still remains. A complicating factor is that each bidder in a coalition has an interest not to increase his/her bid. Indeed, the forgone revenue from unilaterally increasing one's bid falls entirely on the cooperating bidder while the benefits extend to the non-cooperating bidders as well. Note that, as [1] point out, the threshold problem is strictly speaking not a free-rider problem. In a true free-rider problem, the dominant strategy is never to cooperate. However, in the context of the threshold problem, coalition members may still have some incentive to contribute to the effort to overcome the threshold, since being the only bidder to cooperate will typically still be preferable to not cooperating and winning nothing. Our contribution is the clarification of the hitherto vaguely used concepts of coordination and threshold problems. Moreover, we develop a quantitative measure to express the severity of both problems. To the best of our knowledge, we are the first to do this.

References

[1] Bykowsky, M. M., Cull, R. J., and Ledyard, J. O. (2000). Mutually Destructive Bidding: The FCC Auction Design Problem. *Journal of Regulatory Economics*, 17(3):205–228.

[2] Cramton, P. and Schwartz, J. (2000). Collusive Bidding: Lessons from the FCC Spectrum Auctions. *Journal of Regulatory Economics*, 17(3):229–252.

[3] Kagel, J. H., Lien, Y., and Milgrom, P. (2010). Ascending Prices and Package Bidding: A Theoretical and Experimental Analysis. *American Economic Journal: Microeconomics*, 2(3):160–85.

[4] Scheffel, T., Ziegler, G., and Bichler, M. (2012). On the impact of package selection in combinatorial auctions: an experimental study in the context of spectrum auction design. *Experimental Economics*, 15(4):667–692.

Scheduling time-relaxed double round-robin tournaments with availability constraints

David Van Bulck
Ghent University, Faculty of Economics and Business Administration
e-mail: david.vanbulck@ugent.be

Dries Goossens
Ghent University, Faculty of Economics and Business Administration
e-mail: dries.goossens@ugent.be

Over the last four decades, operations-research has successfully been applied to many sports scheduling problems. Most of these problems, however, concern time-constrained round-robin tournaments, i.e. tournaments where all teams meet all other teams a fixed number of times and teams play once in a round (if the number of teams is even). In professional leagues, time-constrained formats have the advantage of producing accurate rankings since, after every round, each team will have played the same number of games. However, time-constrained formats are not very well suited for amateur competitions since they lack the flexibility to incorporate player and venue availability. Indeed, in amateur competitions venues are typically shared with other organizations, and players have other activities as well. In these situations, time-relaxed schedules are more suited since they contain (many) more slots than there are matches per team. Although time-relaxed schedules are widely used in practice, only a few papers (e.g. [2, 4, 6]) consider this format. This is in sharp contrast with the large strand of literature that deals with time-constrained scheduling [3].

In the time-relaxed double round-robin problem with availability constraints (`TRDRRA`), we are given a set of teams and a set of slots. To cope with venue and player availability, each team can provide a set of dates on which it can host a game, and a set of dates on which it cannot play at all. To avoid injuries, a team should ideally rest for at least R days between two consecutive matches. However, if this is not possible, we penalize the solution with a value of p_r each time a team has only $r < R$ days between two consecutive matches. In addition, a team can only play up to M games within a period of $R+1$ days. To increase the fairness of the generated schedules, we also explain how various metrics from Suksompong [7] can be integrated. The goal is to construct a schedule with minimum cost, scheduling all matches and respecting the availability constraints.

Since most traditional methods, such as the circle method or first-break-then-schedule [1, 3], focus on the alternation of home and away matches in a time-constrained setting, they are not appropriate to solve the `TRDRRA` problem. Therefore, we propose two new heuristics to solve our problem. The first is a genetic algorithm backed by a local improvement heuristic which is able to both repair and improve schedules, resulting in a memetic algorithm. Basically, the improvement heuristic sequentially

solves a transportation problem which schedules (or reschedules) all home games of a team. In this transportation problem, the set of supply nodes consists of all slots on which this team can play a home game, and the set of demand nodes consists of all opponents against whom it can play a home game.
The second heuristic is a fix-and-relax heuristic procedure [5] based on the teams as well as on time-intervals. This constructive method initially relaxes all variables. Next, it gradually replaces the fractional variables by resolving the model with the integrality constraints enabled for a small subset of relaxed variables. After each iteration, it then fixes the variable values of the selected group and repeats the procedure until all variables have an integral value. Finally, a simple ruin-and-recreate procedure tries to improve the generated solution.
We assess the performance of both algorithms over multiple independent runs for several instances from the literature [2, 6]. Overall, the quality of the generated schedules from both heuristics is comparable with the optimal solutions obtained with integer programming (GUROBI) but the computational effort required to obtain these solutions is considerably less.

References

[1] Dominique de Werra. Scheduling in sports. In P. Hansen, editor, *Studies on graphs and discrete programming*, pages 381–395, 1981.

[2] Dries Goossens and Frits Spieksma. Indoor football scheduling. In E. Özcan, E. Burke, and B. McCollum, editors, *Proceedings of the 10th International Conference on the Practice and Theory of Automated Timetabling*, pages 167–178. PATAT, 2014.

[3] Graham Kendall, Sigrid Knust, Celso C. Ribeiro, and Sebastián Urrutia. Scheduling in sports: An annotated bibliography. *Computers & Operations Research*, 37(1):1–19, 2010.

[4] Sigrid Knust. Scheduling non-professional table-tennis leagues. *European Journal of Operational Research*, 200(2):358–367, 2010.

[5] Beatriz Brito Oliveira, Maria Antónia Carravilla, José Fernando Oliveira, and Franklina MB Toledo. A relax-and-fix-based algorithm for the vehicle-reservation assignment problem in a car rental company. *European Journal of Operational Research*, 237(2):729–737, 2014.

[6] Jörn Schönberger, Dirk C Mattfeld, and Hubert Kopfer. Memetic algorithm timetabling for non-commercial sport leagues. *European Journal of Operational Research*, 153(1):102–116, 2004.

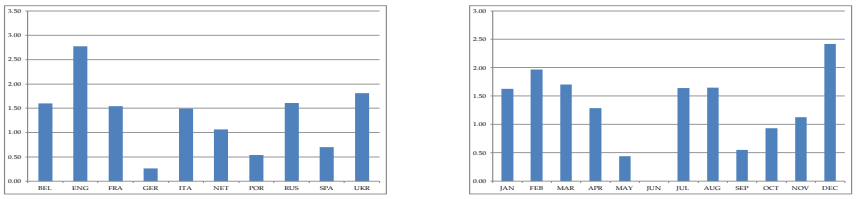
[7] Warut Suksompong. Scheduling asynchronous round-robin tournaments. *Operations Research Letters*, 44(1):96–100, 2016.

Combined proactive and reactive strategies for round robin football scheduling

X. J. Yi
Ghent University, Faculty of Economics and Business Administration
e-mail: xiajie.yi@ugent.be

D. Goossens
Ghent University, Faculty of Economics and Business Administration

This paper focuses on the robustness of football schedules, and how to improve it using combined proactive and reactive approaches. Although hard efforts are invested at the beginning of the season to create high-quality initial football schedules [1], these schedules are rarely fully played as planned. According to our data collection, reasons behind this are bad weather conditions, conflicts with other events like Champions League football, or political elections, and so on. We refer those reasons as *uncertainties*. Games that are postponed or canceled, inducing deviations from the *initial schedule* are defined as *disruptions*. Fixtures as they were effectively played are referred to as the *realized schedule*. We assume that disruptions can only be rescheduled on a date that constitutes a so-called *fictive round*, which is actually not present in the initial schedule. We studied ten main European leagues for the last fifteen seasons, from 2002 to 2017. Figure 3 shows an overview of the percentage of disruptions per league (a), and per month (b). It can be noticed that the frequency of disruptions differs considerably between different leagues. Moreover, a mild climate does not guarantee fewer disruptions.



(a) Average % of disrupted matches per league (b) Average % of disrupted matches per month

Figure 3: Overview of the disruptions

For the purpose of evaluating the quality of initial and realized schedules, measures based on *breaks*, *balancedness* and *failures* are used. A break corresponds to two consecutive games without an alternating pattern of home-away advantage [2]. Inspired by Knust and von Thaden [3], Nurmi *et al.* [4] introduced $k-$*balancedness* of a schedule as the biggest difference, k, between the number of home games and

away games after each round among all teams. Sometimes, there are problems with rescheduling disruptions because there is no fictive round available (also taking into account that a team can play at most once in each fictive round). We call an unsuccessfully rescheduled disruption a failure. A fair tournament should have each game (re)scheduled with as few breaks and smallest k value as possible. Nevertheless, in almost all cases without failure disruptions, the number of breaks increased, and in half of the cases, the balancedness of k worsened as well in the realized schedule. In order to develop schedules and rescheduling policies that are robust with respect to the aforementioned quality measures, we investigate a number of combined proactive and reactive approaches [5]. The proactive approaches focus on developing an initial schedule that anticipates the realization of some unpredicted events during the whole season. We consider two strategies in which fictive rounds are spread equally throughout the season and the second half of the season, respectively. Meanwhile, the reactive approaches revise and re-optimize the initial schedule when a disruption occurs. More specifically, after each round, it becomes clear how many games are disrupted and on which of the selected fictive rounds we want to reschedule the game. Three strategies are developed here: first available fictive round, best available fictive round based on breaks and best available fictive round based on k. Particularly, we discuss these strategies under two assumptions: (i) each disrupted match should be rescheduled right away and permanently, and (ii) disruptions can be rescheduled to a fictive round and this arrangement can be changed later if more information becomes available. Considering a double round robin (DRR) tournament with twenty teams, a mirrored canonical schedule [6] is used as an initial plan. We randomly generate disruptions to this schedule based on probability distributions resulting from our empirical study (Figure 3). Combining each proactive strategy with each reactive strategy, we compare the results on three performance indexes, i.e., breaks, k, and failures.

References

[1] Goossens, D., Spieksma, F., 2012. Soccer schedules in Europe: an overview. Journal of Scheduling, 15(5), 641-651.

[2] Rasmussen, R., Trick, M., 2008. Round robin scheduling – a survey. European Journal of Operational Research 188(3), 617-636.

[3] Knust, S., von Thaden, M., 2006. Balanced home-away assignments. Discrete Optimization 3, 354-365.

[4] Nurmi, K., *et al.*, 2010. A framework for a highly constrained sports scheduling problems. Proceedings of the international MultiConference of Engineers and Computer Scientists. Vol. 3, pp. 1991-1997. Newswood Limited.

[5] Herroelen, W., Leus, R., 2005. Project scheduling under uncertainty: Survey and research potentials. European Journal of Operational Research 165, 289-306.

[6] De Werra, D., 1981. Scheduling in sports, In: Hansen, P. (Ed.), Studies on Graphs and Discrete Programming. North-Holland, Amsterdam, pp. 381-395.

A constructive matheuristic strategy for the Traveling Umpire Problem

Reshma C. Chandrasekharan
KU Leuven, Department of Computer Science, CODeS - imec Belgium
e-mail: ReshmaCC@kuleuven.be

Túlio A. M. Toffolo
KU Leuven, Department of Computer Science, CODeS - imec Belgium
& Federal University of Ouro Preto, Department of Computing, Brazil
e-mail: tulio.toffolo@kuleuven.be

Tony Wauters
KU Leuven, Department of Computer Science, CODeS - imec Belgium
e-mail: tony.wauters@cs.kuleuven.be

Traveling Umpire problem (TUP) is a sports scheduling problem concerning the assignment of umpires to the games of a fixed round robin tournament. Introduced by Michael Trick and Hakan Yildiz in 2007, the problem abstracts the umpire scheduling problem in the American Major League Baseball (MLB). A typical season of MLB requires umpires to travel extensively between team venues and therefore, the primary aim of the problem is to assign umpires such that the total travel distance is minimized. Constraints that prevent assigning umpires to consecutive teams or team venues make the problem challenging. Instances comprising of 8 to 30 teams have been proposed and since the introduction, various exact and heuristic techniques have been employed to obtain near optimal solutions to small and medium-sized instances. However, exact techniques prove to be inefficient in producing high quality solutions for large instances of 26 to 32 teams which resembles the real world problem. The present work focuses on improving the solutions for the large instance and presents a decomposition based method implementing constructive matheuristics (CMH). A given problem is decomposed into subproblems of which IP formulations are solved sequentially to optimality. These optimal solutions of the subproblems are utilized to construct a solution for the full problem. Various algorithmic parameters are implemented and extensive experiments are conducted to study their effects in the final solution quality. The algorithm being constructive, parameters have to be tuned such that the solution for the current subproblem does not prevent the feasibility of the future subproblems. In addition, design parameters are utilized to ensure feasibility of constraints that cannot be locally evaluated with in each subproblem. Parameters such as size of the subproblems and amount of overlap between the subproblems are tuned such that the solution quality is maximized while the runtime lies within the benchmark time limit of 5 hours. Design parameters such as the objective function and future of subproblems ensure that the solution constructed from the optimal solutions of the subproblems continues to be feasible

in terms of the full problem.
The proposed method has been able to improve the current best solutions of all the large instances within the benchmark time limits. In addition, CMH is also able to improve solutions of two medium sized instances of 18 teams. Furthermore, CMH generates solutions those are comparable or better than the solutions generated by other similar heuristics. Experiments conducted so far on the TUP suggest the possibility of CMH being applied on other similar problems. Apart from the innovations in terms of design parameters that may improve the CMH algorithm, the CMH has the inherent property of getting faster with the evolution of better IP solvers. Insights on the applicability of CMH to similar optimization problems and the expected advantages or shortcomings may also be discussed.

Maximum likelihood estimation of discrete and continuous parameters: an MILP formulation

Anna Fernández Antolín, Virginie Lurkin, Michel Bierlaire
Transport and Mobility Laboratory,
School of Architecture, Civil and Environmental Engineering,
Ecole Polytechnique Fédérale de Lausanne
e-mail: {anna.fernandezantolin,virginie.lurkin,michel.bierlaire}@epfl.ch

The state-of-the-art for the mathematical modeling of disaggregate demand relies on choice theory inspired from micro-economics. In the estimation of discrete choice models, in general, only continuous parameters are considered, although advanced models include also discrete ones. The most typical example of a discrete parameter that is usually disregarded from the estimation process is the nest allocation parameter in nested logit models. Nesting structures are used in discrete choice models when correlation between alternatives is suspected. They are used in a very broad range of transportation contexts such as airline itinerary choice, car-type choice, route choice, and in mode choice among others. In route choice, for instance, two paths are correlated if they share a physical segment of the route. However, in other contexts, the partition of alternatives into different nests is less obvious and there are several nesting structures that make intuitive sense. In practice, to determine the most appropriate nesting structure, the analyst has several options: (i) to enumerate all the possible values, and estimate the continuous parameters for each combination, and (ii) to make the problem continuous by relaxing the integrality of the discrete parameters. For instance, a membership indicator becomes a continuous variable between 0 and 1 (like in the cross-nested logit model), or by making the membership probabilistic (like in latent class models). In both cases, however, the likelihood function features several local optima, so that classical nonlinear optimization methods may not find the (global) maximum likelihood estimates.

In this work, we propose a new mathematical model that is designed to find the global maximum likelihood estimates of a choice model involving both discrete and continuous parameters. We call our approach *discrete-continuous maximum likelihood* (DCML) because we introduce into the maximum likelihood framework binary parameters. We rely on simulation to formulate our problem as a mixed integer linear problem (MILP). This is a first attempt towards a complete MILP formulation of the maximum log likelihood, which results in a problem with high computational complexity. The goal of this presentation is to show under which circumstances our approach is computationally feasible, and to study its strengths and limitations. To do so, we use a stated preferences mode choice case study collected in Switzerland in 1998. The respondents provided information in order to analyze the impact of the modal innovation in transportation represented by the Swissmetro, a mag-lev underground system, compared to the usual transport modes of car and train.

Our contributions are multiple. First, to the best of our knowledge, we are the first

to include discrete parameters estimation in the maximum likelihood framework in the context of discrete choice models. Second, our model is formulated as an MILP. We use simulations and piecewise linear function approximation to dispose of the non-linearity of the log likelihood. We believe that it is the first time that the log likelihood is linearized. Finally, the proposed mathematical model is general and can be used with any choice model, as long as the distribution of the error terms can be simulated (e.g.: cross-nested, logit or latent class models).

Preliminary results

As the proposed framework relies on simulation, it is important to start by determining the minimum number of draws needed to obtain reliable values of the final log likelihood. To do so, we evaluate the objective function (which is the log likelihood function) at the values of the parameters obtained by a continuous estimation software. We do so for the logit model, and for the three possible nested logit models (N1,N2 and N3). The results are shown in Table 6, together with the value of the final log likelihood (FLL) obtained with the continuous estimation. The table also shows the relative error between the real FLL and the value obtained with the MILP. Let FLL be the true value of the final log likelihood, and $\widehat{FLL}_R$ the value obtained for R draws.

		N1		N2		N3		Logit	
	R	FLL	e_{FLL} [%]	FLL	e_{FLL} [%]	FLL	e_{FLL} [%]	FLL	e_{FLL} [%]
MILP	5	-1648	958	-1560	866	-1558	860	-1344	729
	10	-358.1	130	-678.2	320	-369.8	128	-657.9	306
	20	-152.8	1.93	-180.9	12.1	-172.4	6.28	-160.1	1.29
	50	-153.7	1.32	-169.1	4.78	-171.2	5.54	-159.3	1.79
	100	-154.0	1.12	-168.6	4.46	-170.8	5.31	-161.0	0.757
Cont. est.	-	-155.8	-	-161.4	-	-162.2	-	-162.2	-

Table 6: Final log likelihood and relative errors for each of the nesting structures

The relative error e_{FLL} is calculated as follows

$$e^R_{FLL} = \left| \frac{FLL - \widehat{FLL}_R}{FLL} \right| \cdot 100.$$

As expected, the relative error decreases with the number of draws. For both the logit and for N1 the difference between the true FLL and the value obtained using the MILP is of less than 2% for 20 draws. For N3, the difference between the true FLL and the approximation using the MILP is a bit larger, but is also stable from 20 draws. For N2 the gap between the value obtained with the MILP and the value obtained with continuous estimation decreases from 12% to 5% when increasing the draws from 20 to 50. More advanced results and analysis will be presented as part of the presentation.

Integrating advanced discrete choice models in mixed integer linear optimization

Meritxell Pacheco Paneque*, Shadi Sharif Azadeh,
Michel Bierlaire, Bernard Gendron
*Transport and Mobility Laboratory,
School of Architecture, Civil and Environmental Engineering,
École Polytechnique Fédérale de Lausanne
e-mail: meritxell.pacheco@epfl.ch

Introduction

Mixed integer linear programming (MILP) models are typically used in Operations Research (OR) to represent supply decisions such as the price or the schedule of a public transportation service. The integration of discrete choice models in MILP is appealing to the entities in charge of taking the supply decisions (the *operator*) since it provides a better understanding of the potential demand while designing and configuring their systems. However, the complexity of choice models leads to mathematical formulations that are highly nonlinear and nonconvex in the variables of interest, and therefore difficult to be included in MILP.

In the literature, we find that various authors have placed simplistic assumptions on the choice model, which might be inappropriate in reality, in order to come up with tractable and more efficient formulations. In this research, we provide a disaggregate representation of the choice model that can be embedded into an MILP model. Our approach is general in the sense that any assumption can be made on the probability distribution of the error term from the choice model, and that it can be included in any MILP formulation that requires a demand representation. To illustrate its usage, we consider a case study from the recent literature in which an advanced choice model is developed.

Choice-based optimization

Concerning the choice model, the preference structure of customers is represented with a *utility function*. Briefly, each customer n associates a score with each service/product i, denoted by U_{in}. The most common specification consists of a deterministic term V_{in}, that includes everything that can be modeled by the analyst, and an additive random term ε_{in}, that captures everything that has not been included explicitly in the model. The behavioral assumption is that the customer chooses the service/product with the highest utility.

Operational choice models are obtained by assuming a distribution for ε_{in}. For example, the well-known logit model is obtained by assuming that ε_{in} are independent and identically distributed (across both i and n), with an extreme value distribution. The only assumption to be placed on V_{in} is the linearity on the variables appearing

both in the choice model and the MILP model. This is not required as such for the derivation of the choice model, but important in our context for its integration in a MILP formulation. A typical example of such a variable is the price of a service/product.
In this context, the choice model provides the expected demand of each service/product, which is obtained as the sum of the individual probabilities associated with the service/product, which have typically highly nonlinear and nonconvex expressions. As U_{in} are random variables due to the presence of ε_{in}, we rely on simulation. For each ε_{in} we draw from its distributional assumption. Each draw can be seen as a different and equiprobable behavioral scenario. The number of scenarios determines both the precision of the approximation obtained by simulation and the computational complexity of the resulting model. With this representation, the expected demand is then computed by averaging the choices of the customers over the number of considered scenarios.
Our approach can be employed to model several applications, such as the maximization of passengers' satisfaction in a transportation context or the optimization of assortment in retail. We assume that the MILP model is composed of a linear objective function that relates the supply decisions to an aggregate performance of the system, and a set of linear constraints that identifies the feasible configurations of the variables, as well as the characterization of the choice of each customer for each behavioral scenario. For the sake of illustration, we define a concrete MILP formulation by modeling one operator that sells services to a market and aims at deciding on the price and the capacity of each service (maximum number of customers who can access it) in order to maximize its benefit, which is defined as the difference between the generated revenue and the operating costs.

Application

For the proof-of-concept, we consider the case study of a parking services operator, which is motivated by the availability of a published disaggregate choice model for parking choice. The obtained results exhibit that this formulation is a powerful tool to configure systems based on the heterogeneous behavior of customers. Some properties of the systems, such as the price, can be set specifically for different market segments, which tailors the systems to the users at the same time that the benefit is maximized.
Nevertheless, the disaggregate representation of customers' preferences, together with the linearity of the formulation, implies that the dimension of the resulting problem is high, and therefore solving it is computationally expensive. This is an issue that needs to be addressed because in practice, populations are large and a high number of draws is desirable to be as close as possible to the true value. Decomposition techniques are convenient in this case in order to speed up the solution approach, and represent an alternative to valid inequalities since these techniques can be applied in a general way. More precisely, Lagrangian relaxation for integer programming is currently being considered to define simpler subproblems that can be solved in parallel.

A density-based decision tree for one-class classification

Sarah Itani
University of Mons, Department of Mathematics and Operations Research
e-mail: sarah.itani@umons.ac.be

Fabian Lecron
University of Mons, Department of Engineering Innovation Management
e-mail: fabian.lecron@umons.ac.be

Philippe Fortemps
University of Mons, Department of Engineering Innovation Management
e-mail: philippe.fortemps@umons.ac.be

One-Class Classification (OCC) aims at predicting a single class on the basis of its lonely training representatives, called *positive* or *target* instances [1, 2]. This form of classification is thus opposed to traditional prediction problems with two or several classes. As the availability of data is not necessarily ensured for a good number of medical and industrial applications, OCC presents itself as an interesting alternative methodology of classification. Concretely, the problem related to OCC consists of isolating a class from the rest of the universe; the resulting model allows to predict target patterns and to reject *outlier* ones.

Statistic-based approaches are most commonly used to deal with OCC. In particular, the Kernel Density Estimation (KDE) technique, also called Parzen Windows, approximates the probability density function of a sample [3]. Thresholded at a given level of confidence, the latter function achieves OCC in rejecting any instance located beyond the decision boundary thus established. Though intuitive and simple in nature, KDE may be sensitive to high dimensional data [4].

Our proposal consists of a one-class decision tree; it integrates a multi-dimensional KDE within a more intuitive, readable and structured reasoning scheme. Under a greedy and recursive approach, the underlying learning algorithm splits each training node based on one or several interval(s) of interest. Our method shows favorable performance in comparison with reference methods of the literature.

References

[1] Moya, M. M., Koch, M. W., & Hostetler, L. D. (1993). One-class classifier networks for target recognition applications (No. SAND–93-0084C). Sandia National Labs., Albuquerque, NM (United States).

[2] Pimentel, M. A., Clifton, D. A., Clifton, L., & Tarassenko, L. (2014). A review of novelty detection. Signal Processing, 99, 215-249.

[3] Silverman, B. W. (1986). Density estimation for statistics and data analysis (Vol. 26). CRC press.

[4] Désir, C., Bernard, S., Petitjean, C., & Heutte, L. (2013). One class random forests. Pattern Recognition, 46(12), 3490-3506.

Comparison of active learning classification strategies

Nouara Bellahdid
Mathématique et Recherche Opérationnelle. Université de Mons, Faculté polytechnique
Et, Université des Sciences et de la Technologie Houari Boumedienne
e-mail: nouara.bellahdid@gmail.com

Xavier Siebert
Mathématique et Recherche Opérationnelle. Université de Mons, Faculté polytechnique
e-mail: Xavier.Siebert@umons.ac.be

Moncef Abbas
Université des Sciences et de la Technologie Houari Boumedienne
e-mail: moncef_abbas@yahoo.com

Modern technologies constantly produce huge quantities of data. Because these data are often plain and unlabeled, a particular class of machine learning algorithms is devoted to help in the data annotation process.

In the setting considered in this paper, the algorithm interacts with an oracle (e.g., a domain expert) to label instances from an unlabeled data set. The goal of active learning is to reduce the labeling effort from the oracle while achieving a good classification.

One way to achieve this is to carefully choose which unlabeled instance to provide to the oracle such that it most improves the classifier performance. Active learning therefore consists in finding the most informative and representative sample. Informativeness measures the impact in reducing the generalization error of the model, while representativeness considers how the sample represents the underlying distribution [3, 6].

In early active learning research the approaches were based on informativeness, with methods such as uncertainty sampling, or query by committee. These approaches thus ignore the distribution of the data. To overcome this issue, active learning algorithms that exploit the structure of the data have been proposed. Among them, approaches based on the representativeness criterion have proved quite successful, such as clustering methods [2] and optimal experiment design [1].

Various approaches combining the two criteria have been studied: methods based on the informativeness of uncertainty sampling or query by committee, and a measure of density to discover the representativeness criterion, others methods combine the informativeness with semi-supervised algorithms that provide the representativeness [4, 5].

In this work, we review several active learning classification strategies and illustrate them with simulations to provide a comparative study between these strategies.

References

[1] R. Chattopadhyay, Z. Wang, W. Fan, I. Davidson, S. Panchanathan, and J. Ye. Batch mode active sampling based on marginal probability distri- bution matching. ACM Transactions on Knowledge Discovery from Data (TKDD), 7(3):13, 2013.

[2] D. Ienco, A. Bifet, I. Žliobaitė, and B. Pfahringer. Clustering based active learning for evolving data streams. In International Conference on Discovery Science, pages 79-93. Springer, 2013.

[3] S. jun Huang, R. Jin, and Z. hua Zhou. Active learning by querying infor- mative and representative examples. In J. D. Lafferty, C. K. I. Williams, J. Shawe-Taylor, R. S. Zemel, and A. Culotta, editors, Advances in Neural Information Processing Systems 23, pages 892-900. Curran Associates, Inc., 2010.

[4] M. Li, R. Wang, and K. Tang. Combining semi-supervised and active learn- ing for hyperspectral image classification. In Computational Intelligence and Data Mining (CIDM), 2013 IEEE Symposium on, pages 89-94. IEEE, 2013.

[5] T. Reitmaier, A. Calma, and B. Sick. Transductive active learning-a new semi-supervised learning approach based on iteratively refined generative models to capture structure in data. Information Sciences, 293:275-298, 2015.

[6] B. Settles. Active learning literature survey. University of Wisconsin, Madi- son, 52(55-66):11, 2010.

Risk bounds on statistical learning

Boris Ndjia Njike, Xavier Siebert
Université de Mons, Faculté polytechnique
Département de Mathématique et recherche opérationnelle
e-mail: borisedgar.NDJIANJIKE@umons.ac.be
e-mail: xavier.siebert@umons.ac.be

The aim of this paper is to study theoretical risk bounds when using the Empirical Risk Minimization principle for pattern classification problems. We review some recent developments in statistical learning theory, in particular those involving minimal loss strategies. We conclude with a discussion of the practical implications of these results.

Motivations

The field of machine learning has considerably developed over the last years, from the use of support vector machines (SVM) and its derivatives to the widespread use of deep neural networks. A better theoretical understanding of learning algorithms is important for the understanding of current algorithms as well as for the development of new ones. In this paper we review some recent risk bounds in statistical learning theory, in particular those involving minimal loss strategies.

Formalization of the learning problem

In defined in [4], the problem of pattern classification consists in finding a function $\widehat{f}$ among set of hypothesis $\mathcal{H} = \{f : \mathcal{X} \to \{0,1\}\}$ which minimizes the probability error

$$R(f) = P(Y \neq f(\mathbf{X})) = \int 1_{y \neq f(\mathbf{x})} dP(\mathbf{x}, y) \tag{11}$$

given a i.i.d sample $(\mathbf{x}_1, y_1), \ldots, (\mathbf{x}_n, y_n)$ and an (unknown) probability $P(\mathbf{x}, \mathbf{y})$. Introducing the regression function $\eta(x) = P(Y = 1 \mid X = x)$, the Bayes classifier defined by $f^*(x) = 1_{\eta(x) \geq \frac{1}{2}}$ achieves the minimum risk (11) over all possible measurable functions $f : \mathcal{X} \to \{0,1\}$, as shown in [2].

Risk bounds and strategies in statistical inference

The principle of empirical risk minimization (ERM) consists in replacing (11) by the empirical risk functional

$$R_n(f) = \frac{1}{n} \sum_{i=1}^{n} 1_{\mathbf{y}_i \neq f(\mathbf{x}_i)}$$

and then approximating the function $f_{\mathcal{H}}$ by the function $\widehat{f}_n$, where $f_{\mathcal{H}}$ and $\widehat{f}_n$ are such that: $f_{\mathcal{H}} = \underset{f \in \mathcal{H}}{\operatorname{argmin}} R(f)$ and $\widehat{f}_n = \underset{f \in \mathcal{H}}{\operatorname{argmin}} R_n(f)$. Risk bounds on the error made by replacing the Bayes classifier f^* with $\widehat{f}_n$ are reviewed in [3] .

The ERM principle implicitely reflects a minimax loss strategy [4], with upper bounds relatively close to the lower bounds on the minimax loss. Indeed, using a loss function ℓ such that: $\ell(f, f^*) = R(f) - R(f^*)$, minimax strategy consists in finding the estimator $\widehat{f}$ that minimizes the supremum (over all P) of expected value of $\ell(\widehat{f}, f^*)$. Two cases can be considered, as detailed in [4] for a set of functions $\mathcal{H}$ with VC dimension V :
In the optimistic case (for set of probability $\mathcal{P}$ for which $R(f^*) = 0,\ \forall P$): for $n > V$,

$$\frac{V+1}{2e(n+1)} \leq \inf_{\widehat{f}} \sup_{P} \mathbb{E}_P(\ell(\widehat{f}, f^*)) \leq \sup_{P} \mathbb{E}_P(\ell(\widehat{f}_n, f^*)) \leq \frac{4}{n} \ln \left(\frac{2en}{V}\right)^V + \frac{16}{n}.$$

In the pessimistic case (for a set of probability $\mathcal{P}$, $\exists P$ such that $R(f^*) \neq 0$), for $n > 2V$:

$$\frac{V}{n}(1 - erf(1)) \leq \inf_{\widehat{f}} \sup_{P} \mathbb{E}_P(\ell(\widehat{f}, f^*)) \leq \sup_{P} \mathbb{E}_P(\ell(\widehat{f}_n, f^*)) \leq 4\sqrt{\frac{V\left(\ln \frac{2n}{V} + 1\right) + 24}{n}} \tag{12}$$

Using geometric and combinatorial quantities related to the class $\mathcal{H}$, it is possible to refine (12) in several ways. First, as in [2], if $n > 2V$ there exists an absolute positive constant c such that:

$$\sqrt{\frac{V}{n}}(1 - erf(1)) \leq \inf_{\widehat{f}} \sup_{P} \mathbb{E}_P(\ell(\widehat{f}, f^*)) \leq \sup_{P} \mathbb{E}_P(\ell(\widehat{f}_n, f^*)) \leq c\sqrt{\frac{V}{n}}$$

Second, as in [1], instead of taking P in some arbitrary set $\mathcal{P}$, one can introduce a parameter $h \in [0, 1]$, such that $P \in \mathcal{P}(h)$ and $\mathcal{P}(h)$ is the set: $\{P \in \mathcal{P},\ |\ 2\eta(x) - 1\ | \geq h \text{ for all } x \in \mathcal{X}\}$. Then, if we assume that $\mathcal{H}$ has a finite VC-dimension $V \geq 2$, for some absolute positive constant k, if $n \geq V$, one has

$$\inf_{\widehat{f}} \sup_{P \in \mathcal{P}(h)} \mathbb{E}_P(\ell(\widehat{f}, f^*)) \geq k \min\left(\frac{V-1}{nh}, \sqrt{\frac{V-1}{n}}\right).$$

We will show that the latter bounds are in fact a particular case of [5] and discuss the practical implications of these results.

References

[1] P. Massart, E.Nédélec,The Annals of Statistics, Risk bounds for statistical learning, 2326-2366, 2006.

[2] G. Lugosi. Pattern classifcation and learning theory. In Principles of nonparametric learning, pages 1-56. Springer, 2002.

[3] O. Bousquet, S. Boucheron, and G. Lugosi. Introduction to statistical learning theory. In Advanced lectures on machine learning, pages 169-207. Springer, 2004.

[4] V. N. Vapnik. Statistical Learning Theory. John Wiley & Sons, 1998.

[5] A. B. Tsybakov. Introduction to nonparametric estimation. Springer Series in Statistics.Springer, New York, 2009.

Author index