# ORBEL 36

## Book of Abstracts

Gent, Belgium, 12-13 September 2022

# Contents

# Welcome to ORBEL 36

We are delighted to welcome you to Gent for the 36$^{\text{th}}$ edition of the ORBEL conference. It has been a long two and a half years since the ORBEL conference last took place in person. Therefore we are especially pleased to be meeting with you all at KU Leuven's Technology Campus.

Thanks to your submissions, ORBEL 36 offers an exciting and diverse program. In total 66 talks are scheduled, divided across 23 parallel sessions. We hope you will also enjoy the two plenary sessions given by our distinguished invited speakers: Prof. Michael Schneider and Prof. Antonio Martinez-Sykora.

We wish to thank the members of the Scientific Committee, the Organizing Committee and the local organizers for all their time and assistance during the preparation of the event. Thanks also to our three sponsors: OMP, GUROBI and N-SIDE.

We hope you have an enjoyable and fruitful conference!

Tony WAUTERS
Conference chair

## Sponsors

# Organization

## Nominations for the board of administrators

El-Houssaine Aghezzaf
Yasemin Arda
Jeroen Beliën
Dries Benoit
Wouter Blondeel
Kris Braekers
Philippe Chevalier
Kristof Coussement
Yves Crama
Bernard De Baets
Patrick De Causmaecker
Yves De Smet
Bernard Fortz
Dries Goossens
Gerrit Janssens
Pierre Kunsch
Martine Labbé
Roel Leus
Hadrien Mélot
Dimitri Papadimitriou
Célia Paquay
Thierry Pironet
Annick Sartenaer
Pierre Schauss
Michaël Schyns
Kenneth Sörensen
Frits Spieksma
Jean-Sébastien Tancrez
Filip Van Utterbeeck
Greet Vanden Berghe
Pieter Vansteenwegen
Wouter Verbeke
Tony Wauters
Sabine Wittevrongel

## Local organizers

Luke Connolly
Jeroen Gardeyn
Tony Wauters

# Program overview

## Monday, September 12

| | | |
|---|---|---|
| 8:30-9:30 | Registration & welcome coffee | Cafetaria |
| 9:30-9:45 | Opening session | E036 (Groot Auditorium) |
| 9:30-10:45 | Plenary talk | E036 (Groot Auditorium) |
| | *Michael Schneider* | |
| 10:45-11:15 | Coffee break | Cafetaria |
| 11:15-12:30 | Parallel sessions | |
| | *MA-1: OR4Logistics -Health Care* | D002 |
| | *MA-2: Scheduling* | E030 |
| | *MA-3: Logistics* | E031 |
| | *MA-4: Data Science 1* | E033 |
| 12:30-14:00 | Lunch break | Cafetaria |
| | *ORBEL board meeting* | E033 |
| 14:00-15:15 | Parallel sessions | |
| | *MB-1: OR4Logistics - Warehousing* | D002 |
| | *MB-2: Cutting & Packing 1* | E030 |
| | *MB-3: Routing* | E031 |
| | *MB-4: Data Science 2* | E033 |
| 15:15-15:45 | Coffee break | Cafetaria |
| 15:45-17:00 | Parallel sessions | |
| | *MC-1: OR4Logistics - Production* | D002 |
| | *MC-2: Sports* | E030 |
| | *MC-3: Graphs* | E031 |
| | *MC-4: Data Science 3* | E033 |
| 19:00-23:00 | Conference dinner | Brasserie Pakhuis |

## Tuesday, September 13

| | | |
|---|---|---|
| 9:00-9:30 | Welcome coffee | Cafetaria |
| 9:30-10:45 | Parallel sessions | |
| | *TA-1: OR4Logistics - Passenger Transport* | D002 |
| | *TA-2: Timetabling* | E030 |
| | *TA-3: Tools* | E031 |
| | *TA-4: OR Methods 1* | E033 |
| 10:45-11:15 | Coffee break | Cafetaria |
| 11:15-12:30 | Parallel sessions | |
| | *TB-1: OR4Logistics - Freight Transport* | D002 |
| | *TB-2: Cutting & Packing 2* | E030 |
| | *TB-3: Production* | E031 |
| | *TB-4: OR Methods 2* | E033 |
| 12:30-14:00 | Lunch break | Cafetaria |
| 14:00-15:15 | Parallel sessions | |
| | *TC-1: OR4Logistics - Routing* | D002 |
| | *TC-2: Simulation Optimization* | E030 |
| | *TC-4: OR Methods 3* | E033 |
| 15:15-15:45 | Coffee break | Cafetaria |
| 15:45-16:45 | Plenary talk | E036 (Groot Auditorium) |
| | *Antonio Martinez-Sykora* | |
| 16:45-17:30 | Farewell drink | Cafetaria |

# Detailed program

## Monday 11:15-12:30

### MA-1: OR4Logistics - Health Care

**Chair:** A. Delaet



### MA-2: Scheduling

**Chair:** S. Michelini



### MA-3: Logistics

**Chair:** M. Wens

### MA-4: Data Science 1

**Chair:** C. Bockel-Rickermann



## Monday 14:00-15:15

### MB-1: OR4Logistics - Warehousing

**Chair:** S. Teck



### MB-2: Cutting & Packing 1

**Chair:** H. Çalık



### MB-3: Routing

**Chair:** C. Defryn

## MB-4: Data Science 2

**Chair:** J. Van Belle



# Monday 15:45-17:00

## MC-1: OR4Logistics - Production

**Chair:** A. Van Meir



## MC-2: Sports

**Chair:** T. Demeulemeester



## MC-3: Graphs

**Chair:** J. Renders

### MC-4: Data Science 3

**Chair:** S. Weekx



## Tuesday 9:30-10:45

### TA-1: OR4Logistics - Passenger Transport

**Chair:** P. Vansteenwegen



### TA-2: Timetabling

**Chair:** N. Schmid



### TA-3: Tools

**Chair:** M. Van Lancker

### TA-4: OR Methods 1

**Chair:** T. De Munck



## Tuesday 11:15-12:30

### TB-1: OR4Logistics - Freight Transport

**Chair:** I. Van Hoeck



### TB-2: Cutting & Packing 2

**Chair:** M. Thanos



### TB-3: Production

**Chair:** S. Tran



### TB-4: OR Methods 2

**Chair:** M. Baratto

## Tuesday 14:00-15:15

### TC-1: OR4Logistics - Routing

**Chair:** N. Morandi



### TC-2: Simulation Optimization

**Chair:** T. Harbi



### TC-4: OR Methods 3

**Chair:** S. Alaei Jordehi

# Plenary talks

## Prof. Michael Schneider

*RWTH Aachen University*

**Monday, 9:45**

Michael Schneider studied business administration and computer science at the University of Mannheim and received his doctoral degree from the Technical University of Kaiserslautern in 2012. From 2013 to 2016, he worked as assistant professor in Logistics Planning and Information Systems at TU Darmstadt. Since June 2016, Michael is a full professor at RWTH Aachen University and holder of the Deutsche Post Chair – Optimization of Distribution Networks (DPO). Michael's research concentrates on the development of heuristic and sometimes exact methods for logistics optimization problems, especially in the fields of transportation, location planning, warehouse management, and vertical farming. He has published several articles on these topics in renowned international journals such as Operations Research, Transportation Science, INFORMS Journal on Computing, and European Journal of Operational Research. Michael is a board member of VeRoLog (EURO working group on Vehicle Routing and Logistics Optimization) and founder of the Global Challenges Lab at RTWH Aachen University.

### Location Routing Problems: State-of-the-Art Heuristics and Recent Developments

The capacitated LRP (CLRP) jointly takes decisions on the location of capacitated depots and the routing of capacitated vehicles to serve a set of customers with known demands from the opened depots. A large number of heuristic approaches has been proposed for the CLRP, often decomposing the problem into a location stage to determine a promising depot configuration and a routing stage, where a vehicle routing problem is solved to assess the quality of the previously determined depot configuration.

In this talk, we give an overview of the current state-of-the-art heuristics for the CLRP, and we try to shed some light on the question which components are most relevant for the success of these heuristics. In addition, we discuss recent developments with regard to the problem variants studied and the solution methods proposed in the literature. Finally, we outline interesting topics for future research.

## Prof. Antonio Martinez-Sykora

*University of Southampton*


**Tuesday, 15:45**

Antonio Martinez-Sykora, current director of CORMSIS (Centre of Operations Research, Management Science and Information Systems), is an Associate Professor in Business Analytics. He finished his PhD in June 2013 at the University of Valencia (Spain) and joined the Southampton Business School in September 2013 as a research assistant. In 2007, he received his BSc in Mathematics and BSc in Statistics and in 2009 he received the M.A.S in Operations Research. His research focuses in a wide range of combinatorial optimization problems, especially on cutting and packing problems and logistic planning problems such as vehicle routing, scheduling, cutting and packing and revenue management. He is a coordinator of the Euro Working Group on Cutting and Packing (ESICUP) and he has participated in various externally funded projects.

### Irregular Packing (Cutting) Problems: Models, Algorithms and Challenges

Cutting and Packing problems with irregular shapes or objects (2D or 3D) is an area that has been developed at a slower pace than other packing areas in the last 20 years, such as packing boxes within a container (container loading problem). Problems with irregular shapes generally have an extra layer of difficulty on top of the combinatorial optimisation problem, derived on how the shapes are modelled and, more generally, how the actual layout/solution is represented. The computational challenges and the complex geometry of the shapes/objects have split researchers in three main areas to address these problems: (1) polygonal/triangular mesh approximation, (2) pixel/raster/voxel-based approximations or (3) phi-objects/parametric functions.

During this talk we will discuss the relevance and the importance of solving these problems efficiently and the state of the art of 2D irregular (strip) packing problems using the polygon representation. Then, some new advances on 3D irregular (strip) packing problems using the voxel-based solution representation. Also we will discuss some future directions and challenges. I hope, the talk will enthuse researchers to engage with irregular packing problems and its challenges.

# Instructions for presenters

We kindly request that the presenter of each session's final talk to chair their corresponding session and transfer all presentations onto a single laptop before their session begins. All presenters should locate and present themselves to their session chair before the session commences.

- Plenary talks: 45 minutes + 15 minutes for Q&A
- Parallel sessions: 20 minutes + 5 minutes for Q&A

# Abstracts

# An integrated nurse rerostering and routing problem in hospital-at-home

A. Maharani
QuantOM - HEC Liège
e-mail: anisha.maharani@uliege.be

Y. Arda
QuantOM - HEC Liège
e-mail: yasemin.arda@uliege.be

V. François
QuantOM - HEC Liège
e-mail: veronique.francois@uliege.be

Hospital-at-home provides short-term acute care at the patient's home for conditions that would otherwise necessitate in-patient hospitalization. Several benefits of this service are the following: increased capacity of the institutional health care system, improved patient quality of life, and reduced length of stay [4]. The problem addressed in this presentation focuses on operational scheduling decisions for hospital-at-home systems. When having an acute illness, a patient needs care for a limited period. As a result, the patient mix fluctuates. Moreover, the availability of nurses also varies over time. Therefore, a precise scheduling plan is required to better match available resources with patient needs.

Given the baseline nurse roster and patient requirements, several operational decisions are taken simultaneously over the planning horizon: (i) select patients to be admitted and their admission dates, (ii) assign nurses to the admitted patients and schedule the care visits, and (iii) decide whether and how the nurse roster should be updated. The studied problem is a combination of task scheduling, nurse routing, and nurse rerostering problems. Generally, the underlying subproblems are solved independently in the literature, as this approach is less complex and thus more computationally practical. However, this sequential decision-making process may lead to inefficient or even infeasible solutions, since the subproblems are strongly intertwined.

The objective of this work is first to maximize the number of patients treated at home, and second to minimize the total working duration of the nurses. A variety of complex real-world characteristics are considered, yielding a rich integrated problem. In particular, to schedule the patient visits of the routing subproblem, we consider the temporal dependencies between the treatments associated with each patient, as well as the periodicity of those treatments. Such routing problems are classified as the home health care routing and scheduling problem in the literature [1]. Another important aspect of the problem we studied are the rerostering decisions [2]. When rerostering is necessary, several classical rostering constraints must be respected, such as the minimum rest time and the maximum number of shift assignments.

The resulting mixed-integer program is modelled using a stepping horizon approach [3] in order to properly address the connections between consecutive scheduling periods. The model was validated using toy instances. The presentation will discuss several challenges encountered due to the integration of scheduling, routing, and rerostering decisions. Our next step is to develop solution methods to tackle realistic instances.

**Keywords**: nurse rerostering, home health care routing and scheduling, integrated combinatorial optimization, hospital-at-home

## References

[1] Fikar, C., & Hirsch, P. (2017). Home health care routing and scheduling: A review. *Computers & Operations Research 77*, 86-95.

[2] Maenhout, B., & Vanhoucke, M. (2018). A perturbation matheuristic for the integrated personnel shift and task re-scheduling problem. *European Journal of Operational Research 269* or(3), 806-823.

[3] Salassa, F., & Vanden Berghe, G. (2012). A stepping horizon view on nurse rostering. In *Proceedings of the 9th international conference on the practice and theory of automated timetabling.* (pp. 161-173).

[4] Shepperd, S., & Iliffe, S. (2005). Hospital at home versus in-patient hospital care. *Cochrane database of systematic reviews*(3).

# An inventory-routing problem for cooperative hospital supply chain operations

S. Mertens
Hasselt University, Research Group Logistics
e-mail: silia.mertens@uhasselt.be

L. Vanbrabant
Hasselt University, Research Group Logistics, Research Foundation Flanders (FWO)
e-mail: lien.vanbrabant@uhasselt.be

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

A. Caris
Hasselt University, Research Group Logistics
e-mail: an.caris@uhasselt.be

The healthcare sector is facing increasing pressure to cut costs. This is due to the ageing of the population, which raises the need for efficient healthcare systems [1]. Moreover, there is a severe shortage in nursing staff and changing policies lead to tighter health budgets [2]. By making improvements in logistic efficiency, progress can be made in the healthcare sector to reduce costs while maintaining or even improve the service quality.

The cost of logistics (e.g. purchasing, inventory management, distribution) increases up to 45 % of the total operational budget of a hospital [3]. A traditional hospital supply chain is designed as a multi-echelon inventory system, consisting of external suppliers (e.g. pharmaceutical companies, manufacturers, distributors), a central warehouse within the hospital, point-of-use locations (i.e. care units such as nursing units, operating room and emergency departments) and the patients as final users [5]. Moving from a traditional hospital supply chain to a cooperative hospital supply chain can result in efficiency improvements in the logistic processes. Supply chain risk (e.g. product obsolescence) can be shared between hospitals and costs can be reduced (e.g. purchasing and inventory costs) [3]. However, one of the struggles can be the need for an integrated IT system, which involves large investment costs [4]. In a cooperative supply chain, a single central warehouse (care hub) operates for multiple hospitals and replaces the central warehouse of each individual hospital.

In order to fully benefit from moving to a cooperative supply chain where inventory is pooled, it is desirable to integrate decisions on different levels of control (operational, tactical and strategic). Therefore, the aim of this study is to solve an inventory-routing problem for a cooperative supply chain (with a care hub) which is compared with the traditional supply chain to examine whether efficiency improvements can be achieved when making use of a care hub. The problem is applied to a case study in collaboration with a Belgian University Hospital. Real-life features are included to make the case realistic. Multiple products are involved. Uncertainty in demand is included and due to this uncertainty, emergency deliveries are possible. In the hub, inventory is pooled, which reduces the inventory levels and time windows are added.

A multi-period rolling horizon approach is proposed. The initial inventory is known at the beginning of the current period and order quantities are defined. In each planning period (daily) total expected costs are minimized by a mixed integer programming model, consisting of inventory holding costs, shortage costs, routing costs, ordering costs and emergency delivery costs. When the actual demand is observed inventory levels are updated accordingly and possible shortages are corrected using emergency deliveries, which can come either from the hub or the supplier, depending on the amount of inventory that is still left at the hub. Total actual costs are calculated based on the actual demand. This procedure is repeated until the end of the planning horizon (1 year).

First insights show that a central care hub results in reduced costs, while maintaining the same service level. Therefore, it is advantageous to move to a cooperative hospital supply chain. Furthermore, an experimental design is shown, varying different parameters to gain more insight into the problem.

## References

[1] A. Azzi, A. Persona, F. Sgarbossa, and M. Bonin. Drug inventory management and distribution: outsourcing logistics to third-party providers. *Strategic Outsourcing: an International Journal; Bingley*, 6(1):48–64, 2013.

[2] S. Landry and R. Philippe. How Logistics Can Service Healthcare. *Supply Chain Forum: An International Journal*, 5(2):24–30, Jan. 2004.

[3] K. Moons, G. Waeyenbergh, and L. Pintelon. Measuring the logistics performance of internal hospital supply chains – A literature study. *Omega*, 82:205–217, 2018.

[4] R. Pinna, P. P. Carrus, and F. Marras. Emerging Trends in Healthcare Supply Chain Management — An Italian Experience | IntechOpen, 2015.

[5] J. Volland, A. Fügener, J. Schoenfelder, and J. O. Brunner. Material logistics in hospitals: A literature review. *Omega*, 69:82–101, June 2017.

# Integrated decision-making in home health care : review, model and first results

A. Delaet
Hasselt University and Vrije Universiteit Brussel
e-mail : arne.delaet@uhasselt.be

K. Braekers, K. Ramaekers
Hasselt University, Research Group Logistics

Y. Molenbruch
Vrije Universiteit Brussel,
Mobility, Logistics and Automotive Technology Research Centre

Home health care (HHC) is an important component of the health care industry and may be defined as care workers visiting patients following predefined schedules in order to provide medical services in their home. Maintaining a sustainable and effective health care system is a major challenge as a result of two trends : limited resources (budget restrictions and staff shortages) and a rise in demand (population ageing, pandemic outbreaks...). In response to these trends and increasing competitive pressures, HHC providers must discover new ways to decrease costs and enhance productivity by optimizing the use of resources. Efficiently organizing HHC services requires making a wide range of complex decisions on multiple levels, ranging from day-to-day operational decisions (e.g. developing routes and visit schedules for care workers) to tactical and strategic decisions having an impact at a longer time horizon (e.g. determining staffing levels). In addition, some complicating problem-specific characteristics need to be considered, such as matching care workers' medical skills with patients' requirements, continuity of care constraints and work-related constraints (e.g. maximum number of shifts and weekends care workers are allowed to work). For all these reasons, it is clear that applying operations research techniques in HHC is a promising research field.

First, this talk will discuss the findings of a literature review on OR models applied in HHC. The OR-literature concerning HHC is dominated by papers proposing models and solution methods for individual operational decision-making problems. Researchers stress the lack of integrated multi-level planning studies spanning a realistic planning-horizon. Integrated studies are highly relevant because solving independent sub problems separately results in suboptimal decision-making. In this context, a key opportunity for improvement is the integration of decisions at different decision-making levels.

Second, the specific problem setting on which we focus will be defined and its contributions will be stressed. In particular, the goal of this research is to develop innovative models and solution algorithms that enable making better decisions on staffing levels by solving the following decisions in an integrated manner : staff dimensioning (determining the number of care workers with the required set of skills necessary to meet demand), rostering (allocating care workers to shifts), clustering (decomposing a problem into smaller sub problems by grouping patients) and scheduling and routing. A first mixed integer linear programming model in this direction will be presented. We focus on improving staffing decisions, as this is highly relevant due to persistent staff shortages and the fact that staffing levels determine the performance of the health delivery system to a large extent. This project contributes to the current state-of-the-art by integrating tactical and lower decision-making levels and by considering the clustering decision, which is both practically and academically relevant.

Finally, this talk will focus on the solution algorithm that is currently being developed to solve the model under study and the first results obtained by this algorithm. An important academic contribution of this algorithm, is that it tackles the HHC planning problem by applying an integrated approach for the rostering, clustering and routing decisions. In addition, staff is scheduled for a realistic time horizon, which allows for enforcing continuity of care and working-time constraints. More specifically, the solution algorithm first finds an initial feasible solution by using a combination of a tailored k-means heuristic to cluster patients and a binary integer linear programming model to roster care workers. This binary integer linear programming model is solved using a commercial solver. In addition, routings are solved in an exact manner in the first phase. In a second phase of the solution algorithm, the initial solution is improved by executing iterations of a tailored large neighbourhood search heuristic.

# Order picking problem: Exact and heuristic algorithms for the Generalized Travelling Salesman Problem with geographical overlap between c lusters

Farzaneh Rajabighamchi; Stan van Hoesel; Christof Defryn
Maastricht University, School of Business and Economics
Department of Quantitative Economics
e-mail: {f.rajabighamchi; s.vanhoesel; c.defryn}@maastrichtuniversity.nl

Within modern warehouses, identical items are often located in multiple locations throughout the warehouse to improve the efficiency of collecting customers' orders. Since only one of these locations needs to be visited to collect a specific item, the problem is modeled as a Generalized Travelling Salesman Problem (GTSP). GTSP is an extension of the traveling salesman problem (TSP) where the set of nodes is partitioned into clusters, and the salesman must visit exactly one node per cluster.

In this research, we apply the definition of the GTSP to an order picker routing problem with multiple locations per product. As such, each product represents a cluster and its corresponding nodes are the locations at which the products can be retrieved. The objective is to find the shortest tour in the clustered graph starting and ending at the depot such that exactly one node of each cluster is visited.

As all products are scattered throughout the warehouse, the product clusters are not separated geographically, which is a more general form of GTSP, and there exists only paper in the literature taking this aspect of the problem into account and assumes such a definition of clustering. To the best of our knowledge, the only research having considered both were proposed by [1]. The researchers highlight the importance of considering overlapping clusters due to their applications in modern warehouses with multiple picking locations for the same item. They propose a Conditional Markov Chain Search, which, based on a pool of heuristic components, automatically generates a meta-heuristic specifically for warehouse order picking. While this paper considers the topology of a warehouse to calculate the distance between items, it does not address various warehouse configurations. More specifically, their instance generator and computational results are based on a fixed grid and do not account for various warehouse sizes with numerous configurations of aisles and cross-aisles.

We propose an ILP formulation of the problem and present a heuristic solution method for obtaining high-quality pick tours. The heuristic is based on a variable neighbourhood search metaheuristic, embedded in a guided local search framework. Furthermore, in order to to illustrate the efficiency of the exact model and

the solution methods, we implemented them on different size benchmark problems from [2] and [3] consisting of 9 different scenarios: three different number of aisles (5,15,60), three different number of cross aisles (3,6,11) and three different number of products in the order (15,60,240).
All algorithms presented in this paper have been implemented in Java, and the ILP formulations have been solved with IBM CPLEX 22.1.0 with default parameters. The time limit for the exact algorithms has been set to 3600 seconds. Testing has been carried out on a Macbook Air with an Apple Silicon M1 chip and 16GB of RAM.
The computational results demonstrate that the proposed algorithm provide better solution quality compared to existing methods for solving GTSP in a shorter amount of time including on larger GTSP instances.
with the increase in cluster sizes and accordingly the instance sizes, our proposed MILP is not able to give us a feasible solution within the time limit of one hour. However, our meta-heuristic algorithm is able to solve the problem and give us a good solution in a very short time (for our biggest instance with 4800 nodes, the run time is 52.5 seconds which is less than one minute and it is incredibly fast). For the cluster size 1,2 and 5, our proposed MILP can provide optimal or feasible solutions but for the larger instances, the model is not able to give any solutions within 60 minutes. The other notable point in these tables is the comparison between our proposed MILP and the existing MILP in the literature. In average, the run time of our proposed MILP is 20% less than the other MILP in the literature, and the quality of the solutions generated by our MILP is in average 15% better than the existing MILP.
Comparing the solutions of GLS with VNS, we can see that implementing GLS and penalising the features of the solutions has an improvement of 15% in average and this percentage is higher for bigger instances compared to the small instances with 2 products in each cluster which this average improvement is 4%.
for the smaller cluster size (2), the Furthermore, in the smaller instances (cluster size 2 and 5) the run time of GLS is in average 80% less than the run time of VNS, and this value drops to 50% for larger instances with 10 and 20 products in each cluster. These numbers show the efficiency of our proposed GLS algorithm and its significant improvements both in run time and the quality of the solutions.

# 1 References

Nalivajevs, O., & Karapetyan, D. (2019, September). Conditional Markov Chain Search for the Generalised Travelling Salesman Problem for Warehouse Order Picking. In 2019 11th Computer Science and Electronic Engineering (CEEC) (pp. 75-78). IEEE.

Pansart, L., Catusse, N., & Cambazard, H. (2017). Exact algorithms for the order picking problem. arXiv preprint arXiv:1703.00699.

Theys, C., Bräysy, O., Dullaert, W., & Raa, B. (2007). Towards a metaheuristic for routing order pickers in a warehouse. Evolutionary methods for design, optimization and control, 385-390.

# A framework for formulating competence-aware scheduling models

C. Miguel Palao, S. Hoedt, J. Cottyn, E.H. Aghezzaf
Ghent University, Department of Industrial Systems Engineering and Production Design
Flanders Make@UGent, Industrial Systems Engineering
e-mail: `carlos.miguelpalao;steven.hoedt@ugent.be`
`johannes.cottyn;elhoussaine.aghezzaf@ugent.be`

Lean manufacturing stresses the need for multi-skilled operators which provide flexibility by being capable of working in multiple stations [1]. A multi-skilled workforce is deemed to be "robust", able to cope with unexpected occurrences such as worker unavailability and operator competence variations driven by practice or lack of it [2]. Attaining a robust workforce through multiskilling is not an easy task. Studies have shown that a fully cross-trained floor can be beneficial but comes at a cost in terms of creation, maintenance and product quality [3]. Finding the correct balance of skills and cross-training is thus of capital importance. Yet, workforce development does not happen in a vacuum and needs to be inserted correctly into schedules.

The inclusion of competence awareness into scheduling models brings further accuracy in production and workforce development planning [4]. Taking into consideration different skill levels among operators allows for different productivities in the formulation of the model and results in more precise solutions. Production planning improvement and workforce development planning can be seen as two extremes in a spectrum of competence-aware scheduling models. The former features competences solely for *efficiency measuring* reasons whereas the latter intends to not just track competence evolution but also for *competence buildup* to happen in the workforce through work assignments and training scheduling.

Scheduling models can be placed along this spectrum depending on how their formulation seeks to fulfil one of these two goals. This can be done in varying degrees, as many models and problems raised in literature seek to achieve both workforce development and realistic planning [4] [5]. In this regard we have observed a series of elements that are present in many of formulations and that we have coalesced into the following framework.

**Problem formulation**

| | | |
|---|---|---|
| Cost formulation | Balance mechanism | Competence evolution |
| Process modelling | Workforce development mechanism | Solving algorithm |

Table 1: Elements of the framework for competence-aware scheduling

Competence-aware scheduling problems can be formulated and solved with the six elements featured in table 1. Ubiquitous to linear programming are process modelling and cost formulation, which serve as the basis for the constraints and objective functions the model would need to properly reflect the conditions of the manufacturing process studied. Constraints limiting the number of operators in a station and an objective function that minimises wage expenditures are examples of what we call process modelling and cost formulation respectively.

Furthermore, two *mechanisms* need to be placed as part of this formulation, dealing with workforce development and balance. The former is responsible for achieving robustness in the labour force by fostering or forcing solutions that increase the competence levels of operators over time, both in average value and distribution in order to ensure having highly competent operators that can work in several stations. The latter allows the user to modulate the degree with whom competence development is sought and weighed against short-term manufacturing goals. The two mechanisms depend largely on the presence of some form of competence evaluation that is able to track increases and decreases in the competence pool of operators to be adequately formulated as part of the model. Finally, a solving algorithm (exact, approximate or heuristic) obtains solutions for the proposed problem.

We call the two elements in the second column of table 1 *mechanisms* since these are the parts of the model that make sure operator and job assignments are conducted in such a way that manufacturing goals and workforce development happen in the exact measure as to fulfil the user's intentions. The absence of a workforce development mechanism would confine a model to the *efficiency intended* side of the spectrum, making use of competence evolution but missing the opportunity to foster a robust workforces. Without a balance mechanism the model would be stuck in a specific configuration, limiting the options management has to adapt planning to manufacturing requirements. Both mechanisms can be formulated in a multitude of ways, either as part of the set of constraints, objective functions or both, which we will discuss in future works.

## References

[1] P.Korytkowski, "Competences-based performance model of multi-skilled workers with learning and forgetting," *Expert Systems with Applications*, vol.77, pp.226–235, 2017.

[2] J.Peltokorpi and M.Y. Jaber, "Interference-adjusted power learning curve model with forgetting," *International Journal of Industrial Ergonomics*, vol.88, p.103257, 2022.

[3] C.H. Glock *et al.*, "Applications of learning curves in production and operations management: A systematic literature review," *Computers & Industrial Engineering*, vol.131, pp.422–441, 2019.

[4] Z.Wang *et al.*, "Modeling worker competence to advance precast production scheduling optimization," *Journal of Construction Engineering and Management*, vol.144, no.11, p.04018098, 2018.

[5] E.-A. Attia *et al.*, "Considering skills evolutions in multi-skilled workforce allocation with flexible working hours," *International Journal of Production Research*, vol.52, no.15, pp.4548–4573, 2014.

# A Global Constraint for Modeling Time in Scheduling Problems for Local Search

Stefano Michelini
CETIC asbl, Combinatorial Algorithmics
e-mail: stefano.michelini@cetic.be

Renaud De Landtsheer
CETIC asbl, Combinatorial Algorithmics
e-mail: renaud.delandtsheer@cetic.be

The Iterative Flattening and Relaxing algorithm (IFlatIRelax) [1] is a heuristic aimed at solving the cumulative scheduling problem, which is a scheduling problem where each machine has a certain capacity, and each task may require more than one capacity unit. The heuristic is an improvement of the Iterative Flattening algorithm by [3], and it works by manipulating the directed acyclic graph (DAG)-based model of the problem, where nodes are associated with tasks and arcs with task precedence relations. More specifically, the algorithm explores the solution space by repeatedly adding feasible arcs between tasks in a minimal conflict set associated with a capacity constraint violation in order to repair feasibility (the *flattening* step), or removing arcs belonging to the *critical path*, i.e., the longest path in the DAG determining the value of the makespan, the objective function of the problem (the *relaxing* step).

In their introduction to constraint-based local search (CBLS) [2], the authors present different ways to tackle scheduling problems within their own COMET framework, including IFlatIRelax. However, in order to do so, the search procedure evaluates the makespan through an ad-hoc formula, breaking a design principle according to which the local search neighbourhoods can only evaluate moves by querying the objective function. Therefore, the search procedure and the model cannot be adapted to solve variants of this problem. An relevant variant is the *flexible job-shop problem* where one must also select the appropriate resource for each operation.

In this work in progress, we address this issue by developing a global constraint whose input is the structure of the DAG, as manipulated by the search procedure and whose output is the makespan. The makespan is maintained incrementally by using the algorithms in [4], and the ad-hoc formula used in [2]to achieve $O(1)$-time complexity whenever possible. This further allows us to formulate the IFlatIRelax as a *Destroy&Repair* search procedure using standard neighbourhood combinators from the OSCAR.cbls framework and generic *flatten* and *relax* neighborhoods. The implementation is included in open source OSCAR.cbls framework [5].

**Acknowledgement:** This research is supported by the SLS project (Logistics in Wallonia - nr 8457)

## References

[1] Michel, L. and Van Hentenryck, P., *Iterative Relaxations for Iterative Flattening in Cumulative Scheduling*, ICAPS. Vol. 4, 2004.

[2] Van Hentenryck, P., and Michel, L., *Constraint-based Local Search*, MIT Press, 2005.

[3] Cesta, A., Oddi, A., and Smith, S. F., *Iterative flattening: A scalable method for solving multi-capacity scheduling problems*, AAAI/IAAI, 742–747, 2000.

[4] Katriel, I., Michel, L. and Van Hentenryck, P., *Maintaining longest paths incrementally*, Constraints 10.2, 159-183, 2005.

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

# A multi-period collaborative model for analyzing the impact of a city hub on city logistics

Titi Iswari
Hasselt University, Research Group Logistics
e-mail: titi.iswari@uhasselt.be

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

An Caris
Hasselt University, Research Group Logistics
e-mail: an.caris@uhasselt.be

City logistics aims to optimally plan, manage, and control the vehicle movements within a logistical network in an urban area, considering integration and coordination among involved stakeholders. Since collaboration is a crucial element in city logistics, innovative solutions for city logistics related to this element have been introduced over the years. The widely-used solution is by considering the use of two or multi-echelon systems using one or more intermediate consolidation points, such as urban consolidation centers, city hubs, micro depots, etc. The additional cost due to the double handling process in the implementation of a city hub makes it important to discuss the broader implications and potential cost reductions in other parts of the stakeholders involved. In practice, the main focus of this intermediate consolidation point is on consolidating flows to make the transportation or distribution activities more efficient and more environmentally friendly. Another opportunity in such a system is to use these consolidation points as temporary storage facilities in a B2B setting, such that retailers can save costly storage space in the urban area. However, the study of the inventory aspect in city logistics is relatively unexplored. Therefore, this project aims to study the effect of using a UCC or city hub in a city logistics system in terms of inventory and routing aspects.

Two scenarios are considered to observe the impact of a city hub in a city logistics setting. Both scenarios use a network with multiple retailers in which each supplier is responsible for delivering a different type of product to the retailers over a finite and discrete-time horizon. Each retailer has a deterministic demand for each product type, a daily time window for deliveries, and a maximum inventory level. The objective function is to minimize the total cost, which is the sum of the inventory and transportation cost. The first scenario is the base scenario in which there is no city hub. In this scenario, each retailer uses a replenishment policy to define the order size. Then, each supplier determines its optimal delivery plan and delivers its product to the retailers directly. In the second scenario, a city hub acts as an intermediate consolidation point. The city hub has a role in receiving the products from the suppliers and defining the optimal delivery routes for the retailers in the city. Therefore, a two-echelon network is formed in the second scenario. Since one of the critical elements in city logistics is considering the ecological impact on urban transport, several logistics providers have innovated to use environmentally-friendly vehicles, e.g. cargo bikes, electric vans, etc. Hence, in the second scenario, a heterogeneous vehicle model with cargo bikes and vans is used for the delivery part from the city hub to the retailers, where the cargo bikes can also perform multiple trips if necessary.

In our current approach, the inventory and routing decisions are defined sequentially. For the inventory part, we use five different replenishment methods to determine the order size from each retailer in each period. Then, a metaheuristic algorithm based on Large Neighbourhood Search (LNS) is used to solve the route optimization problem of scenario 1 and the second echelon part of scenario 2.

An experimental study is conducted to investigate the impact of the city hub's implementation in the city logistics context. Some test variables in this experimental study include the number of suppliers, number of retailers, holding cost, and replenishment method. Each test variable has several different values to be tested, and each unique combination of variables represents a scheme. Several data instances that correspond to each combination are generated. With this experimental study, we aim to see the effect of each variable involved and find in which combinations the city hub will be relevant to improve the city logistics performance. Some performance measures, such as total inventory and routing cost, number of travelled distances, loading degree, and number of trips, are used to evaluate each combination.

# Semi-flexible demand-responsive feeder bus systems for suburban areas using VNS

Fábio S. Vieira
KU Leuven Institute of Mobility, KU Leuven
e-mail: fabio.vieira@kuleuven.be

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

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

Public bus transport system in suburban areas are generally designed as feeder systems, with lines covering the whole area and bringing passengers to a transfer station, where passengers can access high-frequency services. These lines follow fixed routes and schedules, mostly with a low-frequency operation. Therefore, passengers need to adapt their trips to the timetable of the service.

To provide efficient services, it may be desirable to switch vehicles' operation between the traditional fixed services during peak hours and flexible services during off-peak hours, when demand is low and unstable (Kim and Schonfeld, 2012). These flexible services are categorized as semi-flexible demand-responsive feeder systems (DRFS), offering an operation in between a personalized taxi-like service, but keeping characteristics of the traditional feeder system (TFS), where passengers need to walk to the closest bus stop and with a certain frequency (Malucelli et al., 1999). The routes and departure time of buses, however, are flexible.

Several types of DRFS have been designed to address different challenges. The idea of combining DRFS with the operability of TFS is also explored by Quadrifoglio et al. (2008). In this system, the vehicle follows fixed routes with mandatory bus stops, but it includes the possibility to deviate according to logic detours, serving other points inside the operation area. Wei et al. (2020) develop a DRFS coordinating routing and frequency-setting of the buses with other services at the hub station, which is the common destination for passengers in feeder systems. Vansteenwegen et al. (2022) present a recent and comprehensive survey on demand-responsive systems.

In this research, a DRFS is developed to offer a flexible service during off-peak hours while keeping some of the characteristics the TFS has during peak hours in suburban areas. There is a regular frequency of operation and lines as in the TFS, and passengers access the service at fixed bus stops. However, departure times for each service will be optimized, together with the route used by the buses. To be more specific, limited detours and shortcuts are considered for each line of the TFS. It results in a set of possible routes that each bus can follow. A fixed number of buses have a number of services within the operation horizon and each service is associated to a line. A list of passengers' requests with their preferred departure time and desired bus stop are used for the optimization. The optimization of the DRFS services occurs just before the operation starts, assigning all requests to one of the services of the buses.

A Variable Neighborhood Search (VNS) is implemented to minimize the total travel time of passengers. In the VNS, the initial solution is the TFS operation, with buses associated to an initial fixed line. The local search moves to improve this initial solution try to assign passengers to a different service and remove empty services or add new services to a bus. Every time a set of requests is assigned to a bus, the route of the bus and the departure time can be optimized. In order to diversity the search, some bus services are randomly deleted and their requests are assigned to other services.

In a preliminary experiment, 15 different instances with 100 requests in the operation horizon were optimized in a network with 25 bus stops, 3 lines, and 3 buses with each 5 services in the operation horizon. The performance of the DRFS is compared with the TFS. The average waiting time for passengers at the bus stops is reduced by 36% in the DRFS, while the in-vehicle times remained roughly the same. In total, the DRFS could offer a service 24% faster than the TFS.

## References

Myungseob Kim and Paul Schonfeld. Conventional, flexible, and variable-type bus services. *Journal of Transportation Engineering*, 138(3):263, 2012.

Federico Malucelli, Maddalena Nonato, and Stefano Pallottino. Demand adaptive systems: some proposals on flexible transit. In *Operational research in industry*, pages 157–182. Springer, 1999.

Luca Quadrifoglio, Maged M Dessouky, and Fernando Ordóñez. Mobility allowance shuttle transit (mast) services: Mip formulation and strengthening with logic con- straints. *European Journal of Operational Research*, 185(2):481–494, 2008.

Pieter Vansteenwegen, Lissa Melis, Dilay Aktaş, Bryan David Galarza Montenegro, Fábio Vieira, and Kenneth Sörensen. A survey on demand-responsive public bus systems. *Transportation Research Part C: Emerging Technologies*, 137:103573, 04 2022. doi: 10.1016/j.trc.2022.103573.

Ming Wei, Tao Liu, Bo Sun, and Binbin Jing. Optimal integrated model for feeder transit route design and frequency-setting problem with stop selection. *Journal of Advanced Transportation*, 2020.

# Optimising Highway Vehicle Trajectories With a MILP; Case Study on the Ring of Antwerp

Maarten Wens
KU Leuven, Institute for Mobility,
e-mail: maarten.wens@kuleuven.be

Mohammed Ali Arman
KU Leuven, Institute for Mobility,
e-mail: mohammadali.arman@kuleuven.be

Ismail Abuamer
KU Leuven, Institute for Mobility,
e-mail: ismailma.abuamer@kuleuven.be

Chris Tampère
KU Leuven, Institute for Mobility,
e-mail: chris.tampere@kuleuven.be

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

Vehicle trajectory optimisation is a complex problem. Due to this complexity, many current models make unrealistic simplifications in order to get a computable result [1]. These simplified models are consequently solved using heuristic methods, putting a focus on decreasing calculation times instead of optimality [2]. Furthermore, this problem is often modelled using control theory. However, research from other domains indicate that a mixed integer linear program (MILP) might increase the quality of the solution and could be an attractive alternative to model this problem [3].

In order to be realistic, our model allows vehicles to overtake each other (violating the FIFO principle of classical queuing systems), it considers a heterogeneous traffic fleet, and safety headways between vehicles are speed dependent. These features are not taken into account in all attempts to optimise trajectories. To solve this problem, a 'first come first serve' principle is used in a first iteration, modelling selfish driving behaviour. Afterwards, this situation is reoptimized as a whole, modelling cooperative driving behaviour. This is done under the assumption of full knowledge. Therefore, it is possible in cooperative driving that a single vehicle does a certain action that is not beneficial for that specific vehicle, but improves traffic throughput as a whole.

In order to optimise this situation, a MILP is developed and solved using Gurobi. This model discretises time into frames. This discretisation imposes constraints every time frame for each vehicle on the road. Every vehicle is assigned an integer lane and a longitudinal position on the road. This position will realistically change every timeframe, depending on the speed and acceleration of each vehicle. Meanwhile, a safety distance is imposed around each vehicle, prohibiting other vehicles to be allocated too close to the current one, as long as both vehicles are driving in the same lane. Each vehicle is assigned a starting lane and time. The program makes sure each vehicle reaches its destination lane as fast as possible. The main bottlenecks are on- and off-ramps. These ramps force vehicles to do certain lane changes at a specific place on the highway, reducing the degrees of freedom of those vehicle trajectories. These forced lane changes can only happen if there is sufficient space available on the involved lanes. Because those trajectories are the hardest to implement, they can be given a priority.

The model works on simple highway layouts, where lane changes could only be beneficial but are not mandatory, or more complex road layouts, including merges, diverges, and weaving segments. As a case study, a weaving segment with three main lanes, two on-ramps, two-off ramps, and a lane reduction is modelled on the ring-road of Antwerp, one of Belgium's most complex highway sections. Our approach shows a high quality solution.

To speed up the process, multiple formulations for the same problem are tried by adding valid inequalities or using general, pre-defined Gurobi commands. An example of this is the distance between vehicles. In order to model this, the absolute difference between vehicle positions must be calculated. These absolute values can be modelled through big-M constraints or SOS-constraints. A second technique is a form of symmetry breaking: in many instances it does not matter where on the road a certain lane change happens. In order to solve this problem, a small incentive is added to the objective function, penalising earlier lane changes.

These techniques allow for bigger instances to be solved. In simple environments, a relatively steady flow of traffic can be modelled, as long as there is no real congestion occurring. In a more complex situation, with different vehicle types and a more complex road layout, a local cluster of 50 vehicles is modelled. Individual trajectories are added with a high accuracy (optimisation gap of 0.01%), while the re-optimisation has a gap of 2.5%. The results show that no unique optimum exist. Different driving behaviours could still lead to an equally valued solution. However, the cooperative model shows different driving behaviour compared to the selfish one. In selfish driving behaviour, the number of voluntary lane changes decreases, indicating that some lane changes are beneficial.

In a next step, this model could impose policies on the selfish driving behaviour in order to get closer to the cooperative situation. These regulations are incentivised by adding a certain penalty in the objective function, or as an additional constraint. An example of such an incentive penalises driving on the right hand side. This clears space around the ramps and makes it easier for the forced lane changes to happen. This policy opposes the current practice of right hand side driving. As a consequence, the system could be considered less safe.

# References


[1] Ehsan Amini, Aschkan Omidvar, and Lily Elefteriadou. "Optimizing operations at freeway weaves with connected and automated vehicles". In: *Transportation Research Part C: Emerging Technologies* 126 (2021), p. 103072.

[2] Bai Li et al. "Incrementally constrained dynamic optimization: A computational framework for lane change motion planning of connected and automated vehicles". In: *Journal of Intelligent Transportation Systems* 23.6 (2019), pp. 557–568.

[3] Marcos G Judewicz et al. "Linear programming based mpc of a grid-connected vsi". In: *2017 XVII Workshop on Information Processing and Control (RPIC)*. IEEE. 2017, pp. 1–6.

# Prescriptive maintenance with causal machine learning

Toon Vanderschueren
Research Centre for Information Systems Engineering, KU Leuven
toon.vanderschueren@kuleuven.be

Bart Baesens
Research Centre for Information Systems Engineering, KU Leuven
bart.baesens@kuleuven.be

Robert Boute
Research Centre for Operations Management, KU Leuven
robert.boute@kuleuven.be

Tim Verdonck
Applied Mathematics, University of Antwerp
tim.verdonck@uantwerpen.be

Wouter Verbeke
Research Centre for Information Systems Engineering, KU Leuven
wouter.verbeke@kuleuven.be

**Abstract**

Machine maintenance is a challenging operational problem, where the goal is to plan sufficient preventive maintenance to avoid machine failures and overhauls. Maintenance is often imperfect in reality and does not make the asset as good as new. Although a variety of imperfect maintenance policies have been proposed in the literature, these rely on strong assumptions regarding the effect of maintenance on the machine's condition, assuming the effect is (1) deterministic or governed by a known probability distribution, and (2) machine-independent. This work proposes to relax both assumptions by learning the effect of maintenance conditional on a machine's characteristics from observational data on similar machines using existing methodologies for causal inference. By predicting the maintenance effect, we can estimate the number of overhauls and failures for different levels of maintenance and, consequently, optimize the preventive maintenance frequency to minimize the total estimated cost. We validate our proposed approach using real-life data on more than 4,000 maintenance contracts from an industrial partner. Empirical results show that our novel, causal approach accurately predicts the maintenance effect and results in individualized maintenance schedules that are more accurate and cost-effective than supervised or non-individualized approaches.

Machine maintenance constitutes an intricate operational problem. The challenge is to avoid machine failures and costly overhauls, while simultaneously minimizing the cost of preventive maintenance (PM). Moreover, maintenance is often imperfect in practice since it does not restore the machine to a state as good as new. In fact, a broad spectrum of maintenance effects have been studied in the literature, ranging from perfect maintenance, which restores the system to a state as good as new, to worst maintenance, where maintenance causes the machine to fail.

Existing approaches in imperfect maintenance rely on strong assumptions regarding the effect of PM. First, the effect is modelled as either deterministic or stochastic assuming a certain probability distribution. These assumed effects, however, might not correspond to the actual effect. Second, the effect is typically assumed to be machine-independent, i.e., identical for all machines. In reality, the effect of the same type of PM intervention could be very different for different machines. For example, changing a gear would likely have a different impact on a brand new machine compared to the exact same maintenance intervention on an old, worn down machine.

This work relaxes both assumptions by proposing a completely data-driven maintenance policy that learns the effect of maintenance conditional on a machine's characteristics. The benefit of this approach is that it allows (1) to flexibly learn the maintenance effects from observational data (instead of assuming a certain deterministic or stochastic effect based on expertise), and (2) to design a machine-specific PM schedule based on these learned effects.

These benefits are achieved by framing maintenance as a problem of causal inference. We argue that the challenge in maintenance is that, for each specific machine, we only observe one outcome for the maintenance frequency that was administered in practice. We never observe the counterfactual outcomes – what would have happened if that machine received more or less maintenance in the past. Therefore, we never know whether the optimal maintenance frequency was prescribed. This is exactly the aim of causal inference, i.e. to predict each individual machine's potential outcomes in terms of failures and overhauls for different levels of PM. By learning a model that predicts the number of overhauls and failures given the PM frequency, we can optimize the PM schedule to minimize the total estimated cost. Essentially, we propose using observational data to learn a machine-specific digital twin for maintenance that predicts what would happen if a machine is prescribed a certain maintenance schedule.

| | **MISE** | |
|---|---|---|
| | Overhauls | Failures |
| SCIGAN | **7.71 ± 0.60** | **14.16 ± 1.68** |
| MLP | 10.25 ± 1.33 | 18.27 ± 3.65 |

| | **PE** | **PCR** |
|---|---|---|
| SCIGAN–ITE | **2.40 ± 0.46** | **1.07 ± 0.01** |
| MLP–ITE | 4.36 ± 1.25 | 1.11 ± 0.02 |
| SCIGAN–ATE | 8.77 ± 1.07 | 1.24 ± 0.04 |

Table 1: **Empirical evaluation.** We compare performance for the different policies over five runs. We evaluate each model's ability to predict the potential outcomes $o_i(t)$ and $f_i(t)$ (MISE), as well as each policy's ability to accurately prescribe the maintenance frequency (PE) and minimize costs (PCR). For all metrics, a lower value is better. Our proposed methodology outperforms both the non-causal (MLP–ITE) and the non-individualized approach (SCIGAN–ATE).

# Using Deep Learning for Life Event Prediction in the Financial Industry

S. Beyer Diaz
IESEG School of Management, Univ. Lille, CNRS, UMR 9221,
LEM - Lille Economie Management
e-mail : s.beyerdiaz@ieseg.fr

Kristof Coussement
IESEG School of Management, Univ. Lille, CNRS, UMR 9221,
LEM - Lille Economie Management
e-mail : k.coussement@ieseg.fr

Arno De Caigny
IESEG School of Management, Univ. Lille, CNRS, UMR 9221,
LEM - Lille Economie Management
e-mail : a.de-caigny@ieseg.fr

In the financial services industry, personalization of services has become essential for improving relationships with customers, increasing customer lifetime value and reducing attrition. To facilitate this type of personalization, it is important to be aware of a customers' needs. For this, life event detection and prediction models have been proposed as a way to better understand user preferences and behavior through time, as a means of improving cross-selling and loyalty. Thus, the objective of this paper is to advance the research of life event prediction in the financial services sector. The contributions are the evaluation of different deep learning models, specifically within the framework of comparing longitudinal and cross-sectional approaches. The study uses monthly customer snapshot data over a period of 18 months, obtained from an international financial services provider, where 12 months of data are used to predict whether a life event will occur or not in the following 6 months. The data used consists of 788,141 customers and 4 different life events. The results are compared using a cross-validated F1 measure. Preliminary results show the added value of longitudinal approaches. The study also provides important recommendations for financial services providers, specifically regarding method effectiveness, time horizon depth, and model optimization. To the best of our knowledge, our study is the first to evaluate deep learning model performance in both longitudinal and cross-sectional setups in the financial service sector.

# Causal Machine Learning for Loan Pricing

Christopher Bockel-Rickermann
KU Leuven, Research Centre for Information Systems Engineering (LIRIS)
e-mail: christopher.rickermann@kul euven.be

Tim Verdonck
University of Antwerp, Department of Mathematics

Wouter Verbeke
KU Leuven, Research Centre for Information Systems Engineering (LIRIS)

Pricing loans and mortgages is a challenging task for banks and lenders. A typical pricing mechanism must correctly assume and balance several uncertainties in the current and future relationship with a client to be profit-maximizing in the long run. On the one hand, a price must be sufficiently high to account for default and repayment risks, as well as to cover the running costs of the borrower. On the other hand, the price must match the potential customer's price sensitivity to maximize the probability of accepting the product on offer.

Hence, a loan price will typically combine two major constituents:

- A customized base price: A price that balances risk and reward for the lender, taking into account the characteristics of an applicant, the market sentiment and associated costs to the lender.
- A discretionary discount: Deductions from the base price granted by a sales person based on their judgement of the customer's willingness-to-pay (WTP) during origination to increase the probability of the customer to accept the product on offer.

In our work, we focus specifically on the effect of the discretionary discount. Unfortunately, the price sensitivity of the potential customer is not known ex-ante, so the true effect of the discretionary discount on the probability of loan acceptance is not known. Additionally, price sensitivity is diverse across customer types and the assignment of discounts is discretionary and potentially biased.

Our goal is to learn individual price sensitivity curves from observational data. To counter the potential bias in the assignment mechanism of discounts, we apply state-of-the-art methods from causal machine learning to enrich the input data and learn unbiased estimates of individual price sensitivity curves.

We work with a real-world data set from an institutional lender in Belgium. Our approach aims to support lenders in offering an optimal combination of the customized and discretionary price component of the loan rate.

We add to the literature in several ways. First, we apply individual continuous treatment effect modeling to pricing and provide a novel, causal perspective. Second, we provide use case for does-response curve modeling in business.

# Optimisation of order picking operations in a spare parts warehouse with dynamic order arrivals

Ruben D'Haen
Hasselt University, Research Foundation Flanders (FWO)
ruben.dhaen@uhasselt.be

Kris Braekers, Katrien Ramaekers
Hasselt University,
{kris.braekers, katrien.ramaekers}@uhasselt.be

Warehouses are an important part of many supply chains. To improve operational performance, optimisation of the order picking operations is considered indispensable [1]. Our research focuses on a manual, picker-to-parts, order picking system. In this setting, the order picking operations can be divided into several planning problems, i.e., the order batching, the picker routing and the batch scheduling problem. The order batching problem looks at which orders should be combined in a single pick tour throughout the warehouse. The picker routing problem decides on the most efficient path an order picker should follow to visit the storage locations of all items in his batch. Finally, the batch scheduling problem handles the assignment of batches to pickers, and determines the sequence in which the batches are picked. In the past years, research efforts shifted from the optimisation of these individual problems towards the optimisation of the integrated order batching, picker routing and batch scheduling problem (IBRSP). Since these planning problems are interrelated, solving the integrated problem leads to better overall results [2].

So far, studies have focused on a static problem setting, in which all orders are known at the start of the planning horizon. In practice, however, new and possibly urgent orders arrive throughout the day. To offer a good customer service level, these new orders should be included in the picking schedule as soon as possible. Therefore, our study extends the static IBRSP to account for dynamically arriving orders. By re-optimising previous schedules when new orders arrive, urgent orders can be handled very quickly in a cost-effective way.

A new metaheuristic algorithm, based on a large neighbourhood search, was developed to solve this optimisation problem. After developing and testing the algorithm in previous research, it is now applied to a real-life case study. We received order picking data from a spare parts warehouse and analysed the benefits of using our optimisation algorithm compared to the existing operational practices. The results show that large efficiency improvements are obtained by using an integrated optimisation algorithm. Moving from the company's current picker routing policy towards optimal picker routing leads to a considerable reduction in tardiness and will also improve the required order pick time. Changing the batching policy from the currently used earliest due time batching towards optimal batching, leads to even larger improvements regarding both tardiness and order pick time. By using the large neighbourhood search, the company can pick all orders in time while using less order pickers than currently required.

## 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(2), 481–501.

[2] van Gils, T., Caris, A., Ramaekers, K., & Braekers, K. (2019). Formulating and solving the integrated batching, routing, and picker scheduling problem in a real-life spare parts warehouse. European Journal of Operational Research, 277(3), 814–830.

# A proactive scheduling approach for a robotic mobile fulfillment system with processing time variability

Sander Teck
KU Leuven, Department of Mechanical Engineering - CIB
e-mail: sander.teck@kuleuven.be

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

Pieter Vansteenwegen
KU Leuven, Institute for Mobility – CIB
e-mail: pieter.vansteenwegen@kuleuven.be

The e-commerce sector is facing a rapid transition towards more automation and technological evolution [1]. The increased complexity of the warehouse operations requires control systems that can efficiently and effectively manage this complexity. Many researchers focus on the development of deterministic algorithmic frameworks to solve and study warehousing and manufacturing scheduling problems. However, no predetermined plan of operations will be accurate long after it is first taken into operation. The reason being that unexpected events and statistical fluctuations occur continuously during operations in warehousing/manufacturing environments. This uncertainty is an inevitable aspect of real-world scheduling problems and leads to scheduling disruptions. For instance, a typical cause of disruptions is the processing time variability [2]. In this study, we consider uncertain order picking times at the workstations. Furthermore, the possibility of failures of mobile robots to lift/lower inventory pods, resulting in a need for robot repositioning under the pod, is included. The disruptions may result in unnecessary delays or increasing waiting times which will only accumulate during the schedule execution. Therefore, this uncertainty has to be addressed during the scheduling process to ensure the generation of schedules that do not deteriorate the system performance as a result of process variability.

To this aim, we propose a proactive scheduling approach for solving the simultaneous scheduling and routing of mobile robots and human pickers in a robotic mobile fulfillment system (RMFS) subject to processing uncertainties. The proposed approach incorporates an evolutionary algorithm with Monte Carlo simulation sampling to find good schedules under uncertain conditions. The simulation sampling is used to determine the quality of candidate solutions, generated during the evolutionary process, for different stochastic scenarios. Thus, it acts as a fitness evaluation method to determine the solution quality and robustness to process variability. However, the use of simulations to evaluate the quality of a schedule requires considerable computational effort [3]. To reduce some of this computational effort, we incorporate an Optimal Computing Budget Allocation heuristic to efficiently allocate the available computation budget over the potential candidate schedules [4]. As an alternative, we develop and study a metamodel [5] to approximate the evaluation process of a candidate solution during the evolutionary process. This approximation method allows to reduce the computational time compared to the simulation approach. The objective of the proposed scheduling approach is to minimize the operational and capital costs related to the operation of the workstations and mobile robots in the warehouse environment [6]. The proposed frameworks will be extensively tested on a wide range of different instances. The instances will vary in both the number of picking orders that have to be scheduled, various warehouse layouts, and a varying number of available resources (e.g. mobile robots, workstation, inventory pods, etc.) in the warehouse system.

## References

[1] Boysen, N., de Koster, R., Weidinger, F. (2019). Warehousing in the e-commerce era: A survey. European Journal of Operational Research, 277: 396-411.

[2] Deblaere, F., Demeulemeester, E., Herroelen, W. (2011). Proactive policies for the stochastic resource-constrained project scheduling problem. European Journal of Operational Research, 214: 308-316.

[3] Ghasemi, A., Ashoori, A., Heavey, C. (2021). Evolutionary learning based simulation optimization for stochastic job shop scheduling problems. Applied Soft Computing, 106: 107309.

[4] Xiao, H., Lee, L., Ng, K. (2014). Optimal Computing Budget Allocation for Complete Ranking. IEEE Transaction on Automation Science and Engineering, 11(2): 516-524.

[5] Rojas-Gonzalez, S., Van Nieuwenhuyse, I. (2020). A survey on kriging-based infill algorithms for multiobjective simulation optimization. Computers and Operations Research, 116: 104869.

[6] Teck, S., Dewil, R. (2022). Optimization models for scheduling operation in robotic mobile fulfillment systems. Applied Mathematical Modelling, 111: 270-287.

# Adapting collision detection techniques from computer graphics for use in 2D irregular-shaped nesting problems

Jeroen Gardeyn
KU Leuven, CODeS, e-mail: jeroen.gardeyn@kuleuven.be

Tony Wauters
KU Leuven, CODeS, e-mail: tony.wauters@kuleuven.be

Nesting problems are widespread and occur in a range of different contexts, such as laser cutting, 3D printing and garment cutting. Determining whether a placement of an item is valid or not, which typically involves ensuring that no intersecting or overlapping occurs, is an essential part of all nesting algorithms. The efficiency and accuracy of these collision detection methods is therefore crucial. Since using exact trigonometry is typically too expensive, common techniques for collision detection in nesting problems include discretization techniques such as rasterization or employing no-fit polygons (NFP) by computing Minkowski sums. However, these techniques are not without their limitations. Discretization quickly becomes memory intensive when targeting for high precision while NFP generators often lack robustness. Moreover, a unique NFP has to be computed for every possible rotation of every pair of polygons.

Another field where collision detection is very relevant is computer graphics. In particular video games, simulations and CAD. However, while the essence of the collision detection problem at hand is the same, the context differs greatly from nesting problems. Videogames, for example, generally have a more dynamic but also much parser environment when compared to 2D nesting. Techniques used in computer graphics generally consist of a combination of broad- and narrow-phase collision detection. The goal of the broad phase is to efficiently eliminate as much of the required computational work as possible. Afterwards, what remains can be solved with precise (and therefore more expensive) checks during the narrow phase.

The goal of this research is to investigate whether such a two-phased approach can be adapted into a robust, fast and accurate collision detection system for use in 2D nesting heuristics. In a first implementation, a datastructure based on quadtrees is used for the broad phase, combined with exact intersection and inclusion tests in the narrow phase.

While this research is still in its early stages, initial experiments are already show promising results. The next steps are to develop a heuristic on top of this collision detection system which supports free rotation of items, irregular-shaped bins and holes.

# Efficient heuristics for irregular volume maximisation problems

Jonas Tollenaere
KU Leuven, CODeS
email: jonas.tollenaere@kuleuven.be

Tony Wauters
KU Leuven, CODeS
email: tony.wauters@kuleuven.be

The objective of volume maximisation problems is to find the biggest given item(s) that can be extracted from a larger three-dimensional object. This requires determining the placement of the necessary items so they can be maximally scaled up without violating any physical constraints. These problems appear in several real-world contexts, such as 3D-printing and gem cutting. Here, we study the category where both the item(s) and the larger object can be irregular. The combination of the three-dimensional nature and the irregular geometries results in challenging problems, with large solution spaces and compute-intensive evaluations. In this study, we will limit ourselves to instances where the volume of a single item must be maximised. While this may seem like a trivial optimisation problem, it is important to recognise its difficulty and usefulness as a building block for algorithms that solve harder variants of this problem. The goal of this research is to find good solutions using methods that scale well with the complexity of the item and the material. These objects are represented by triangular meshes. Previous research showed that working with exact and mathematical formulations soon requires an excessive amount of computation time as the number of triangles in the meshes increases. Given that we need to be able to handle real-world instances with meshes consisting of tens of thousands of triangles, (meta)heuristic approaches are better suited. We define a baseline approach that uses established techniques like bounding volume hierarchies for collision detection and bisection search to determine the maximum scale for an item with a given position and rotation. This subroutine can then iteratively be called by (meta)heuristic algorithms in search for the optimal position and rotation. While this already proves to be a successful approach, we can introduce improvements to significantly speed it up. A large portion of the computation time is spent in the subroutine for detecting collisions between the items and the material, and by exploiting the geometric properties of this optimisation problem we can reduce both the number of times we have to call the routine as well as the time it takes to execute it. Finally, we compare our improvements to each other and to a matheuristic approach. A dataset, consisting of both instances from related works and new real-world data, is used.

# Cutting trapezoids with rotations and angle-dependent loss

Hatice Çalık
KU Leuven, Department of Computer Science
e-mail: hatice.calik@kuleuven.be

Tony Wauters
KU Leuven, Department of Computer Science

Have you ever wondered how your window frames are produced? Four trapezoid metal or wooden parts with 45° end angles are connected two-by-two at the frame corners. It is highly likely that these four material parts are cut from one big stock. The customized sizes of frames require an optimization tool that will minimize the material waste when cutting from multiple stocks.

Frames are not the only products that require cutting stocks with angles. In many real-world applications such as roofing and pipe fitting, the end angles of parts may have different magnitudes as well as signs. Distinct angle values make the problem even more complicated to obtain good solutions with rules of thumb. In fact, the underlying optimization problem is NP-Hard as it is a generalization of the *bin packing problem*.

When cutting with angles, nesting parts with identical angles reduces both the number of cuts needed and the resulting additional waste. Hence, determining the sequence and the rotation of parts becomes essential. Problems of this nature have been addressed in the literature as *trapezoid packing* and *truss cutting* by Lewis and Holborn [2] and Lewis et al. [3], respectively. The truss cutting problem was also addressed by Garraffa et al. [1] who proposed methods for solving one-dimensional cutting stock problem with sequence-dependent cut losses (1D-CSP-SDCL).

We investigate a generalization of the aforementioned problems that integrates several additional practical real-world restrictions which cannot be tackled effectively with the 1D-CSP-SDCL methods in the literature. Consider a material profile that is not fully symmetric and can also be cut by rotating 90°, 180° or 270°. There may be multiple categories of parts to be cut. For example, category 1 parts can only be cut when the material is not rotated, category 2 parts can only be cut when when the material is rotated 90° and category 3 parts can be cut in either material orientation. In such a case, the optimization problem also requires deciding which parts to be cut from which orientation of the stocks. With this additional problem feature, certain part rotations will not be feasible for certain stock orientations. Therefore, one must only consider the feasible material-part rotation combinations.

All the practical restrictions mentioned in the previous paragraph create a very challenging optimization problem, which we refer to as the *Cutting Trape-*

*zoids problem with Rotations and Angle-dependent loss (CTRA)*. We initially develop a mixed integer programming formulation for the CTRA. In order to tighten the dual bounds of the formulation, we introduce area-based valid inequalities and utilize commodity-flow type of sub-tour elimination constraints. We evaluate small-scale instance solutions under different objective functions including minimum number of bins utilized, minimum total bin length utilized and maximum total reusable leftover length. In order to obtain high quality solutions efficiently, we developed an iterative best fit algorithm. The algorithm is able to obtain high quality solutions for medium- and large-scale instances in a few minutes if not in seconds.

## References

[1] Garraffa, M., Salassa, F., Vancroonenburg, W., Vanden Berghe, G., and Wauters, T. (2016). The one-dimensional cutting stock problem with sequence-dependent cut losses. *International Transactions in Operational Research*, 23(1-2):5–24.

[2] Lewis, R. and Holborn, P. (2016). How to pack trapezoids: exact and evolutionary algorithms. *IEEE Transactions on Evolutionary Computation*, 21(3):463–476.

[3] Lewis, R., Song, X., Dowsland, K., and Thompson, J. (2011). An investigation into two bin packing problems with ordering and orientation implications. *European Journal of Operational Research*, 213(1):52–65.

# Robust Alternative Fuel Refueling Station Location Problem with Routing under Decision-Dependent Flow Uncertainty

Özlem Mahmutoğulları
KU Leuven, Faculty of Economics and Business, ORSTAT
e-mail : ozlem.mahmutoullar@kuleuven.be

Hande Yaman
KU Leuven, Faculty of Economics and Business, ORSTAT
e-mail : hande.yaman@kuleuven.be

## 1 Introduction

Transportation is heavily dependent on fossil fuels. Using alternative fuel vehicles is a solution to break the transportation sector's reliance on consuming fossil fuels. The lack of alternative fuel station (AFS) infrastructure and the rather limited range of alternative fuel vehicles (AFVs) are two significant obstacles that are slowing down the introduction of AFVs. In this regard, the refueling station location problem (RSLP) has recently started to be studied in the literature. In the RSLP, the AFSs are located on the drivers' predetermined paths, which are usually the shortest paths. Since the drivers may tolerate deviating from their paths to refuel their vehicles, the RSLP with routing (RSLP-R) extends the RSLP and determines the locations of stations and routes of drivers simultaneously.

It is likely to observe uncertainties in the flows because the rollout of AFVs and the development of the AFS network are still at their initial stages. Moreover, the statistical data shows that the number of AFSs has a significant impact on the number of AFVs. It is thus important to consider that the availability of AFSs in the neighborhood affects the proliferation of AFVs during the development of infrastructure.

In this study, we incorporate robustness and decision-dependency into the problem and introduce the robust RSLP-R under decision-dependent polyhedral vehicle flow uncertainty. We derive mathematical programming formulations and propose a Benders reformulation and a branch-and-cut algorithm for the reformulation. We perform the following computational experiments : We first compare the performances of the proposed mathematical models and the Benders reformulation. Then, we investigate the changes in station locations and total covered flows when the optimal solutions of the deterministic, robust and decision-dependent robust problems are employed. We observe, under different parameter settings, that recognizing the uncertainty in flows and the decision-dependency of uncertain flow realizations may lead to significant gains in the total AFV flows covered.

# 2 Problem Definition and Solution Methods

The RSLP-R is defined on a road network and aims to maximize the total amount of AFV flows that can be refueled by locating a predetermined number of AFSs on the network by considering the willingness of drivers to deviate from their shortest paths to refuel their vehicles as well as the limited range of the vehicles. We use the deterministic problem introduced by [1] and introduce our flow uncertainty set using the hybrid model ([3]). The hybrid model comprises a hose model and an interval model. We define the hybrid uncertainty set of the vehicle flows under the impact of station location decisions. We suppose that, when a new station is opened, vehicle flows in the neighborhood increase because the drivers will be more willing to use AFVs if there are AFSs nearby. We derive two mathematical programming formulations. As the problem size grows, we encounter difficulties in solving these models, and thus we propose a Benders reformulation. We solve this formulation using a branch-and-cut algorithm. The separation, which is exact and polynomial, is done by inspection.

# 3 Computational Results

We use four different sized data sets to perform our computational experiments. The first one is a commonly used data set in the RSLP literature. We generated the other data sets based on the road network of Belgium. The nominal flow volumes are computed using the gravity model and the decision-dependency parameters are chosen using a similar way to that presented by [2].

We first compare the performances of the proposed solution methods. We observe that the Benders reformulation outperforms the other formulations. Then, we compare the station location decisions obtained by solving the deterministic, robust (without decision-dependency), and decision-dependent robust problems. We assess the importance of considering only uncertainty and uncertainty and decision-dependency simultaneously. In these experiments, we also examine the effect of different parameter settings on the results. Under all settings, we highlight the gain of incorporating uncertainty and decision-dependency into strategic-level decisions.

# Références


[1] Arslan, O., Karaşan, O. E., Mahjoub, A. R., & Yaman, H. (2019). A branch-and-cut algorithm for the alternative fuel refueling station location problem with routing. *Transportation Science*, 53 (4), 1107–1125.

[2] Basciftci, B., Ahmed, S., & Shen, S. (2021). Distributionally robust facility location problem under decision-dependent stochastic demand. *European Journal of Operational Research*, 292 (2), 548–561.

[3] Meraklı, M., & Yaman, H. (2016). Robust intermodal hub location under polyhedral demand uncertainty. *Transportation Research Part B : Methodological*, 86, 66–85.

# Multi-trip Vehicle Routing Problem with Access Restrictions

Munise Kubra Sahin
KU Leuven, Faculty of Economics and Business
e-mail: munisekubra.sahin@kuleuven.be

Hande Yaman
KU Leuven, Faculty of Economics and Business
e-mail: handeyaman@kuleuven.be

To cope with the negative effects of freight vehicles on the life quality in urban areas, many municipalities in Western Europe have implemented road accessibility restrictions such as *limited traffic zones*. The *limited traffic zones* represent the areas accessible only during a given time window. Implementing these time windows aims to reduce the congestion and pollution in the central areas. The problem of determining the vehicle routes in a city where vehicles are not allowed to enter certain areas during a given time period is called the Vehicle Routing Problem with Access Time Windows (VRPATW). The time-dependent restrictions in the VRPATW result in different waiting times depending on the arrival time at a street with access restrictions. These waiting times can create a trade-off between different arcs, even when the arcs have one distinctive attribute, which requires using a graph that contains the information of all Pareto-optimal paths to avoid jeopardizing the optimality. Therefore, we formulate the problem on a road network. We also allow vehicles to perform more than one trip per day to compensate the effect of longer travel distances caused by the access restrictions on the number of used vehicles. The resulting problem is multi-trip VRP with access restrictions on a road network. We formulate this problem using workday variables and propose a branch and price algorithm.

# Decentralized optimization of multi-commodity flows over time

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

Ilgin Efe Şenyuva
Eindhoven University of Technology
Industrial Engineering and Innovation sciences department
Operations Planning Accounting & Control

Tim Oosterwijk
Vrije Universiteit Amsterdam
School of Business and Economics, Operations Analytics
e-mail: t.oosterwijk@vu.nl

In today's logistics system, shippers and third party logistics service providers rely on carriers to execute a set of transport requests (referred to as shipments). These carriers operate a logistics network that consists of hubs — which can act both as the source or the destination of the shipments — and a set of links that connect these hubs with each other. It is up to the carrier to deploy resources (trucks, trains, barges,...) on the different links of their network to make sure all shipments can be routed from their origin to their destination within their respective time window.

The problem of the carrier can be represented mathematically by a *multi-commodity flow problem over time.* The multi-commodity flow problem is defined by a set of commodities (shipments) — each with their origin, destination, release time, deadline and volume — and a graph $G(V, A)$, in which $V$ represents the set of vertices (hubs) and $A$ denotes the set of arcs that connecting these hubs with each other.

With on average one-fifth (20.2%) of all road freight transports running with empty vehicle within Europe in 2020 (ranging between 6% in Belgium and 44.3% in Cyprus) and an average loading rate of the loaded vehicles around 50%, the performance, efficiency and sustainability of our logistics system is at least questionable. These poor results are mainly caused by a very fragmented sector, and limited economies of scale as well as an unbalanced configuration of the transport requests at individual company level. The recent COVID crisis, the high energy prices and increased uncertainty on the international markets increase the pressure on the sector even more.

Over the last decade, the field on collaborative logistics has proven the potential of horizontal collaboration in logistics with double-digit savings in both operational cost and environmental impact (e.g. $CO_2$ emission) by means of theoretical/computational studies as well as by reporting on actual case studies. These results, however, assume a central decision maker at the level of the coalition who has full information (in practice this is likely a neutral third party). Such a construction not only lacks scalability, but creates suboptimal plan for the individual members of the coalition.

Two streams of research can be distinguished within the context of our multi-commodity flow problem. First, there is the *network-focused research*. From the point-of-view of a central decision maker (in our example: the carrier) all flows are optimally routed through the network. Methods range from ILP-based formulations in combination with general solvers to specific algorithms for flow optimization using time-expanded networks [3]. Second, there is the *commodity-focused research*, in which researchers study the behaviour of single flow units while moving through the network, often based on game theoretical principles in the context of (a variant of) the congestion game [1,2].

In this research, we aim to bridge the gap between the network-focused and the commodity-focused research by means of mixed strategies that cover different scenarios of decentralized decision making and combine the best of both worlds.

Next to the common configurations in which individual carriers operate their own network (no collaboration) or a centralized planner at system level (full collaboration), we investigate the potential of simple local routing protocols at the level of the carrier, the hub and the individual shipment. Adopting a game theoretical approach, we are interested in the efficiency (with respect to social optimum and price of anarchy) and stability (convergence to equilibrium) of these protocols.

As such, our goal is to generate recommendations on how we could facilitate the creation of an integrated logistics network (often referred to as the physical internet) in a way that its protocols are intuitive, transparent, allow for free entry and exit of entities in the system, and are easily scalable.

## References

[1] Hara, T., Sasabe, M., & Kasahara, S. (2020, December). Distributed route selection under consideration of time dependency among agents' road usage for vehicular networks. *In Proc. 2020 International Conference on Emerging Technologies for Communications (ICETC2020)*, Virtual, E2-1.

[2] Holzman, R., & Monderer, D. (2015). Strong equilibrium in network congestion games: increasing versus decreasing costs. *International Journal of Game Theory*, 44(3), 647-666.

[3] Nagaraj, S. V. (2021). Review of Network Flow Algorithms David P. Williamson. ACM SIGACT News, 52(1), 12-15.

# Human Resource Analytics: Employee Journeys from a Process Perspective

Simon De Vos
KU Leuven, Faculty of Economics and Business
simon.devos@kuleuven.be

Johannes De Smedt
KU Leuven, Faculty of Economics and Business
johannes.desmedt@kuleuven.be

Wouter Verbeke
KU Leuven, Faculty of Economics and Business
wouter.verbeke@kuleuven.be

In recent years, many organizations focus more and more on exploiting the vast amount of available data to support decision making in their daily activities for gaining a competitive advantage [1, 2]. The field of human resources (HR) is strongly anticipating the transformative potential of big data, analytics and machine learning. Hence, albeit with a certain delay [3, 4], HR departments within organizations are actively looking for ways to adopt analytical techniques to execute activities where current evidence-based techniques are supplemented with objective and data-driven insights [5, 6]. Consequently, HR analytics as a discipline and research field is strongly gaining importance. In this proposal, we look at HR analytics from a distinct perspective where career paths, which are characterized by a trace of subsequent events, are considered as processes.

Despite the notable rising enthusiasm on HR analytics, the adoption rate by businesses of insights from scientific literature to model-based tools in practice remains limited. The reason for this lack of adoption is threefold. First, HR practitioners need to be able to work with these analytical tools. This is not evident as we often observe a gap between the data-related knowledge and skills of HR departments who use them on the one hand and data science teams who design them on the other hand. Second, the availability of data within an HR department is often limited. Hence, the application of complex algorithms that require lots of data to train a model that for example automatically screens incoming CVs is often not feasible and therefore inadequate in practice. Third, the goal of analytical tools in the setting of HR is to provide insights to support rather than to automate decision making. The deployed model needs to be explainable. The proposed methods based on process analytics can cope with these three constraints as they are accessible in use, can provide insight based on a relatively low amount of data and have high explainability. On top of that, careers are highly behavioral. Consequently, modeling them requires highly dynamic methods, making process mining techniques a great fit with career analytics.

Process mining techniques are applied on HR data to automatically discover employee journey maps. The purpose of an employee journey map is twofold. First, it is used to give an overview of all job positions and transitions between these positions within an organization. Second, it can be used to gather insights on opportunities of how decision-making can be improved. For example, employee maps can detect hard-to-fill positions and shed light on the profiles of employees that are attracted to these positions. The methods in this proposal are not intended to automate decision-making, but rather to support and improve the quality of human-made decisions. Given the longitudinal structure of HR data, a series of subsequent activities can be defined per employee. Hence, the HR data can be transformed to the format of an event log on which it is convenient to apply process mining techniques [7]. This results in a comprehensive end-to-end process view of career paths of employees within an organization, including decision points indicating how they traverse within this job-function space.

Concepts in the field of process mining can be translated to HR concepts in the following way. (i) A case is conceptually the equivalent of an employee. Over time, the employee can transfer to another function, undertaking a journey. (ii) An activity translates to occupying a function within an organization. It is executed by an employee and is characterized by a timestamp on which the activity starts and ends. (iii) A trace covers all the activities performed in a particular process instance by a certain case (employee). Hence, it is the equivalence of an employee journey. Each trace is defined as the ordered set of subsequent functions that an employee occupies throughout their career within an organization. (iv) Applying process discovery techniques on an event log consisting of longitudinal HR data is conceptually equal to the discovery of employee journey maps. The event log is defined as the collection of traces and thus contains the information of all employee journeys. An employee journey map is the aggregate of all possible paths that an employee can journey within an organization

This research focuses on the generation and descriptive analysis of employee journey maps. We can look at this employee journey map from two perspectives. We can adopt the perspective of an individual employee and analyze individual employee journeys and the various variants of these different paths. Alternatively, we can zoom in on specific functions and analyze their inflow and outflow, as well as other properties like average lead time.

In this research, we explore the potential of applying process mining methods for employee journey mapping. By adopting a process perspective to the analysis of HR data, we aim to further leverage the use of data analytics for decision-making in HR Management, to the benefit of both employers and employees. Based on this preliminary analysis, we identify a wide variety of process analysis methods that can be applied on HR data in the format of an event log, extending upon the descriptive methods by also leveraging predictive and prescriptive approaches.

## References

[1] Daniel J Power. Understanding Data-Driven decision support systems. Information Systems Management, 25(2):149–154, March 2008.

[2] Foster Provost and Tom Fawcett. Data science and its relationship to big data and Data-Driven decision making. Big Data, 1(1):51–59, March 2013.

[3] Vicenc Fernandez and Eva Gallardo-Gallardo. Tackling the HR digitalization challenge: key factors and barriers to HR analytics adoption. Copmetitiveness Rev. J., 31(1):162–187, July 2020.

[4] John Boudreau and Wayne Cascio. Human capital analytics: why are we not there? Journal of Organizational Effectiveness: People and Performance, 4(2):119–126, January 2017.

[5] David Angrave, Andy Charlwood, Ian Kirkpatrick, Mark Lawrence, and Mark Stuart. HR and analytics: why HR is set to fail the big data challenge. Hum. Resour. Manag. J., 26(1):1–11, January 2016.

[6] Thomas Rasmussen and Dave Ulrich. Learning from practice: how HR analytics avoids being a management fad. Organ. Dyn., 44(3):236–242, July 2015.

[7] Wil van der Aalst. Process mining: Overview and opportunities. ACM Trans. Manage. Inf. Syst., 3(2):1–17, July 2012.

# Capturing complexity over space and time via deep learning: An application to real-time delay prediction in railways

Léon Sobrie
Ghent University
e-mail: leon.sobrie@ugent.be

Marijn Verschelde
IÉSEG School of Management, LEM (UMR-CNRS 9221), Ghent University
e-mail: m.verschelde@ieseg.fr

Veerle Hennebel
KU Leuven
no e-mail

Bart Roets
Infrabel, Ghent University
e-mail: bart.roets@infrabel.be

Predictive analytics is increasingly used in managerial decision-making. These decisions are mostly driven by approaches that are inflexible over space and time. However, the decision environment is often dynamic and complex. In addition, business systems are becoming increasingly integrated, complex and heterogeneous. This development creates a growing need to frequently update the models and increase the flexibility of the implemented models. It is within this context that we investigate whether the use of deep learning results in more added value in real-world settings. Moreover, previous research on deep learning stipulates that its adoption in business applications is still in its infancy. In particular, we focus on train delays that are proactively monitored in digital control rooms. Herein, we show that the added value generated by a flexible deep learning model, in comparison with an industry-validated, real-world, rule-based system, increases with the real-world complexity of the managerial decision setting over space and time. The proposed deep learning model is based on a customized recurrent neural network structure (i.e., an LSTM (Long Short-Term Memory) encoder-decoder model) including uniquely rich information on the spatio-temporal interdependence between trains. Besides presenting the new comparative findings, we show the usefulness of the proposed approach by implementing the deep learning model as a proof of concept within Belgium's railway infrastructure company Infrabel. In this setting, the proof of concept provides decision support for operators, supervisors and managers.

# Probabilistic forecasting with modified N-BEATS networks

J. Van Belle
KU Leuven, LIRIS
e-mail: jente.vanbelle@kuleuven.be

R. Crevits
OMP
e-mail: rcrevits@omp.com

W. Verbeke
KU Leuven, LIRIS
e-mail: wouter.verbeke@kuleuven.be

In this paper, we present a modification to the state-of-the-art N-BEATS deep learning architecture for the univariate time series point forecasting problem for generating probabilistic forecasts. Next, we propose an extension to this probabilistic N-BEATS architecture to allow optimizing probabilistic forecasts from both a traditional forecast accuracy perspective as well as a forecast stability perspective, where the latter is defined in terms of a change in the forecast distribution for a specific time period caused by updating the probabilistic forecast for this time period when new observations become available, i.e., as time passes. For the M4 monthly data set, we show that this leads to more stable forecast distributions without leading to a significant deterioration in their accuracy. Finally, we present a second extension to the probabilistic N-BEATS model for use in an inventory management context in that also probabilistic forecasts of temporal aggregates, i.e., cumulative demands, are produced by the network. Results are reported for the M4 monthly data set and indicate that large improvements in accuracy can be obtained over basic but well-established methods to produce probabilistic cumulative demand forecasts.

**Keywords:** Aggregates, Deep learning, Forecast instability, Global models, Inventory management, Quantiles

# Aggregate planning for multi-product assembly systems with reconfigurable cells

M. Uzunosmanoglu*,†, V. Limère‡,†, B. Raa*
Ghent University,
*Industrial Systems Engineering (ISyE)
‡Department of Business Informatics and Operations Management
†Flanders Make@Ghent University
e-mail: mehmet.uzunosmanoglu@ugent.be

**Introduction** Reconfigurable Manufacturing Systems (RMS) are proposed to find a balance between the efficiency of Dedicated Manufacturing Systems (DMS) and the flexibility of Flexible Manufacturing Systems (FMS), and thus to adapt to the demand volatility and changing product functionality at minimum effort and cost.[1] For these systems, traditional optimization problems, such as scheduling and system design, have been considered, and different solution methods have been proposed in the literature.[2] However, in the project AssemblyRecon[3] the aim is to design a Reconfigurable Assembly System to meet the demand across multiple periods at minimum cost, and this problem has not been tackled by previous research considering such a system. This design problem consists of three different subproblems which are solved simultaneously: cell configuration, system configuration and product routing. Cell configuration focuses on selecting resources and allocating them to the hexagonal cells to ensure that all assembly operations can be performed. System configuration considers opening and positioning the cells and conveyors, such that parts can visit the different cells along the assembly system, taking into account possible specific layout requirements. Product routing, finally, aims to find a routing for all parts and products through the system, ensuring the availability of resources is not exceeded and operation precedence restrictions are not violated. Note that the solution of the problem serves as an input for further detailed scheduling in the project, which is beyond the scope of this work.

**Reconfigurable Assembly System** The system in focus consists of a layout with possible positions for hexagonal cells. The cells are connected with unidirectional conveyors such that the part flow from the first to the last cell is ensured.

[1] Y. Koren et al. "Reconfigurable Manufacturing Systems". In: *CIRP Annals* 48.2 (1999), pp. 527–540. ISSN: 0007-8506. DOI: https://doi.org/10.1016/S0007-8506(07)63232-6

[2] Abdelkrim Yelles-Chaouche et al. "Reconfigurable manufacturing systems from an optimisation perspective: a focused review of literature". In: *International Journal of Production Research* 59 (Oct. 2020). DOI: 10.1080/00207543.2020.1813913.

[3] This research was supported by Flanders Make within the scope of the project AssemblyRecon.

It is also possible to add an additional conveyor. The hexagonal cells consist of a central module and multiple side modules. Side modules can be equipped with a resource and then used for performing a particular assembly operation, whereas the central module is a robotic arm that moves parts between side modules, as well as between a side module and a conveyor. It is assumed that any side module can be equipped with any resource, but a resource can only perform one type of operation. The time and cost of an assembly operation are part/product-dependent and resource-dependent. In addition to this, the central modules might be able to handle only a subset of parts/products due to geometrical restrictions. The system is reconfigurable because all the resources can be reallocated, removed or inserted from one period to the next. However, these actions have certain costs. In the base problem instance, there are three cell positions and four different products, each following a fixed sequence of operations.

**Solution Approach** The problem is formulated as an Integer Quadratic Programming model. In the model, three types of decision variables are defined: resource assignment, product routing and scenario variables. The scenario variables are used to assess different (re)configuration scenarios between periods. The objective of the model is to minimize the total cost which consists of processing and (re)configuration costs. There are four different types of constraints. First, resource assignment constraints ensure meeting specific layout and logical rules as well as production requirements. Secondly, scenario constraints make sure that an appropriate (re)configuration scenario is selected for each resource considered in the system. Thirdly, product constraints ensure that the products are assembled according to the assembly requirements and the demand. Finally, availability constraints avoid exceeding the capacity of the resources in the system. After modelling the problem, it is solved for small instances using Gurobi.

**Conclusions** Small problem instances are solved to optimality in limited time (a few seconds). However, for larger instances, this is no longer the case. Therefore, (meta)heuristic approaches are being developed as the next step. Along with this, the model is also being generalized. This involves considering automatic guided vehicles (AGVs) or other material handling systems for inter-cellular transportation, and extending the number of cell positions to more than three. In addition, while a fixed assembly operation sequence for each product is considered in this paper, a flexible sequence of operations (with precedence constraints) is often possible in practice. Therefore, the model will be extended to adopt this flexibility by allowing alternative sequences.

# Modeling supermarket sizing and placement in the assembly line feeding problem

Ebenezer Olatunde Adenipekun, Veronique Limère
Ghent University, Department of Business Informatics and Operations Management
FlandersMake@UGent — core lab CVAMO, Ghent, Belgium
e-mail: ebenezerolatunde.adenipekun@ugent.be, veronique.limere@ugent.be

Nico André Schmid
IESEG School of Management, Univ. Lille, CNRS,
UMR 9221 - LEM - Lille Economie Management, F-59000 Lille, France
n.schmid@ieseg.fr

With the steep rise in number of parts required in an assembly line due to customers' ever-increasing preference for customized products, many different part variants are required to assemble the corresponding products. Consequently, practitioners and academic researchers face the challenge of managing assembly system more efficiently, while considering the limited space at the assembly line. Hitherto, researchers have dealt with the assembly line feeding problem (ALFP) from different perspectives through various descriptive and optimization-based models (see, e.g., [2, 1]), however, incorporating placement and sizing of cells into the ALFP has received little or no attention. In this study, a cell is defined as a preparation area located next to each assembly station (line-integrated) or between the warehouse and the assembly line (regular), where parts are repackaged. Each of these cell kinds have specific characteristics in operation and impact the overall feeding costs differently. The collection of all cells constitutes the supermarket. This study aims to fill this research gap by including cell sizing and placement as well as supermarket sizing decisions into the ALFP, i.e., the assignment of every part to the most suitable feeding policy. Placement and sizing decisions are concerned with the determination of the cell's size and the positioning of dedicated cells. The integration of cell placement and sizing decisions into the ALFP will lead to accurate computation of transportation distances to (and from) the assembly line and in turn, minimize the transportation cost. Furthermore, dealing with the sizing problem which involve proper design of the cells will lead to the minimization of space cost. The size of every cell is decided based on the parts allocated to it, whereas the supermarket size is rectangular and depends on placement and sizing of the cells contained.
We propose a mixed-integer programming model that simultaneously assigns every part to one of five feeding policies, i.e., line stocking, boxed supply, sequencing, stationary and traveling kitting and a suitable cell (line-integrated or regular) to every part that requires preprocessing. The proposed model minimizes total ma-

terial handling costs incurred through four subprocesses: (1) replenishment of the cells; (2) preparation; (3) transportation to the line; and (4) usage at the line. We propose some constructive algorithms that provide ample solutions to the transportation problem arising from navigating, firstly, within the supermarket and, secondly, between the supermarket and the assembly line. In addition, the result provided by these algorithms are compared with results obtained from the Concorde solver. The proposed algorithms will be used within a callback procedure while solving the model. The model will be solved for some artificial datasets that have been created based on an industrial case study.

## References

[1] Adenipekun, E. O., Limère, V., and Schmid, N. A. 2021. The impact of transportation optimisation on assembly line feeding. *Omega* 102544.

[2] Limère, V., van Landeghem, H., Goetschalckx, M., Aghezzaf, E. H., and McGinnis, L. F., 2012. Optimising part feeding in the automotive assembly industry: Deciding between Kitting and Line Stocking. *International Journal of Production Research* 50 (15): 4046–4060.

# The influence of the target delivery lead time on the design and capital cost of a make-to-order multiproduct batch plant

Amy Van Meir
University of Antwerp, Department of Engineering Management
e-mail: amy.vanmeir@uantwerpen.be

Trijntje Cornelissens
University of Antwerp, Department of Engineering Management
e-mail: trijntje.cornelissens@uantwerpen.be

Johan Springael
University of Antwerp, Department of Engineering Management
e-mail: johan.springael@uantwerpen.be

Most batch plant design models in academic literature focus on optimizing capital or operational costs. However, having a responsive production environment is equally important in today's business world. A metric to define responsiveness is the delivery lead time (DLT), i.e. the time a customer waits between ordering and receiving an order. Having a responsive plant can be expressed as having a competitive DLT for each customer order. In case of batch plant design, the transport to the customer is excluded from the DLT. In this presentation, we focus on the responsiveness of a multiproduct batch plant in a make-to-order (MTO) environment. We will introduce a target delivery lead time for all customer orders to the plant design model and observe the impact of this lead time on the capital cost of the production equipment units and storage tanks needed to cover the total demand over a specific planning horizon. In our MTO batch plant model, each customer order is produced in a single batch. Also, orders finished earlier than their delivery date/due date are stored in non-dedicated storage tanks until loaded for transport and remain there for at least a few hours, to account for administration and quality control. Since customer orders have a due date, and their delivery lead time is constrained by the target delivery lead time, the earliest start date for production of all batches are known. The number and sizes of production equipment units and storage tanks are obtained by minimizing the total capital cost of the plant. The effect of the target delivery lead time is analysed for different characteristics such as the length of the planning horizon, the number of customer orders and the total volume of all orders to be produced. For each unique combination of the characteristics, different sets of customer orders are considered. The mathematical model is solved by means of exact methods and heuristics in case of respectively small and large instances.

# Probability based models to build football fixtures

Cristian Aguayo
Université libre de Bruxelles, Unité de recherche en Graphes et Optimisation mathématique — e-mail: cristian.aguayo.quintana@ulb.be

Football is one of the most popular sports worldwide and, given this popularity, it is not surprising that there are social and economic interests in the sport. One of the main activities associated with the development of football competitions is the scheduling of matches, which takes place prior to the start of a championship. It is possible that more than one match may be feasible and, in that case, one of them must be chosen.

This paper presents probabilistic models for simulating football matches. Combining these models with mathematical programming models, we seek to generate different types of fixtures using different objective functions. The models were calibrated and tested using data from the South American qualifiers.

# Optimizing the selection of a soccer team in Fantasy Premier League

Jens Goemaere
KU Leuven, Department of Computer Science
e-mail: jens.goemaere@kuleuven.be

Lennert Claerhout
KU Leuven, Department of Computer Science
e-mail: lennert.claerhout@student.kuleuven.be

Fantasy Premier League [1] is a game where a player acts as a football team manager. The fantasy game is based on the English Football League, the Premier League. Points in the game can be scored based on real-life performances of the selected team. To boost points, specific game chips can be used to for example allow free transfers during one week. This study aims to create a model that predicts the expected points of the players, optimizes the selected team and examines multiple strategies how to use the game chips.

The first part of this study gives a Mixed Integer Programming (MIP) formulation of the game. The MIP is used to extract interesting insights from previous seasons, such as the best formations or the optimal transfer strategy. Furthermore, the MIP is used as a benchmark for the prediction model for seasons from the past. For passed seasons, solutions of the MIP provide the ultimated strategy.

The second part of this study, puts the focus on the prediction of the points of each player. The prediction model uses a Recurrent Graph Evolution Neural Network (ReGENN) [2]. The ReGENN architecture combines the information of multiple performance statistics of every player recorded every game week over past seasons. The predicted scores together with a mathematical model to select our initial team. Running both models, the prediction model and the MIP, provides an overview of transfers that should be performed for the most optimal team.

The third part of the study gives a comparison of three strategies for the use of the game chips. A stochastic strategy was compared with a deterministic strategy and a strategy based on the Secretary Problem [3].

In the last part of the stury, the results of the model can be found. The proposed model is validated on both the 2020-21 and the 2021-22 season.

## References

[1] Premier League, *Fantasy Premier League*, `https:/fantasy.premierleague.com/`. Accessed August 19, 2022.

[2] Gabriel Spadon and Shenda Hong and Bruno Brandoli and Stan Matwin and Jose F. Rodrigues-Jr and Jimeng Sun, *Pay Attention to Evolution: Time Series Forecasting with Deep Graph-Evolution Learning* (2020), available at `2008.12833`.

[3] Thomas S. Ferguson, *Who Solved the Secretary Problem?*, Statistical Science **4** (1989), no. 3, 282 – 289, DOI 10.1214/ss/1177012493.

# A pessimist's approach to one-sided matching

Tom Demeulemeester
KU Leuven, Research Center for Operations Research & Statistics
Ghent University, Department of Business Informatics and Operations Management
e-mail: tom.demeulemeester@kuleuven.be

Dries Goossens
Ghent University, Department of Business Informatics and Operations Management
FlandersMake@UGent - corelab CVAMO
e-mail: dries.goossens@ugent.be

Ben Hermans
KU Leuven, Research Center for Operations Research & Statistics
e-mail: ben.hermans@kuleuven.be

Roel Leus
KU Leuven, Research Center for Operations Research & Statistics
e-mail: roel.leus@kuleuven.be

Inspired by real-world applications such as the assignment of pupils to schools or the allocation of social housing, the one-sided matching problem studies how a set of agents can be assigned to a set of objects when the agents have preferences over the objects, but not vice versa. For fairness reasons, most mechanisms use randomness, and therefore result in a probabilistic assignment. We study the problem of decomposing these probabilistic assignments into a weighted sum of ex-post (Pareto-)efficient matchings, while maximizing the worst-case number of assigned agents. This decomposition preserves all the assignments' desirable properties, most notably strategy-proofness. Next to discussing the complexity of the problem, we obtain tight lower and upper bounds on the optimal worst-case number of assigned agents. Moreover, we propose two alternative column generation frameworks for the introduced problem, which prove to be capable of finding decompositions with the theoretically best possible worst-case number of assigned agents, both for randomly generated data, and for real-world school choice data from the Belgian cities Antwerp and Ghent. Lastly, the proposed column generation frameworks are inherently flexible, and can therefore also be applied to settings where other ex-post criteria are desirable, or to find decompositions that satisfy other worst-case measures.

# Developing Heuristic Algorithms for Graph Optimisation Problems using Tree Decompositions

Louis Carpentier
KU Leuven, Department of Computer Science
e-mail: louis.carpentier@student.kuleuven.be

Jan Goedgebeur
KU Leuven, Department of Computer Science
e-mail: jan.goedgebeur@kuleuven.be

Jorik Jooken
KU Leuven, Department of Computer Science
e-mail: jorik.jooken@kuleuven.be

Many combinatorial graph optimisation problems are $\mathcal{NP}$-hard. One way to deal with $\mathcal{NP}$-hardness is by exploiting special structures in the input graph. The *tree decomposition* of a graph, introduced by Robertson and Seymour (1984), is one such structure. Every graph can be converted into a tree decomposition. This is a tree in which every node of the tree contains a subset of the vertices of the graph. For simplicity, we call the vertices of the graph effectively vertices, but adopt the term node when referring to the tree decomposition. A tree decomposition must satisfy three conditions:

1. every vertex of the graph is contained in at least one node of the tree,
2. for every edge in the graph, there exists a node that has both endpoints,
3. for every vertex, the set of nodes containing that vertex are connected.

The width of a tree decomposition is the maximal number of vertices contained in a node, minus one. Generally the tree decomposition of a graph is not unique. The treewidth of a graph equals the minimum width over all tree decompositions of that graph. The treewidth measures how much the graph resembles a tree.

A tree decomposition allows for a dynamic programming approach to solve multiple problems, by computing tables of partial solutions at each node of the tree decomposition in a bottom-up fashion. This approach yields exact solutions, but is impractical when the treewidth of the graph is too large. We modify this procedure by only storing a limited number of all dynamic programming states, and thus only considering a subset of all partial solutions. By doing so we drop the exactness guarantee of our algorithm in favour of a shorter running time.

Concretely, we developed such algorithm for the Maximum Happy Vertices (MHV) problem, introduced by Zhang and Li (2015). Given a graph with coloured vertices. A vertex is said to be *happy*, if all its neighbours share the same colour. Otherwise, the vertex is said to be *unhappy*. Given a graph of which some vertices are coloured, the goal of the MHV problem is to colour every vertex, such that the number of happy vertices is maximised.

Agrawal (2017) proposes an exact algorithm for solving the MHV problem, using tree decomposition. For each node in the tree decomposition, all colour- and happiness-assignments of all vertices in the node are considered. This results in $(2k)^{w+1}$ dynamic programming states for a tree decomposition with width $w$ and an instance to the MHV problem with $k$ colours. Our algorithm only constructs $W$ partial solutions in each node, with $W$ a user defined parameter. By setting $W$ small, we drastically reduce the number of dynamic programming states to consider, and consequently the running time. However, if

$$(2k)^{w+1} < W, \tag{1}$$

then our heuristic algorithm considers all partial solutions, and constructs an exact solution. Furthermore, it proves exactness of the constructed solution. We obtain either an exact or heuristic algorithm, only by setting the value of $W$.

Our algorithm constructs exact solutions more efficiently than the algorithm by Agrawal (2017). Furthermore our algorithm constructs higher quality colour functions than Greedy-MHV and Growth-MHV when more than 40% of the vertices are initially coloured. These instances require a more precisely constructed colouring in order to find those vertices that can be happy. Greedy-MHV and Growth-MHV do have shorter running times however, because our algorithm performs a relatively complex operation at each node of the tree decomposition.

The performance of our algorithm depends on the width of the tree decomposition. The number of possible partial solutions is exponential in the treewidth, as is shown in Equation (1). Our algorithm is able to construct high quality solutions for graphs with a small treewidth, since the space of partial solutions has a limited size. Moreover, it is possible to prove exactness of the solution, which is not possible with other heuristics.

We propose a new approach for developing algorithms using tree decomposition, specifically for the MHV problem. A wide variety of exact algorithms using tree decompositions exist in the literature. Our approach can extend each of these, resulting in numerous new heuristic algorithms.

## References

[1] A. Agrawal. “On the parameterized complexity of happy vertex coloring”. In: *International Workshop on Combinatorial Algorithms*. Springer. 2017, pp. 103–115.

[2] N. Robertson and P. D. Seymour. “Graph minors. III. Planar tree-width”. In: *Journal of Combinatorial Theory, Series B* 36.1 (1984), pp. 49–64.

[3] P. Zhang and A. Li. “Algorithmic aspects of homophyly of networks”. In: *Theoretical Computer Science* 593 (2015), pp. 117–131.

# Maximal Acyclic Subgraph Optimization For Gene Regulatory Networks

K. Stoop Ghent University - imec, kenneth.stoop@ugent.be
T. Michoel University of Bergen - Computational Biology Unit, tom.michoel@uib.no
V. Vermeirssen Ghent University - CBIGR, CRIG, vanessa.vermeirssen@ugent.be
M. Pickavet Ghent University - imec, mario.pickavet@ugent.be
P. Audenaert Ghent University - imec, ghentuniversity@duck.com
D. Colle Ghent University - imec, didier.colle@ugent.be

**Introduction** Weighted digraphs pop up in many different areas of research, e.g. directed social networks, natural language semantic dependency parsing, or regulatory and signaling networks in biology. In many cases one wants the directed graph modelling the data to be acyclic, as it often represents an underlying causal structure, e.g. Bayesian networks. Determining the "most probable" underlying acyclic subgraph from a given fully connected weighted digraph, where weight reflect the confidence of arcs or relations being present, amounts to the *maximum acyclic subgraph problem* (MDAG), i.e.
*Given a complete weighted digraph $G = (V, A)$, find a subset of arcs $A^*$ such that the sum of their weights is maximal and the resulting subgraph $G^*$ is acyclic.*
This is an NP-hard problem, which beyond its interest in practical applications, is also of theoretical interest: it is a classic example of an independence system for which there are not necessarily polynomial approximation algorithms with a strong guaranteed error bound.

Recently a method was developed that estimates the probability of a directed interaction of gene $X$ on gene $Y$ from genetic (DNA-based) and gene expression (RNA-based) data, and that is sufficiently efficient to handle data with measurements for thousands to ten-thousands of genes simultaneously, i.e. on the order of millions to hundreds of millions of pairwise interaction scores[1]. Finding the optimal acyclic subgraph within this large matrix of arc weights for subsequent Bayesian network reconstruction is a promising approach for modelling large-scale gene networks.

**Algorithms** In order to tackle the MDAG problem[2], 3 heuristics are considered: 1. Heuristic sorting (HS); 2. Greedy insertions (GI); 3. Simulated annealing (SA) based on GI. In HS, a topological ordering is constructed based on a value $f(i)$ which denotes the "importance" of node $i$. More precisely, we defined it to

[1] Wang, Lingfei, and Tom Michoel. "Efficient and accurate causal inference with hidden confounders from genome-transcriptome variation data." PLoS computational biology 13.8 (2017): e1005703.

[2] The MDAG problem will be formulated based on the concept of topological ordering. The order of the nodes is kept in a list $l$, where $l(x)$ denotes the label of the node at position $x$ in the list. The objective then becomes $E = \sum_{x<y}^{n} W_{l(x)l(y)}$ with $W_{ij}$ the adjacency matrix.

be $f(i) := \frac{\delta^+(i)}{\delta^-(i)+\epsilon}$, where $\delta^+(i) = \sum_j W_{ij}$ and $\delta^-(i) = \sum_j W_{ji}$ are respectively the (weighted) out- and indegree of node $i$ and $\epsilon$ a small constant. The time complexity of this scheme is $O(n^2)$ for $n$ nodes.

In GI the node ordering is improved iteratively by the insertion of the node with current position $x$ in a new position $y$ (and shifting the other nodes accordingly). The change in the objective function is then given by: $\Delta E_{x\to y} = \text{sgn}(x-y)\sum_{z=x}^{y} W_{l(x)l(z)} - W_{l(z)l(x)}$. The new position $y$ is selected such that it yields the maximal possible gain $\Delta E_{x\to y}$, i.e. $y = \arg\max_z \Delta E_{x\to z}$. The order in which these elementary operations are applied to the list $l$ is a bit arbitrary, here we sweep over the list, where one sweep is taking $x = 1, 2, ..., n$ in that order. This sweeping procedure is repeated until $\forall x, y : \Delta E_{x\to y} \leq 0$. The time complexity of this scheme is $O(Sn^2)$, where $S$ is the number of sweeps needed to reach convergence, which is small in practice.

Finally, an SA variant of GI was tested, where the insertion is selected in a probabilistic manner in the following way: $p(x \to y) = \frac{e^{\Delta E_{x\to y}/T}}{\sum_z e^{\Delta E_{x\to z}/T}}$. Depending on the annealing schedule $T_t$, this optimization scheme will take a certain number of sweeps $S'$ to settle such that the time complexity is $O(S'n^2)$.

**Gene regulatory networks (GRNs)** We analyzed two GRNs. The first one represents directed edge probabilities between 3172 genes (nodes) obtained from genomic data from humans, let us call it graph $A$. The second one represents arc probabilities between 2884 yeast genes, and was obtained by applying the Findr software on data from [3]; let us name it graph $B$. The 3 presented algorithms were applied to both graphs. The results are presented in table 1. All presented algorithms were coded in Java 11.0.7 and the experiments were performed on a computer with an Intel Core i7-8650U CPU @ 1.90GHz×8 processor and 16 GB of RAM, under Ubuntu 18.04 x64. For graph $A$ the GI scheme converged after 18 sweeps, for $B$ convergence is reached after 30 sweeps.

Table 1: Objective and running time of the algorithms for the two GRNs.

| | $E_{\text{low}}$ $(= \frac{1}{2}\sum_{i,j} W_{ij})$ | $E_{\text{up}}$ $(= \sum_{i<j} \max(W_{ij}, W_{ji}))$ | $E_{\text{HS}}$ | $E_{\text{GI}}$ | $E_{\text{SA}}$ |
|---|---|---|---|---|---|
| $A$ | 78 649 | 146 210 | 139 358 | 139 938 | 140 002 |
| | - | - | *0.20 s* | *6.31 s* | *234 s* |
| $B$ | 1 180 397 | 1 829 665 | 1 602 905 | 1 621 242 | 1 622 023 |
| | - | - | *0.16 s* | *5.98 s* | *153 s* |

**Conclusion** The first and simplest scheme, HS, has been shown to be able to find a very good candidate solution. The solution can be improved even further by applying the GI scheme. We found that for the applications considered here it converges quickly and improves the result by about 1 % for graphs of sizes $n \sim 1000$. For the SA variation of this GI approach, there are a lot of parameters which can be tuned. We found that applying this improved the result only by a very small fraction, at the cost of much longer computation times. We could thus conclude that for practical purposes HS and GI have a good balance between the optimality of the obtained result and the computation time.

**Acknowledgements** *This research has been partially funded by the Ghent University BOF-project BioGraph and the UGent-project BOF/STA/202009/039.*

[3] Albert, Frank Wolfgang, et al. "Genetics of trans-regulatory variation in gene expression." Elife 7 (2018): e35471.

# On $K_2$-Hamiltonian Graphs

Jan Goedgebeur
KU Leuven Kulak, Department of Computer Science
e-mail: jan.goedgebeur@kuleuven.be
Jarne Renders
KU Leuven Kulak, Department of Computer Science
e-mail: jarne.renders@kuleuven.be
Gábor Wiener
Budapest University of Technology and Economics,
Faculty of Electrical Engineering and Informatics
e-mail: wiener@cs.bme.hu
Carol T. Zamfirescu
Ghent University,
Department of Applied Mathematics, Computer Science and Statistics
e-mail: carol.zamfirescu@ugent.be

Hamiltonian and hypohamiltonian graphs have been intensely studied since the 1960s [1, 3]. In this abstract we only consider connected simple graphs. A graph is *hamiltonian* if it contains a cycle which visits every vertex. Similarly, a graph is *hypohamiltonian* if it is not hamiltonian, but every vertex-deleted subgraph is. Motivated by a conjecture of Grünbaum and a problem of Katona, Kostochka, Pach and Stechkin, we investigate $K_2$-hamiltonian graphs based on earlier work by Zamfirescu [4].

Grünbaum [2] defined $\Gamma(j,k)$ with $k \geq j$ to be the family of graphs whose order and *circumference*, i.e., the length of a longest cycle, differ by $k$ and in which any $j$ vertices are missed by some longest cycle. $\Gamma(1,1)$ are then precisely the hypohamiltonian graphs. In 1974, Grünbaum conjectured that $\Gamma(j,j)$ must be empty for $j \geq 2$. So far little is known about the truth of this conjecture. We relax this problem further by investigating the following class of graphs.

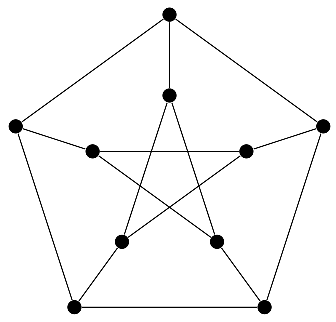

Figure 1: Petersen's graph: the smallest $K_2$-hypohamiltonian graph.

A graph is called $K_2$-*hamiltonian* if the deletion of any two adjacent vertices yields a hamiltonian graph. If this graph is also non-hamiltonian, we call it $K_2$-*hypohamiltonian*. An example of such a graph is Petersen's graph, which we show to be the smallest $K_2$-hypohamiltonian graph (Fig. 1).

We have designed and implemented an algorithm that checks whether a given input graph is $K_2$-(hypo)hamiltonian and which turns out to be quite efficient in practice. In particular, we use a backtracking algorithm with some heuristics to restrict the search space.

Using this algorithm together with mathematical operations preserving $K_2$-hypohamiltonicity, we were able to classify for which orders there exist $K_2$-(hypo)hamiltonian graphs for all orders except two. In the class of cubic graphs, i.e., graphs where every vertex has three incident edges, and *snarks*, i.e., cubic graphs with a 4-edge coloring, we completely characterize all orders and we have shown that for every order $n \geq 177$ there exists a $K_2$-hypohamiltonian planar graph. We have also found the smallest planar $K_2$-hypohamiltonian graph of *girth* 5, i.e., the length of a shortest cycle is 5, which has 48 vertices and the smallest cubic planar $K_2$-hypohamiltonian graph, which has 68 vertices (Fig. 2).

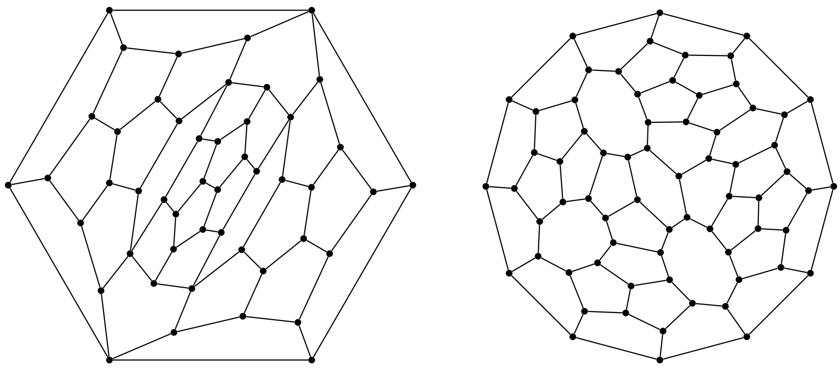

Figure 2: Left: The smallest planar $K_2$-hypohamiltonian graph of girth 5. Right: The smallest cubic planar $K_2$-hypohamiltonian graph.

# References

[1] R. J. Gould. Recent Advances on the hamiltonian Problem: Survey III. *Graphs and Combinatorics*, 30(1):1–46, 2014.

[2] B. Grünbaum. Vertices missed by longest paths or circuits. *Journal of Combinatorial Theory, Series A*, 17(1):31–38, 1974.

[3] D. A. Holton and J. Sheehan. *Hypohamiltonian graphs*, pages 214–248. Australian Mathematical Society Lecture Series. Cambridge University Press, 1993.

[4] C. T. Zamfirescu. $K_2$-hamiltonian graphs: I. *SIAM Journal on Discrete Mathematics*, 35(3):1706–1728, 2021.

# Exploring the beneficial impact of the LLM for credit scoring industry

K. Idbenjra
IÉSEG School of Management, IÉSEG Centre for Marketing,
LEM, UMR CNRS 9221
e-mail : k.idbenjra@ieseg.fr

K. Coussement
IÉSEG School of Management, IÉSEG Centre for Marketing,
LEM, UMR CNRS 9221
e-mail : k.coussement@ieseg.fr

A. De-Caigny
IÉSEG School of Management, IÉSEG Centre for Marketing
LEM, UMR CNRS 9221
e-mail : a.de-caigny@ieseg.fr

Despite advances in the credit scoring literature, most models lack interpretability. This requirement is crucial for the financial legislative regulators'. Therefore, simplistic logistic regression remains the golden standard in this field, mainly due to its simplicity and transparency. In this study, we introduce the Logit Leaf Model (LLM) as an alternative credit scoring algorithm. The LLM is a segmentation-based hybrid model that aims to improve the predictive power of the logistic regression model by segmenting the customer population, while keeping the predictions easily interpretable. The experiment benchmarks the LLM to the logistic regression and eight other popular classification techniques, using a real-world credit scoring data set of 65,536 active customers of a European financial institution. An extended visualization for LLM is also introduced in order to offer insights into the customers' drivers on a segment level. Empirical results show that the LLM is an encouraging solution for credit risk problems, since it competes with best performing approaches while offering deep insights vital for managerial decisions.

**Keywords :** Credit scoring, Interpretability, Logit Leaf Model, Hybrid segmentation-based model, Managerial decisions.

# A Proactive Inventory Policy based on Real-time degradation data

N. Al Khoury, D. Claeys, E. Aghezzaf
Universiteit Gent, Department of Industrial Systems Engineering and Product Design
Industrial Systems Engineering (ISyE), Flanders Make, www.FlandersMake.be
e-mail: naim.alkhoury@ugent.be,dieter.claeys@ugent.be
elhoussaine.aghezzaf@ugent.be

D. Fiems
Universiteit Gent, Department of Telecommunications and Information Processing
e-mail: dieter.fiems@ugent.be

Spare parts demand forecasting has been an interesting area of research over the past years. This area of research is getting increasing attention due to the enormous costs associated with managing and owning spare parts, while spare parts availability is critical in industries where any process blockage will cost from thousands up to millions of euros. To this extent, various spare parts forecasting techniques exist to decrease the stock level, without sacrificing the required service availability. These techniques mostly use historical demand data (time-series) to forecast future demand.
Lately, the fast technological development in sensors and communication technologies has paved the path towards real-time condition monitoring, which in turn enabled Condition-Based maintenance (CBM). In CBM, real-time condition information allows planners to anticipate future component failures, and maintenance actions are planned accordingly. A common practice is to implement a threshold policy where part replacement is triggered when degradation crosses a threshold.
The implementation of condition-based maintenance is in an advanced stage where major commercial corporations are putting enormous efforts into developing software to benefit from condition information data [1]. However, few researchers have considered exploiting condition information in spare parts decision-making [2]. Doing so could be beneficial because real-time condition data gives real-time information about future spare parts demand. Furthermore, in an inventory serving M machines sharing the same critical component, inventory pooling leads to higher fill rates, or one can achieve the same fill rate with fewer stock [3], [4].
We propose an inventory policy that exploits the condition information and pooling effect in spare parts decision-making. It anticipates the need for a spare to perform the maintenance action by ordering one when the degradation crosses an Order Threshold ($X_o$) smaller than the part replacement threshold. We hypothesize that the proposed policy will exploit both condition information and the pooling effect to reduce the average stock level. Simply put, the proposed

policy is a proactive variant of the base stock policy, where real-time condition information is used instead of only historical demand data.
The two decision variables driving this policy are the initial stock level (N) and the Order Threshold ($X_o$). We evaluate this policy using Discrete Event Simulation (DES), and we develop a custom algorithm that explores the search space intelligently to find the optimal parameters. We show that the proposed proactive policy reduces the average stock level by 15%-60% compared to the traditional base stock policy, with an average reduction of 35%. Another interesting result is that an optimal policy will adapt to a varying number of machines without a significant change in the Fill Rate, and thus provide the service providers with the flexibility needed in a changing market.

## References

[1] Lei, Y., Li, N., Guo, L., Li, N., Yan, T., & Lin, J. (2018). Machinery health prognostics: A systematic review from data acquisition to RUL prediction. Mechanical Systems and Signal Processing, 104, 799–834. https://doi.org/10.1016/j.ymssp.2017.11.016

[2] Van der Auweraer, S., Boute, R. N., & Syntetos, A. A. (2019). Forecasting spare part demand with installed base information: A review. International Journal of Forecasting, 35(1), 181–196. https://doi.org/10.1016/j.ijforecast.2018.09.002

[3] S. van Wijk (2012). Pooling & polling: creation of pooling in inventory and queueing models. PhD dissertation; Eindhoven University of Technology (supervisors: I. Adan and G. van Houtum).

[4] F. Karsten, R.J.I. Basten (2014). Pooling of spare parts between multiple users: how to share the benefits. European Journal of Operational Research 223: 94-104.

# The effect of airline dominance on airport performance: empirical evidence from medium-sized European airports

Simon Weekx
Department of Business Technology and Operations, Vrije Universiteit Brussel
Department of Transport and Regional Economics, University of Antwerp
e-mail: simon.weekx@vub.be

Sven Buyle
Department of Transport and Regional Economics, University of Antwerp
e-mail: sven.buyle@uantwerpen.be

The Air Transport Industry is a competitive and volatile market, creating a challenging operating environment for both airlines and airports. With respect to the airline market, various market events are causing continuous changes in the airline's market structure. These include the liberalization of the airline market industry, the introduction of the low-cost business model, and more recently, the COVID-19 pandemic. With respect to the airport market, the competition with more sustainable transport modes (e.g., high-speed railways), the rise of airport congestion and the fierce competition with other airports are forcing airports to make more efficient use of their scarce infrastructure.

Given the vertical relationship in which airports and airlines cooperate [1, 2], the question arises how airline dominance can have an impact on the airport's technical efficiency. Existing research is still lacking on this matter, and from a theoretical point of view, arguments for both a positive and negative relation between airline dominance and airport efficiency can be formulated.

Therefore, in this research we empirically analyse this relationship based on a panel dataset of 20 European airports observed from 2012-2018. The two-staged double bootstrapped truncated regression model as proposed in [3] is applied to estimate the effect of airline dominance on the airport's technical efficiency. In the first stage, Data Envelopment Analysis (DEA) is used to estimate the efficiency score of each airport. However, in the conventional DEA approach, efficiency estimates are too optimistic because efficiency is measured relative to the estimated frontier and not the true frontier. Thus, [3] propose to apply bootstrapping to calculate bias-corrected (BC) technical efficiency. In the second stage, these BC efficiency scores are used in a truncated regression model and then enter another bootstrapping procedure to estimate standard errors and confidence intervals. We furthermore identify BC input and output slacks, allowing to calculate the input and output-increasing potentials of the airports in the sample.

Our model proves a significant positive relationship between airline dominance and the airport's technical efficiency. We also find that this relationship is not dependant on the operating strategy of the dominant carrier (i.e., being a Full-Service Carrier or Low-Cost Carrier). Furthermore, policy-makers are advised to take possible improvements in technical efficiency at medium-sized airports and the related effects on consumer prices into consideration when reviewing airline consolidation cases. Medium-sized regional airports owned by regional governments and typically dominated by a LCC should not necessarily introduce a second dominant carrier. Doing so might decrease the airports efficiency, possibly increasing costs to consumers or taxpayers.

**Acknowledgement** This research contains enhanced contents from S.W.'s Master's Thesis [4]. Furthermore, S.W. acknowledges a fellowship from the Research Foundation Flanders (FWO-Vlaanderen) under Grant No. 11I5122N.

## References

[1] H.-K. Ha, Y. Wan, Y. Yoshida, and A. Zhang. Airline market structure and airport efficiency: Evidence from major northeast asian airports. *Journal of Air Transport Management*, 33:32–42, 2013.

[2] S. Hidalgo-Gallego and I. Mateo-Mantecón. Effect of concentration in airline market on spanish airport technical efficiency. *Journal of Air Transport Management*, 76:56–66, 2019.

[3] L. Simar and P. W. Wilson. Estimation and inference in two-stage, semi-parametric models of production processes. *Journal of Econometrics*, 136(1):31–64, 2007.

[4] S. Weekx. Empirisch onderzoek naar het effect van dominante luchtvaartmaatschappijen op de technische efficiëntie van luchthavens. Master's thesis, University of Antwerp, 2020.

# The integrated on-demand bus routing problem: combining on-demand buses with a high-frequency metro-network

Lissa Melis
University of Antwerp, Engineering Management Department - ANT/OR
e-mail: lissa.melis@uantwerpen.be

Michell Queiroz
University of Antwerp, Engineering Management Department - ANT/OR
email: michell.queiroz@uantwerpen.be

Kenneth Sörensen
University of Antwerp, Engineering Management Department - ANT/OR
email: kenneth.sorensen@uantwerpen.be

Considerable high travel times and accessibility issues, are some of the reasons for discontent with the current operation of traditional public transportation. Consequently, people are increasingly drawn to use private cars, which provide more convenience and flexibility. The rise in number of vehicles in the streets inevitably leads to many negative effects, such as traffic congestion and increased emissions. Integrated transit systems, which include traditional, fixed route public transportation and demand responsive transport, hold the potential to address these problems. The first consists of a set of predefined fixed routes and schedules and is operated by (trolley)buses, trams, trains and subways. The second adapts its routes and schedules according to bookings and is usually operated by a fleet of minibuses.

In this work we analyse the performance of integrating a large-scale on-demand bus system with a high-frequency fixed line public transport network in an urban context. Given are a high-speed metro network, a set of bus station locations, a fleet of fixed capacity minibuses and a set of real-time requests. Users issue a transportation request informing their departure and arrival locations, of which we compile a set of possible departure/arrival bus stations within walking distance of these locations. A time window consisting of an earliest departure time and a latest arrival time is also specified. The system responds to the user with an itinerary as a proposal that fits the time window constraints.

Passengers can have five possible trip types: (1) only bus, (2) bus + metro, (3) metro + bus, (4) bus + metro + bus, or only metro. Note that when the itinerary has a metro-leg, it is possible for passengers to have to switch lines within the metro network to reach their assigned transfer or destination station. These transfers *within the metro network* are presumed to be highly efficient and unlimited in number.

The aim of the solution algorithm for this problem is to simultaneously (1) decide on the trip type for each passenger (only bus, metro or mixed), (2) route the on-demand buses, (3) assign each passenger to a departure and arrival bus station (bus station assignment), and (4) in the case of a metro-leg in the trip, decide the assigned transfer station(s) and used metro lines (transfer station assignment). We call this problem *the integrated on-demand bus routing problem or I-ODBRP*.

We propose a quick and scalable insertion-based heuristic to solve the I-ODBRP. We smartly assign passengers to buses and to pieces of the metro network. The results found by the heuristic are further used to compare the performance of an integrated system, to a system that only uses on-demand buses. It is concluded that the integrated system always performs better regarding the service rate or number of served requests. Depending on the speed and layout of the metro network, also the average user ride time per passenger improves by the integration.

Lastly, we present a case study of the city of Lisbon, Portugal, to investigate the integration of on-demand buses with the existing fixed public transport network.

# Towards better service quality with the dynamic feeder service with a maximum headway at mandatory stops

B.D. Galarza Montenegro
University of Antwerp, Department of Engineering Management (ENM)
e-mail: bryan.galarzamontenegro@uantwerpen.be

K. Sörensen
University of Antwerp, Department of Engineering Management (ENM)
e-mail: kenneth.sorensen@uantwerpen.be

P. Vansteenwegen
KU Leuven, KU Leuven Institute for Mobility - CIB
e-mail: pieter.vansteenwegen@kuleuven.be

Public transport enables accessibility to social activities, goods and services, which reduces social exclusion and poverty [1]. For secluded areas in particular, like residential areas and suburbs, a feeder service enables connectivity to major transit networks. There are two types of public transportation services: traditional transport services (TTS) and demand-responsive transportation services (DRFS). TTS are composed of a set of bus lines that follow predefined routes and schedules. A crucial advantage of TTS is their low operational costs and their predictable nature which makes them highly accessible to most commuters. The inflexibility of TTS is a limitation as well, since it renders TTS inadequate in settings where demand for transportation is sparse and constantly changing. DRFS are better able to cope with an ever-changing demand for transportation; these services operate only when there is demand for transportation and meet the passenger's expectations accordingly [2]. DRFS can be further divided into two groups: fully flexible DRFS and semi-flexible DRFS [3]. Fully flexible DRFS are services where all aspects are dependent on the demand for transportation. Semi-flexible DRFS are flexible services that have one or more predetermined components, such as (partially) fixed routes and/or timetables, and offer a middle-ground between TTS and fully flexible DRTS. As such, this research focuses on a semi-flexible DRFS, integrating the positive characteristics from both fully flexible DRFS and TTS.

The feeder service considered in this paper is denoted as the Feeder Service with Mandatory Stops (FSMS). The FSMS serves two sets of bus stops: mandatory stops and clustered optional stops. The mandatory stops are placed along the main route, typically a highway, and are visited by a bus within a maximum allowable headway, i.e., these stops are visited by a bus within a certain time interval. For example, a maximum headway of 20 minutes means that passengers never have to wait longer than 20 minutes at a mandatory stop. It should be noted that this is more strict than having three buses per hour. The optional stops are only visited by a bus when a client nearby makes a request for transportation. These stops are scattered across several residential areas, which are close to the main route. Potential passengers make a request for transportation by stating their current location and their desired arrival time at the destination or departure time from their origin location. Passengers are assigned to a departure stop and a bus. Afterwards, they are informed of their pick-up time and are expected to walk to their assigned bus stop in order to board a bus. This means that the routes and the timetables of the buses are not fixed but change according to the demand. The mandatory stops provide some predictability and serve as a safety net for the "passengers without reservation" because these passengers can still board a bus at the mandatory stops. Furthermore, passengers at mandatory stops do not have to wait more than a certain amount of time because at least one bus departs from these stops within a certain time interval. The use of fixed transit stops is a driving factor in transit-oriented urban development [4, 5], which makes the use of the mandatory stops more attractive as well. Buses that arrive at the hub are expected to return to be reused for subsequent trips, which means that fewer buses are required to meet demand. It is found that attributes such as comfort and arrival time at the destination, but most importantly the reliability of a service, are highly valued by passengers [6]. These attributes are present in the FSMS.

To optimize the performance of the service, we develop a two-phase heuristic. The heuristic re-optimizes the service whenever new requests are received. In the first phase, an insertion algorithm adds the new incoming passenger request to the existing planning. In the second phase, we implement an iterative greedy construction heuristic to improve the current solution, if possible. In each iteration of the improvement heuristic, a complete solution is reconstructed, while respecting all system constraints and additional real-time restrictions. Since the construction of solutions is greedy, infeasible solutions are possible. These infeasible solutions are discarded. To bring more variability into the reconstruction of the solutions, construction parameters are introduced. These parameters are (semi-)randomly sampled in each iteration. Consequently, different solutions are generated in each iteration. These parameter bring a balance between the greediness and the feasibility of the construction and allow us to obtain better results, as well as a larger number more feasible solutions. The heuristic stops when a certain number of feasible solutions are generated.

The real-time, dynamic optimization method is compared to a previously developed heuristic that optimized the service in a static manner, i.e., when all requests are known beforehand and the planning is fixed after a certain deadline. It is then concluded, through preliminary results, that the dynamic method performs quite well since the difference in objective function values between the dynamic and the static model is not large.

## Acknowledgments

This project was supported by the FWO (Research Foundation Flanders) project G.0759.19N.

## References

[1] J. Hine and F. Mitchell, “Better for everyone? Travel experiences and transport exclusion”, Urban Studies, vol. 38, no. 2, pp. 319–332, 2001.

[2] M. J. Alonso-González, T. Liu, O. Cats, N. Van Oort, and S. Hoogendoorn, “The Potential of Demand-Responsive Transport as a Complement to Public Transport: An Assessment Framework and an Empirical Evaluation”, Transportation Research Record, vol. 2672, no. 8, pp. 879–889, 2018.

[3] P. Vansteenwegen, L. Melis, D. Aktaş, B. D. Galarza Montenegro, F. Veiera, and K. Sörensen, “A survey on demand-responsive public bus systems”, Transportation Research Part C: Emerging Technologies, vol. 137, p. 103573, 2022.

[4] R. D. Knowles, F. Ferbrache, A. Nikitas,“ Transport’s historical, contemporary and future role in shaping urban development: Re-evaluating transit oriented development", Cities, 99, p. 102607, 2020.

[5] A. Ibraeva, G. Homem, A. Correia, C. Silva and A. Pais Antunes, “Transit-oriented development: A review of research achievements and challenges", Transportation Research Part A: Policy and Practice, 132, p. 110–130, 2020.

[6] G. Beirão, J. A. Sarsfield Cabral, “Understanding attitudes towards public transport and private car: A qualitative study", Transport Policy 14, 6, p. 478–489, 2007.

# Speeding up the Passenger Assignment Problem in public transport network design

Pieter Vansteenwegen
KU Leuven, Institute for Mobility – CIB
e-mail: pieter.vansteenwegen@kuleuven.be

Dilay Aktas
KU Leuven, Institute for Mobility – CIB
e-mail: dilay.aktas@kuleuven.be

Evert Vermeir
De Lijn, Flemish public bus company
e-mail: evert.vermeir@delijn.com

When designing a public transport network with the passenger travel time in mind, the *passenger assignment problem (PAP)* needs to be solved and explicit passenger routing has to be considered. In the PAP, expected decisions of passengers are aimed to be modelled and optimized. Since it is a complex problem on its own, public transport planners typically treat the PAP as a subproblem of larger problems, such as line planning problem (LP) (Durán-Micco and Vansteenwegen, 2022).

In LP, typically all passengers are assigned to their shortest paths to minimize their in-vehicle time and transfer penalties i.e. all-or-nothing transit assignment. Several other assumptions are made sometimes to further simplfy the problem such as ignoring the transfers or operators assigning routes to the passengers. However, in order to avoid unrealistic assumptions, LP is rather solved as a bi-level optimization problem where the aim is to first come up with a line plan for the given network, and then to solve the PAP for the given line plan (Vermeir et al., 2021; Durán-Micco and Vansteenwegen, 2022).

To solve the all-or-nothing PAP, the network representation has to be properly adjusted when transfers are considered. A typical method used in liturature is to add a dummy node for every bus stop on every bus line. This extended network is then called 'Change and Go' (CNG) (Schöbel and Scholl, 2006) or 'Train Service Network' (TSN) (Liu et al., 2020). Therefore, the number of nodes significantly increases especially when the line plan includes many lines with many transfer possibilities. Thus, the computation time required to solve PAP also increases as the complexity of shortest path algorithms (e.g. Dijkstra, Floyd Warshall) mainly depends on the number of nodes. According to Farahani et al. (2013), solving the PAP is often considered to be the most time-consuming part of solving LP. Moreover, since many (meta-)heuristics addressing the LP require to solve this PAP in every iteration, a time-consuming PAP significantly increases the computation time of these (meta-)heuristics (Durán-Micco and Vansteenwegen, 2022).

In this study, a more efficient "Direct Link Network" (DLN) representation for the uncapacitated line planning problem is presented. The new network representation adds a link between a pair of nodes only if there exists a *direct link*, i.e., a single bus line that connects those two stops. The travel time or weight of the link corresponds to the shortest travel time, without taking a transfer, between the two nodes. Thus, for any origin and destination pair, if there does not exists a 'direct link', this means that the passengers need to make a transfer. By adding a transfer penalty to the weight of each link in the DLN and then subtracting one from each shortest path determined, all 'real' transfers are automatically considered without any additional transfer nodes.

According to the experimental results, DLN representation significantly fastens the passenger assignment for the most commonly used networks of Mandl's Swiss Network and Mumford Networks. For the largest Mumford3 network with 127 stops and a line plan of 60 lines, the PAP is solved at least 10 times faster when DLN is used compared to when CNG is used. The fast evaluation of passenger assignment promises a significant improvement in the computational efficiency of the algorithms that solve the line planning problem with passenger travel time as an objective.

## References

Durán-Micco, J. and Vansteenwegen, P. (2022). A survey on the transit network design and frequency setting problem. *Public Transport*, 14(1):155–190.

Farahani, R. Z., Miandoabchi, E., Szeto, W. Y., and Rashidi, H. (2013). A review of urban transportation network design problems. *European Journal of Operational Research*, 229(2):281–302.

Liu, D., Durán Micco, J., Lu, G., Peng, Q., Ning, J., and Vansteenwegen, P. (2020). A matheuristic iterative approach for profit-oriented line planning applied to the chinese high-speed railway network. *Journal of Advanced Transportation*, 2020.

Schöbel, A. and Scholl, S. (2006). Line planning with minimal traveling time. In *5th Workshop on Algorithmic Methods and Models for Optimization of Railways (ATMOS'05)*. Schloss Dagstuhl-Leibniz-Zentrum für Informatik.

Vermeir, E., Engelen, W., Philips, J., and Vansteenwegen, P. (2021). An exact solution approach for the bus line planning problem with integrated passenger routing. *Journal of Advanced Transportation*, 2021.

# Scheduling non-professional leagues based on a Swiss tournament system

Miao Li
e-mail: miao.li@ugent.be

Dries Goossens
e-mail: dries.goossens@ugent.be

David Van Bulck
e-mail: david.vanbulck@ugent.be

Ghent University, Department of Business Informatics and Operations Management

## 1 Introduction

Scheduling non-professional (youth) sports competitions is extremely daunting due to the large number of teams and matches typically involved. For instance, in the U15 youth football series of Easter Flanders alone, the Royal Belgian Football Association organizes approximately 1,500 matches each year for about 200 teams from over a hundred clubs. To deal with problem instances of such size, a very common two-phase approach is to first group teams into leagues, and then have each league play a round-robin tournament (i.e., each participant plays a fixed number of times against all other participants in the same league). Often, the objective in the first phase is thereby to minimize pairwise distances of the teams grouped into the same league (see e.g. [1]), whereas the second phase is to respect the venue capacity of the clubs as much as possible.

To better cope with the above challenging issues, we suggest to organize non-professional competitions based on a Swiss tournament system. Originally designed for chess tournaments, the Swiss tournament is nowadays a popular tournament format in many professional and amateur tournaments such as badminton and bridge. Similar to round-robin tournaments, the Swiss system is a non-eliminating tournament format that features a predefined number of rounds. However, the pairing of players for a round in the Swiss system depends on their results against each other in earlier rounds. Hence, at the end of the season, all teams will have played the same number of matches, but probably against a different set of opponents. The contribution of this work is to show how a Swiss-system based approach can be a very effective way compared with the traditional approach to optimize travel distance while respecting the venue capacities of the clubs.

## 2 Problem description and proposed method

We propose an alternative two-phase decomposition approach to organize the competition. The input of the first phase consists of a set of clubs $C$, a set of teams $T$, a set of rounds $R$, and a set of home-away-patterns (HAP) $\mathcal{H}$ indicating for each round whether a team plays a home game (H) or an away game (A). An integer programming (IP) model is used to assign a suitable HAP to each team. The objective is to minimize the venue capacity violations over all clubs and all rounds. Capacity violations happen whenever the number of teams from the same club simultaneously playing at home exceeds the number of venues of that club [2]. We define binary decision variables $x_{t,h}$ that are equal to one if team $t \in T$ is assigned to HAP $h \in \mathcal{H}$ and zero otherwise. The constraints can be summarized as follows:

(C1) each team $t \in T$ is assigned to a HAP $h \in \mathcal{H}$.

Given an assignment of teams to HAPs, the second phase attempts to find opponents for all teams, which gives a schedule or proves that no such schedule exists (in which case a new HAP assignment is generated). This phase aims at minimizing the travel distance of the teams, while only pairing teams with more or less similar strength at the beginning of the season. The set of eligible opponents for team $t$ is denoted by $O_t$. The problem is again formulated as an IP model. We introduce binary decision variables $y_{t,p,r}$ which are equal to one if team $t \in T$ plays against team $p \in O_t$ at home on round $r \in R$, and zero otherwise. The constraints taken into account are as follows:

(C2) each team $t \in T$ plays exactly once on each round $r \in R$,

(C3) each team $t \in T$ plays at most one home game against each other team $p \in O_t$ during the season,

(C4) each team $t \in T$ meets each other team $p \in O_t$ at most once in each half of the season,

(C5) each team $t \in T$ plays according to the HAP assigned to that team during the first phase.

As the second stage is computationally challenging to solve, we propose a relax-and fix algorithm, combining heuristics and exact mathematical programming techniques to obtain a good solution. More in particular, the proposed algorithm constructs a schedule by maintaining a rolling horizon of time intervals (i.e., consecutive rounds), in which the integrality constraints of all $y_{t,p,r}$ variables within the current time interval are activated, while relaxing all others. At the end of each iteration, the $y_{t,p,r}$ variables of the last considered time interval are fixed, and the algorithm moves on to the next interval.

# References

[1] Toffolo, T.A., and Christiaens, J., and Spieksma, F.C., and Vanden Berghe, G. (2019) The sport teams grouping problem. *Annals of Operations Research*, 275(1), 223–243.

[2] Davari, M, and Goossens, D., and Beliën, J., and Lambers, R., and Spieksma, F.C. (2020) The multi-league sports scheduling problem, or how to schedule thousands of matches. *Operations Research Letters*, 48(2), 180–187.

# A traditional Benders' approach to sports timetabling

D. Van Bulck, D. Goossens
Ghent University, Department of Business Informatics and Operations Management
e-mail: david.vanbulck@ugent.be

## 1 Introduction

In a double round-robin (2RR) sports timetable, each team plays against every other team once at home and once away. As the construction of a fair and practical sports timetable turns out to be extremely challenging, most approaches from the literature decompose the problem via the so-called first-break-then-schedule (FBTS) approach: first decide when teams play home or away (the so-called home-away pattern or HAP set), then determine a compatible assignment of opponents (see e.g. [1]). Although the efficiency of FBTS heavily depends on the ability to avoid and recover from infeasibility in the second phase, research on this topic is scarce. The contribution of this work is to show how classical Benders' decomposition with integer subproblems relates to FBTS, and how Benders' cuts can be used to organize backtracking. Compared to the classic FBTS approach, our approach is exact and offers more flexibility to cope with real-life constraints and can even deal with situations where the objective is (partly) determined by the assignment of opponents.

## 2 Problem description and proposed approach

The aim of this section is to show how to use Benders' decomposition to construct a so-called compact 2RR timetable, where there are $2n$ teams and $4n-2$ rounds during which the games can take place (denoted by sets $T$ and $R$, respectively). To emphasize the generality of the approach, this section does not assume any particular side constraints (e.g. at most three home games in a row) nor objective but instead assumes that the objective is to minimize $z = z_1 + z_2$. We further assume that $z_1$ only depends on the HAP set associated with the timetable (e.g. the minimization of consecutive home or consecutive away games), and that $z_2$ (linearly) depends on the assignment of opponents (e.g. the minimization of costs related to the assignment of games). With only very few exceptions, the vast majority of sports timetabling problems previously presented in the literature adhere to this assumption.

In order to formulate this problem, let binary variables $h_{i,r}$ indicate whether team $i \in T$ plays home in round $r \in R$ and $x_{i,j,r}$ whether $i$ plays at home against $j \in T \setminus \{i\}$ in $r$.

$$\begin{aligned}
&\mathbf{min} \;\; z_1 + z_2 && (1)\\
&z_1 = \mathbf{a}^\mathsf{T}\mathbf{h} && (2)\\
&z_2 = \mathbf{b}^\mathsf{T}\mathbf{x} && (3)\\
&\sum_{r\in R} h_{i,r} = 2n-1 & \forall i \in T \;\; & (4)\\
&\sum_{i\in T} h_{i,r} = n & \forall r \in R \;\; & (5)\\
&\sum_{j\in T\setminus\{i\}} x_{i,j,r} = h_{i,r} & \forall\, i \in T, \forall\, r \in R \;\; & (6)\\
&\sum_{j\in T\setminus\{i\}} x_{j,i,r} = 1 - h_{i,r} & \forall\, i \in T, \forall\, r \in R \;\; & (7)\\
&\sum_{r\in R} x_{i,j,r} = 1 & \forall i, j \in T : i \neq j \;\; & (8)\\
&h_{i,r} \in \{0,1\}, x_{i,j,r} \in \{0,1\} & \forall i, j \in T : i \neq j, \forall\, r \in R \;\; & (9)
\end{aligned}$$

The objective is to minimize the sum of $z_1$ and $z_2$, whose value is regulated by Constraints (2) and (3), in which **a** and **b** are application dependent cost vectors and **h** and **x** are the vectors of all $h_{i,r}$ and $x_{i,j,r}$ variables. Constraints (4) state that each team plays half of its games at home, and Constraints (5) state that exactly half of the teams in each time slot play home. Constraints (6) and (7) require that all teams play according to their HAP, and Constraints (8) ensure that all games are scheduled. Finally, constraints (9) are the binary constraints.

Instead of solving the above formulation at once, we propose to apply Benders' decomposition by projecting out the complicating $x_{i,j,r}$ variables. The master problem then corresponds to constructing a HAP set, whereas the subproblem is to find a compatible assignment of opponents or to prove that none exists. As the $x_{i,j,r}$ variables need to be integral, we complement the traditional Benders' feasibility and optimality cuts with no-good and integer optimality cuts.

Among other applications, we will show how to apply the proposed approach to the constant-distance travelling tournament problem (C-TTP), which aims to minimize the total number of times the teams have to travel from one venue to another. More in particular, Urrutia and Ribeiro [2] show that the total number of travel trips can be minimized by maximizing the total number of breaks (i.e. consecutive home or consecutive away games). In order to generate trip-minimal timetables, we apply the proposed Benders' approach by first constructing a break-maximal HAP set and then finding a compatible assignment of opponents. Although the C-TTP is a well researched problem and despite the Benders' approach being broadly applicable, the proposed algorithm is competitive with several tailored-made ones and even finds some new best solutions.

## References

[1] G. L. Nemhauser and M. A. Trick, *Scheduling a major college basketball conference*, Oper. Res., 46 (1998), pp. 1–8.

[2] S. Urrutia and C. C. Ribeiro, *Maximizing breaks and bounding solutions to the mirrored traveling tournament problem*, Discrete Appl. Math., 154 (2006), pp. 1932–1938.

# Optimizing invigilator assignments for higher education examination periods

Nico André Schmid[a,b], Veronique Limère[b,c]
[a]IESEG School of Management, Univ. Lille, CNRS,
UMR 9221 - LEM - Lille Economie Management, F-59000 Lille, France
[b]Ghent University, Department of Business Informatics and Operations Management
Ghent, Belgium
[c]Flanders Make@Ghent University — core lab CVAMO, Ghent, Belgium
e-mails: n.schmid@ieseg.fr

Higher education institutions typically utilize assessments to measure their students' progress and to assess whether they obtain the required competences to proceed with their studies and ultimately receive a degree. Due to the ever-increasing number of students and the variety of (elective) courses, the execution of examinations needs to be carefully orchestrated. Avoiding students having to take multiple exams simultaneously, considering the availability of room capacities [2] and ideally spreading the exams over the examination period such that students have some time to study in between different exams [6] are aspects considered and optimized in the well-established research domain of examination timetabling ([1, 5] provide excellent surveys on this problem). After this, all exams need to be staffed with sufficient invigilators. This staffing problem may be considered as part of the general examination timetabling problem. However, to this day, only a few studies are concerned with this problem [3, 4]. Furthermore, the COVID-19 pandemic has complicated this problem. Many countries impose social distancing measures, reducing the number of students allowed to take an exam in the same room. Therefore, more rooms and invigilators are needed than before. Thus, this study is concerned with the staffing of invigilators to exams while considering their availability and preferences. Moreover, we explore the opportunity of reducing contacts during pandemic situations to limit the spread of highly-contagious diseases such as COVID-19.

Specifically, we investigate the invigilator assignment problem and propose some problem variations along with mathematical programming formulations. $(i)$ We propose a model that assigns invigilators to exams based on their preferences for specific time slots. That is, each invigilator is provided with a list of time slots and indicates its preference for these time slots, where high priorities have low values, and low priorities have high values. The mathematical program minimizes the sum of all preference values while ensuring sufficient staffing of all exams. Due to some practical constraints such as equal distribution of supervisions and non-availability of invigilators, some exams might not be fully staffed. Therefore, a penalty cost is induced in such cases. The penalty cost can be parameterized such that the problem is optimized lexicographically. First, as many exams as possible are staffed, and secondly, the sum of preferences is optimized. This model can be solved very efficiently. (*ii*) Due to the contagious nature of COVID-19 and other diseases, one may be interested in reducing the number of contacts among all parties involved, i.e., invigilators and students. We present a first model that minimizes the contacts between invigilators as they are likely to interact more closely during the examination supervision. Due to the intractable nature of this problem, we propose a heuristic approach based on a large neighborhood search (LNS). For this, a starting solution is created based on the model discussed in (*i*) and iteratively improved by destroying the invigilator assignments of either a set of exams and/or a set of time slots and fixing it by solving a mixed-integer linear programming (MILP) problem for this neighborhood only. We will discuss and demonstrate how these models can be complemented with multiple practical considerations and extensions.

The problem variants and some practical constraints have been applied to the invigilator assignment of a large faculty of a Belgian university. We will show the results of this planning model and discuss practical considerations that arose during the implementation over multiple examination periods.

## References


[1] Carter, M. W. (1986). A survey of practical applications of examination timetabling algorithms. *Operations Research*, 34(2): 193-202.

[2] Dimopoulou, M.. Miliotis, P. (2001). Implementation of a university course and examination timetabling system *European Journal of Operational Research*, 130: 202–213.

[3] Kahar, M.N.M., Kendall, G. (2014). Universiti Malaysia Pahang examination timetabling problem: scheduling invigilators *The Journal of the Operational Research Society*, 65(2): 214-226.

[4] Matci, D.K., Acar, I. (2019). Modelling and optimization of the exam invigilator assignment problem based on preferences *Turkish Journal of Engineering*, 3(3): 140-148.

[5] Qu, R., Burke, E.K., McCollum, B., Merlot, L.T.G., Lee, S.Y. (2009). A survey of search methodologies and automated system development for examination timetabling. *Journal of Scheduling*, 2009(12): 55-89.

[6] Woumans, G.. De Boeck L., Belien, J., Creemers, S. (2016). A column generation approach for solving the examination-timetabling problem *European Journal of Operational Research*, 253: 178-194.

# A Transparent Data Persistence Architecture for the SimJulia Framework

Piet Van Der Paelt
Department of Mathematics
Faculty of Engineering
Royal Military Academy
e-mail: piet.vanderpaelt@mil.be

Ben Lauwens
Department of Mathematics
Faculty of Engineering
Royal Military Academy
e-mail: ben.lauwens@mil.be

Beat Signer
Web & Information Systems Engineering Lab
Department of Computer Science
Vrije Universiteit Brussel
e-mail: beat.signer@vub.be

SimJulia is an open source discrete event simulation framework [1] written in the Julia programming language [2]. The framework transforms processes expressed as functions into resumable functions using the Resumable Functions library [3]. The current SimJulia implementation lacks functionality to persist any variables, characterising the state of such a transformed process during the simulation in a transparent way. To mitigate this shortcoming, we extended both the SimJulia framework and the Resumable Functions library by implementing a transparent probing and persistence architecture, employing language-specific metaprogramming features. Our implementation is based on the Object-Relational Mapping concept (ORM) [4] using the PostgresORM library [5], supported by the PostgreSQL Relational Database Management Systems (RDBMS), and Julia's macro expansion.

A simulation model expressed by means of functions is transformed into resumable functions ready to be run by SimJulia. Our solution uses the code analysis occurring at this stage to store a monitored function's state variables configuration in the database. The user controls which processes are monitored through a flag set on each process. After the code analysis phase, macro expansion takes place, creating the object definitions based upon the configuration saved earlier, before the simulation runs. During the simulation, successive events invoke an array of callback functions being executed. The last function in such an array is the probing function. Monitored processes are probed using the object definitions created just in time. Instantiated objects materialising the state of the concerned process are then persisted in the table specific to the concerned process. The data is subsequently available in the RDMBS, and can be visualised using the technology of choice. Our implementation visualises the data using the VueJS library.

Our contribution consists of a transparent probing and persistence mechanism. Using Julia's metaprogramming capabilities, we were able to divert from a static ORM configuration as often seen in web applications, to a transparent and dynamic configuration. The end user has the evolution of the state variables available throughout the simulation, without having to provide any configuration.

## References

[1] B. Lauwens, "Simjulia: Discrete Event Process Oriented Simulation Framework Written in Julia." `https://simjuliajl.readthedocs.io/en/stable/welcome.html`, 2017. Online; accessed 28-June-2022.

[2] J. Bezanson, A. Edelman, S. Karpinski, and V. B. Shah, "Julia: A Fresh Approach to Numerical Computing," *SIAM review*, vol. 59, no. 1, pp. 65–98, 2017.

[3] B. Lauwens, "ResumableFunctions: C# Sharp Style Generators for Julia," *Journal of Open Source Software*, vol. 2, no. 18, p. 400, 2017.

[4] C. Russell, "Bridging the Object-Relational Divide: ORM Technologies Can Simplify Data Access, But Be Aware Of The Challenges That Come With Introducing This New Layer Of Abstraction," *Queue*, vol. 6, no. 3, pp. 18–28, 2008.

[5] V. Laugier, "PostgresORM.jl: Object Relational Mapping for PostgreSQL." `https://discourse.julialang.org/t/ann-postgresorm-jl-object-relational-mapping-for-postgresql/63847`, 2021. Online; accessed 28-June-2022.

# NLP + network analysis for monitoring social networks: SOCMINT toolbox prototype

J. C. Fernandez-Toledano
Royal Military Academy, Department of Mathematics
e-mail: juan.fernandeztoledano@mil.be

B. De Clerck
Royal Military Academy, Department of Mathematics
Ghent University, Department of Economics

F. Van Utterbeeck
Royal Military Academy, Department of Mathematics

Over the past years, social media has become a powerful tool for communication, education, and sharing information in real-time. However it can also constitute a threat to national security when used to spread misinformation, propaganda, and hate speech. Extremists, both individuals and groups, can use these platforms to increase their presence and their number of sympathizers and to disseminate their messages of hate more broadly [2]. For this reason, it is desirable to develop social media analytics tools, able to pull data from social media channels and to perform sophisticated analyses. Objectives include gaining insight into the current situation, finding opportunities to solve crisis problems, stopping misinformation campaigns and predicting potential upcoming topics and events.

In this work, we present a prototype of a React-based web application toolbox able to combine modern natural language processing (NLP) tools to perform content analysis, artificial intelligence techniques to convert image attachments to captions, and efficient topological analysis of the interaction network created between the different actors or topics on a given social network. NLP techniques as state-of-the-art pre-trained transformers [4] can be used for automatic text translation, to extract key information such as sentiments, topics, named entity recognition (locations, names or corporations) or to do an automatic summarization of particular topics [5]. Then, this information together with the standard metrics provided by the different social network platforms (eg. number of user's followers or likes of a message) can be used to filter the data stored in a database to create the network of interest using the efficient Rapids cugraph library [3]. This library allows the management networks of millions of edges [1] based on parallel GPUs which ensures its scalability. Moreover, the SOCMINT toolbox allows easy navigation between the network and messages as well as the profiling of the users and the different tags. In the current phase, the SOCMINT toolbox operates on the Twitter platform but it will be extended to other social media in the future.

## References

[1] T. Hricik, D. Bader, and O. Green. Using rapids ai to accelerate graph data science workflows. In *2020 IEEE High Performance Extreme Computing Conference, HPEC 2020*, 2020.

[2] J. Shin, L. Jian, K. Driscoll, and F. Bar. The diffusion of misinformation on social media: Temporal pattern, message, and source. *Computers in Human Behavior*, 83:278–287, 2018.

[3] R. Team et al. Rapids: Collection of libraries for end to end gpu data science, 2018, 2022. `https://rapids.ai`.

[4] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin. Attention is all you need. In *Advances in Neural Information Processing Systems (NIPS 2017*, volume 30, Long Beach, CA, USA, 2017.

[5] T. Wolf, L. Debut, V. Sanh, J. Chaumond, C. Delangue, A. Moi, P. Cistac, T. Rault, R. Louf, M. Funtowicz, J. Davison, S. Shleifer, P. von Platen, C. Ma, Y. Jernite, J. Plu, C. Xu, T. L. Scao, S. Gugger, M. Drame, Q. Lhoest, and A. Rush. Transformers: State-of-the-art natural language processing. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, pages 38–45, 2020.

# Neighborhood Enumeration in Local Search Metaheuristics

Michiel Van Lancker
KU Leuven, CODeS
e-mail: michiel.vanlancker@cs.kuleuven.be

Greet Vanden Berghe
KU Leuven, CODeS

Tony Wauters
KU Leuven, CODeS

The definition and implementation of neighborhoods is an essential aspect of the design of local search-based metaheuristics. In addition to serving as the primary manner by which a metaheuristic improves a solution, they also constitute the basis of a wide variety of different metaheuristics. Given the importance of neighborhoods as algorithmic components, it is surprising that academic reporting often lacks the necessary information to enable their exact reimplementation. One design aspect in particular is often overlooked: the order in which neighbors are enumerated. In this presentation, we present a definition for neighborhood enumerations which makes explicit enumeration order. We analyze and evaluate the impact of neighborhood enumeration order on the performance and behavior of heuristic search procedures by means of a series of computational experiments.

# Multi-Area Reserve Dimensioning using Chance-Constrained Optimization

Alberte Bouso
N-SIDE, Energy Department
e-mail : abo@n-side.com

In modern power systems, the integration of variable renewable energy production and electrification of sectors such as heating or transport introduces new degrees of uncertainty, which have to be managed by Transmission System Operators (TSOs). Among other tasks, TSOs need to ensure that a balance between production and consumption is maintained at all times. Such activity is becoming increasingly complex due to more significant imbalances caused precisely by these systems' uncertainties. The procurement of balancing reserves to compensate for the errors, as mentioned earlier, is needed to ensure a secure and stable power system. An important question is how to size the amount of balancing reserves that are required for the system such that one meets the reliability requirements while minimizing the cost of such service procurement.

In this work, the authors present how we precisely cast the problem of dimensioning Frequency Restoration Reserves (FRR) according to a probabilistic criterion that is compliant with EU legislation while accounting for transmission constraints. This novel methodology aims at decreasing the total reserve needs in both upward and downward directions (upward reserves cover a short system with not enough power injection, while downward reserves cover a long system with too much energy in the system) for each area while keeping the desired reliability at the system level, and accounting for reserve exchange between the Load Frequency Control (LFC) areas.

In particular, the authors propose a heuristic algorithm for resolving a chance-constrained model, formulated as a large-scale Mixed-Integer Linear Program (MILP), which is computationally tractable and exact in some instances. The methodology is adapted to size automatic FRR (aFRR) and manual FRR (mFRR) capacity separately while accounting for their interactions. These two sub-products are common across European balancing markets and aim to cover the imbalance's fast and slow components. A second step aims at allocating reserves in a way that limits unnecessary use of the transmission network, i.e., for equal solutions, choose the one that leads to the minor cross-zonal exchanges.

The mixed-integer programming formulation is highly flexible and extensible : it covers simultaneously upward and downward reserve requirements, uncertainty in the form of normally distributed imbalances, contingencies, and uncertain transmission capacities, and can be extended to consider additional features such as minimum zonal reserve requirements or limits on inter-temporal variations of reserve requirements.

The algorithm is applied to the problem of sizing reserves in the Swedish power system's four load frequency control areas, which is possible thanks to the collaboration between N-SIDE and the Swedish TSO, Svenska kraftnät. The algorithm executes within minutes and can also be executed by open-source solvers within an acceptable run time. The resulting sizing decision can capture 76.1% (downward reserve) and 77.7% (upward reserve) of the potential savings achieved by reserve sharing in the absence of transmission constraints, compared against a non-collaborative scenario.

# Exact mixed-integer programming approach for chance-constrained multi-area reserve sizing

Jehum Cho
Uuiversité Catholique de Louvain (UCLouvain),
Center for Operations Research and Econometrics (CORE)
e-mail: jehum.cho@uclouvain.be

Anthony Papavasiliou
National Technical University of Athens (NTUA),
Department of Electrical and Computer Engineering
e-mail: papavasiliou@mail.ntua.gr

An exact algorithm is developed for the chance-constrained multi-area reserve sizing problem in the presence of transmission network constraints. The problem can be cast as a two-stage stochastic mixed integer linear program using sample approximation. Due to the complicated structure of the problem, existing methods attempt to find a feasible solution based on heuristics. Existing mixed-integer algorithms that can be applied directly to a two-stage stochastic program can only address small-scale problems that are not practical. We have found the minimal description of the projection of our problem onto the space of the first-stage variables. This enables us to directly apply more general Integer Programming techniques for mixing sets, that arise in chance-constrained problems. Our method can tackle real-world problems. We specifically consider a case study of the 10-zone Nordic network with 100,000 scenarios where the optimal solution can be found in approximately 5 minutes.

# Priority Management for Service Platforms with Waiting Time Differentiation

Thomas De Munck
UCLouvain, Center for Operations Research and Econometrics
e-mail: thomas.demunck@uclouvain.be

Philippe Chevalier
UCLouvain, Center for Operations Research and Econometrics
e-mail: philippe.chevalier@uclouvain.be

Jean-Sébastien Tancrez
UCLouvain, Center for Operations Research and Econometrics
e-mail: js.tancrez@uclouvain.be

## 1 Problem Description

Recently, multiple on-demand service platforms (e.g., Lyft, UberEats, DiDi) have considered differentiating their customers according to their willingness to wait and to pay. Similar to what is observed in other industries (e.g., mail delivery, airline boarding), the central idea is to offer customers several options for an almost identical service with the quality mainly determined by the experienced waiting time. Such a practice can be valuable. On the one hand, it allows the platform to allocate capacity where it is the most urgent so that impatient customers wait shorter times. On the other hand, customers willing to wait can benefit from a discounted price. Overall, the platform alleviates the total waiting cost endured by the customers and extracts some part of the created value.

There exist several illustrations of waiting time differentiation on on-demand service platforms. For instance, Lyft proposes a priority pickup option to riders in various cities worldwide. According to this option, a rider decides whether to be served by the standard or the express option, which ensures a quick match with the nearest driver. In the context of food delivery, Uber Eats may offer its customers either to pay an extra fee for faster delivery or to get a price discount in exchange for a longer delivery time.

## 2 Methodology

In this work, we consider an on-demand service platform that assigns service providers to two customer classes, called express and regular customers. In this setting, priority management involves two main decisions. First, the platform determines when to admit customers of each class. Behind that decision lies a trade-off between serving as many customers as possible, and rejecting them to avoid system congestion and costly customer abandonment. Second, the platform decides how to allocate its service providers to each customer class. Because express customers are more impatient and provide higher revenues, the platform may purposely avoid serving admitted regular customers and reserve service providers for later arrivals of express customers.

To support on-demand service platforms in addressing these two decisions, customer admission and service provider allocation, we develop a stylized model formulated as a Markov decision process. The model considers two customer classes that arrive following Poisson processes and that may abandon after an exponentially distributed waiting time. Upon arrival, the customers can be admitted to the system or rejected. If admitted, the customers wait in queues dedicated to their class until they are served by a service provider. At each moment, the service providers can be allocated to admitted customers or reserved for future high-priority customers. The costs experienced by the platform include penalties for customer rejection and abandonment, respective to the customer class. Each penalty is associated with a loss of brand image and loss of marginal profit that results from an unsatisfied customer's request. The model's objective is to determine the dynamic policy that minimizes the platform's total costs.

## 3 Contributions

The contribution of this research is threefold. First, from a mathematical modeling perspective, we elaborate an original model to represent waiting time differentiation on on-demand service platforms. To the best of our knowledge, we are the first to model customer admission and service provider allocation jointly to manage two classes of customers. Due to the novelty of the subject, our model includes a set of features that have not been studied together in the extant literature. These features include customer impatience, multiple servers (service providers in our application), customer classes and the combination of customer admission and server allocation control.

Second, from a theoretical perspective, we use the structural properties of the value function to obtain new results concerning the structure of the optimal policy deduced by our model. Specifically, we establish the existence of state-dependent admission thresholds for both customer classes by showing that the value function is supermodular. This result is useful in practice, as it eases the interpretation of the optimal policy, and its implementation through faster algorithms.

Third, from a managerial perspective, we study the benefits of controlling the customer admission and the service provider allocation in numerical experiments. To this aim, we compare the optimal policy found by our model with simpler policies. We observe that controlling jointly the customer admissions and the service provider allocation enables the system to reach lower costs and operate optimally over an extensive range of system parameters. In addition, the optimal policy can also improve the customer welfare. While it limits abandonment by controlling customer admission, the optimal policy serves additional regular customers by controlling service provider allocation.

# A three-stage service network design model for intermodal transport under uncertainty

T. Delbart(1), Y. Molenbruch(2), K. Braekers(1), A. Caris(1)

[1]Hasselt University, Research group logistics

[2]Vrije Universiteit Brussel (VUB), MOBI research centre

e-mail: thibault.delbart@uhasselt.be

Intermodal transport is the transport of freight with multiple modes whereby freight remains in the same loading unit. The combination of road transport with other high-capacity transport modes such as trains and barges is potentially more sustainable than unimodal road transport for long-haul freight shipping. A drawback of those high-capacity transport modes is their lower flexibility compared to trucks, hindering the modal shift. Logistics service providers typically book capacity in advance on those services, at which point complete demand information is missing. This work aims to assist logistics service providers in their implementation of intermodal transport, which is done with a capacity decision support model.

The proposed model tackles a service network design problem with stochastic demand. In this type of problem, the aim is to select transport services based on the estimated demand and route freight through those services. In academic literature, these types of problems are studied with two-stage models. The first stage takes place before demand has materialised and only the demand distribution is known. Capacity on each service is determined at this stage. The second stage occurs in the short term when complete demand information is available. Routing decisions are taken at this stage, as well as recourse actions in case the capacity selected in the first stage is insufficient. A shortcoming of these two-stage models is that recourse actions are only performed once complete demand information is available. In reality, transport orders arrive up to a few days in advance when it might be too late to perform large updates to the transport plan. Logistics service providers revise their capacity gradually over time as new orders arrive. To more accurately represent reality, we propose a three-stage model that includes an additional intermediate stage.

A logistics service provider is the decision maker in our proposed stochastic optimisation model. It considers a rail-road intermodal network in which train services are offered by third parties. Since different logistics service providers can book those same services, the remaining amount of capacity on the market declines over time. Included uncertainties are stochastic demand and a stochastic decrease in available capacity. Regarding demand, it is assumed that only the distribution of total demand is known in the first stage. In the second stage, additional information on the total demand in the transport market leads to more accurate forecasts of demand volume. Demand materialises in the third stage. The model includes capacity decisions taken in each of the three stages and routing decisions in the last stage. Stochasticity is captured with a scenario tree and the objective is to minimise the average cost over all scenarios. The model is solved with an exact algorithm with a time limit.

*Acknowledgement*
This work is supported by VLAIO (cSBO project DISpATch, HBC.2016.0412)

# An improved decomposition-based heuristic for truck platooning

Boshuai Zhao
KU Leuven, ORSTAT
e-mail: boshuai.zhao@kuleuven.be

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

Truck platooning is a promising transportation mode to save fuel consumption in the freight industry. In this paper, we consider a truck platooning system for which we jointly optimize the truck routes and schedules from the perspective of a central platform. We improve an existing decomposition-based heuristic by Luo & Larson [1], which iteratively solves a routing and a scheduling problem, with a cost modification step after each scheduling run. We propose different formulations for the routing and scheduling problems and embed these into Luo and Larson's framework, and we examine ways to improve their iterative process. In addition, we propose several preprocessing techniques to accelerate the solution of the scheduling problem and enable the resulting heuristic to deal with large instances. The computational results show that our procedure achieves better performance than the existing one. Moreover, based on sensitivity analysis, we find that settings with long routes, differing path lengths, clustered demands, and sparse networks allow to achieve the greatest benefits of truck platoons.

## 1 Introduction

Truck platooning is a transport mode in which several trucks drive together at a close distance with a synergistic operation (e.g. acceleration, braking, and steering). Due to the aerodynamics, the following trucks undergo less air resistance and thus save fuel consumption.

Creating a truck platoon is often difficult because trucks usually follow different routes and delivery schedules. In most cases, before joining a platoon, drivers have to wait in a station, hub, or highway entrance, and even change original routes to meet other trucks, which unavoidably leads to additional delivery costs and waiting time. Therefore, there is not only a trade-off between the benefits and costs of the platooning mode but also an adjustment for the truck's time schedule.

In this paper, we focus on the operational planning issues of truck platooning from the perspective of a central platform. Specifically, we study the joint routing and scheduling problem for truck platooning.

## 2 Problem statement

The problem statement is as follows: we wish to obtain a distribution plan that minimizes the total delivery cost of all trucks. In graph $G(N, A)$, each truck $v$

has to travel from its origin $O_v$ to destination $D_v$ within its delivery time period $(T_v^{ed}, T_v^{la})$. If a truck $v$ meets at least one other truck at node $i$, they can form a platoon on arc $(i, j)$. The delivery cost for the first truck on arc $(i, j)$ is $c_{ij}$, whereas for the following trucks in a platoon this cost is reduced by ratio $\eta$.

# 3 Decomposition-based heuristic

We propose a heuristic using the framework of Luo & Larson (2021)[1] by iteratively solving routing first and then scheduling. After each scheduling run, the arc costs for the routing stage are updated. Our algorithmic differences are presented in Table 1.

Table 1: Differences between Luo & Larson [1] and our procedure

| | Routing | Scheduling | Iterative procedure |
|---|---|---|---|
| Luo & Larson | Three sets of variables | Assignment-based formulation | Three types of cost modifications |
| This work | Two sets of variables | Time-indexed formulation | Four types of cost modifications |

# 4 Results

By proposing different formulations for the two decomposed problems (routing and scheduling) of the framework and slightly improving its iterative process, we speed up the computation and reduce the optimality gap. Moreover, we propose several pre-processing techniques to better solve the scheduling problem, and a scheduling heuristic is proposed to handle the large-scale platooning problem. Based on numerical experiments, we find that our procedure is suitable for practical situations because it can quickly coordinate the routes and schedules for many trucks based on a real road network. Our procedure is beneficial especially when delivery windows are tight, i.e. when companies prioritize delivery timeliness and are unwilling to drastically alter the delivery moments. In addition, based on sensitivity analysis, we find that truck platoon formation tends to be interesting especially in case of long delivery paths, varying path lengths, clustered demand, and sparse networks.

# References

[1] Fengqiao Luo and Jeffrey Larson. A Repeated Route-then-Schedule Approach to Coordinated Vehicle Platooning: Algorithms, Valid Inequalities and Computation, Operations Research, Preprint, 2021.

# A MILP model to analyse and optimise the capacity of railway networks

Inneke Van Hoeck
KU Leuven, Institute for Mobility
e-mail: inneke.vanhoeck@kuleuven.be

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

The demand for railway transport keeps growing due to several economic and environmental reasons. In order to answer this growing demand, a thorough understanding of the current capacity of railway networks is necessary. This can lead to more efficient planning by identifying bottlenecks. A commonly used measure for the capacity is the capacity occupation. This is defined as the minimal amount of time that a sequence of trains needs to traverse a network. A well-known method to determine the capacity occupation is timetable compression. The International Union of Railways (UIC) has proposed a general method to do this in practice (UIC, 2013). In this method, a network is decomposed into lines and nodes. The capacity occupation is then determined separately for these different parts. This will actually lead to an underestimation of the capacity occupation of the entire network because not all train dependencies and network effects are included. Thus, to accurately assess the railway capacity of a network, it is necessary to consider the entire network in microscopic detail to include all the train dependencies. This is exactly what we have done in previous research (Uyttendaele et al.). We extended the max-plus automata method for timetable compression developed by Besinovic et al. (2017) to be applicable for large networks. The developed method works well to compress a given timetable.

However, we are not only interested in performing timetable compression for a given timetable, we would also like to examine how the timetable could be changed to obtain a lower capacity occupation. With this in mind, we propose a MILP method to perform timetable compression. To consider the network in microscopic detail, it is divided into small, non-overlapping pieces of infrastructure called resources. A train driving through the network is represented by blocking times of the resources that are used. The model is based on a machine scheduling problem, the resources are the machines and the blocking times of a train are the jobs.

The model can be extended to optimize the capacity occupation for different scenarios. A first change that can be made is to not assume a fixed order of (all of) the trains. The model can then determine which order is optimal from a capacity occupation point of view. A second possible change is to include alternative routes for the trains. For each train, exactly one route needs to be selected such that the chosen set of routes leads to a minimal capacity occupation.

The proposed model is tested using data provided by Infrabel. The problem is implemented in Python and solved with Gurobi Optimizer. Three zones, of increasing size, centered around the station of Halle are considered during one hour in the morning peak. Experiments show that the MILP model can quickly compress a given timetable. When alternative routes are considered, preliminary results show that the computation time increases quickly when more possible routes are considered for each train. Therefore, a heuristic approach will be developed.

## References

Bešinović, N., Goverde, R. M. P., Quaglietta, E. (2017). *Microscopic Models and Network Transformations for Automated Railway Traffic Planning.* Computer-Aided Civil and Infrastructure Engineering, 32(2), 89–106.

UIC, (2013). Capacity (UIC Code 406) (Issue 2nd edition).

Uyttendaele, J., Van Hoeck I., Bešinović, N., Vansteenwegen, P., *Timetable Compression using Max-Plus Automata applied to Large Railway Networks.* TOP, An Official Journal of the Spanish Society of Statistics and Operations Research, under review.

# A fast deepest-bottom-left-fill algorithm to solve 3D nesting problems using a semi-discrete representation

S. Chehrazad
KU Leuven, Department of Computer Science
campus Leuven, Belgium
e-mail: sahar.chehrazad@kuleuven.be

D. Roose
KU Leuven, Department of Computer Science
campus Leuven, Belgium
email: dirk.roose@kuleuven.be

T. Wauters
KU Leuven, Department of Computer Science
Technologiecampus Gent, Belgium
email: tony.wauters@kuleuven.be

We present an efficient algorithm to solve packing problems for 3D possibly non-convex polyhedral pieces. The aim is to place the pieces without overlap in a container, while minimizing the total length ($x$-dimension) of the container. Both the pieces and the container are discretized in the $x$- and $y$-directions. Discretized pieces and container are represented by a set of continuous line segments in the $z$-direction for each $(x, y)$ grid point. We developed an efficient procedure to semi-discretize the pieces and the container that are initially in STL (Standard Triangle Language) file format. We apply extensions to the line segments of the polyhedron to ensure that non-overlapping placement of the line segments guarantees a non-overlapping placement of the polyhedron.

The proposed deepest-left-bottom-fill placement algorithm is an extension of the 2D left-bottom-fill algorithm using semi-discrete representation, that we developed earlier, see [1]. Each semi-discretized $y$-$z$-plane of a 3D piece is placed in a semi-discretized $y$-$z$-plane of a 3D container using this 2D left-bottom-fill algorithm. The placement can be performed very efficiently due to the use of appropriate data structures, simple arithmetic operations to detect and avoid overlap and an optimised ordering of the segment overlap tests.

The performance and the computational cost of the deepest-left-bottom-fill heuristic depends on the resolution (distance between the $(x, y)$ grid points) and the number of rotation angles. Placement of 36 polyhedra ('Stoyan' data set), with the pieces ordered largest-first, with 8 rotations, with resolution 0.1, results in a container length ($x$-dimension) of 37.5 and requires 856 ms on a single core of an Intel i9 CPU. As resolution gets finer, the execution time increases, as the number of line segments representing a piece increases and consequently the amount of computations required for checking whether a piece can be placed in a position increases.

Our parallel implementation of the algorithm reduces the execution time. We evaluate three strategies: a) testing the placement with each rotation angle concurrently as explained in [2]; b) testing the placement of $y$-$z$-planes of an angle concurrently; c) parallel placement of $y$-$z$-planes of all rotation angles. The speedup that can be obtained with strategies a and b is limited due to load imbalance, since the number of operations in the concurrent tasks can differ substantially. Strategy c allows to create many tasks with sufficient grain-size when the resolution is fine. Therefore, by employing dynamic load balancing the load imbalance problem can be solved, leading to a higher speedup. The execution time of the parallel deepest-left-bottom-fill algorithm with dynamic load balancing using the same data set and ordering of pieces, 8 rotations and resolution 0.1, using 10 cores of an Intel i9 CPU is 164 ms (the obtained speedup is 5). Due to the redundant computations in the parallel algorithm, the ideal speedup without considering parallelization overhead is 7.2.

As the computing times are always in the order of milliseconds on an Intel i9 CPU, our algorithm can be considered as a high-performance building block for use in heuristics for optimal 3D placement.

## References

[1] S. Chehrazad and D. Roose and T. Wauters, "A fast and scalable bottom-left-fill algorithm to solve nesting problems using a semi-discrete representation," European Journal of Operational Research, 300, pp. 809–826, 2022.

[2] S. Chehrazad and D. Roose and T. Wauters, "A Parallel Bucket-Based Bottom-Left-Fill Algorithm for Nesting Problems Using a Semi-Discrete Representation," Proceedings of AlioEuro Conference, pp. 11-16, 2022.

# Automating and optimising pallet loading for irregular-shaped parts

Emmanouil Thanos Tony Wauters
KU Leuven, Department of Computer Science
e-mail: emmanouil.thanos@kuleuven.be

Material storage is a crucial element of operations within heavy industry. Maximising pallet storage usage can significantly accelerate transportation and production process and thus reduce operational costs. However, restrictions resulting from items' shapes and physical properties may highly constraint possible packing configurations. In addition, limited available space imposes great difficulties with respect to the automation of the loading process in confined settings.

This study proposes a model and an algorithmic approach to automate and optimise pallet loading of irregular-shaped, identical parts. The general setting involves a press brake environment in which an automatic robot arm bends metal sheets to generate parts of certain shape and specifications. Those parts must then be placed by the robot arm into pallets in order to be transferred to other locations of the production process. Given a specific type of part, the focus is on designing the palletising process such that the number of parts which are stored in each pallet is maximised.

The developed model formulates the accurate layout positions of the parts within a given pallet. Depending on the type of part in effect, the proposed approach may apply either a preselected stacking construction process or a more generic functionality which, based on collision detection, forms groups of parts to minimise their combined volume.

Generated solutions must respect various physical and stability-related restrictions in order to result in pallets which can be safely transferred by appropriate vehicles. Finally, when it comes to automated palletising, they must also respect robot arm moving and motion planning restrictions, each time given the existing stacking configuration.

# Joint Order Batching and Picker Routing Problem including congestion

P. Torrealba-González, M. Ogier, F. Semet
Univ. Lille, CNRS, Inria, Centrale Lille, UMR 9189 CRIStAL,
F-59000 Lille, France
e-mail: {pablo.torrealba-gonzalez,maxime.ogier,frederic.semet}
@centralelille.fr

D. Feillet
Mines Saint-Etienne, Univ. Clermont Auvergne, LIMOS UMR CNRS 6158,
Centre CMP, Gardanne F-13541, France
e-mail: feillet@emse.fr

## 1 Introduction

In warehouse operations, one of the main processes is the order picking. Order picking deals with the preparation of different customer orders, and concentrates the majority of the total operational costs. Picking tasks are usually performed by human operators pushing a trolley and moving around the warehouse to collect products. At the operational level, the main two decisions taken by managers are order batching and picker routing. The order batching problem (OBP) consists in determining the set of orders assigned to each one of the pickers. The picker routing problem (PRP) is to define the minimal distance tour to collect all the products of a given set of orders. In an integrated approach, the joint order batching and picker routing problem (JOBPRP) considers both decisions jointly and can be solved to optimality, with a branch-and-cut approach [3] or with a column generation based approach [1].

## 2 Including and modeling congestion

In the literature a common assumption is the non-existence of congestion produced by pickers, in this work we propose to include the effect of congestion as a delay in the nominal travel time. To quantify the delay, we divide the planning horizon in time intervals and compute the number of pickers that are visiting each sub-aisle during each time interval. If there is more than one picker visiting a sub-aisle during the same time interval, then all these pickers incur a delay in their travel time. The delay function increases with the number of pickers.

In order to include congestion it becomes necessary to know the time at which the picker visits each place, thus, timing variables are necessary. As in an optimal solution for the PRP a given node can be visited more than once, but an arc is visited at most once [2], we propose a graph transformation that allows us to compute the visit time for each movement. The transformation consists in defining a node for each arc in the original graph such that it is possible to assign a single timing for each node in the transformed graph. A Mixed Integer Linear program (MIP) model for the JOBPRP that includes the consideration of congestion is formulated on the transformed graph.

# 3 Solving the JOBPRP including congestion

When we solve this model with a commercial solver, optimal solutions can be reached only for very small size instances. To evaluate the model on larger instances a two-steps procedure has been implemented. First, we solve the JOBPRP without considering the effect of congestion. Then, the routes obtained in the first step are timed to include the effect of congestion. We evaluate and discuss the relevance of the congestion modeling, analyzing the impact of the number of time intervals, and the existence of waiting times. To obtain good quality solutions for larger instances we re-formulate the model and present an exponential linear programming formulation in which variables (or columns), are associated to possible picking routes including the timing and the level of congestion in each sub-aisle. The model is solved using a column generation procedure in which we call a dedicated dynamic programming procedure to obtain a negative reduced cost column.

# References

[1] Briant, O., Cambazard, H., Cattaruzza, D., Catusse, N., Ladier, A., & Ogier, M. (2020). An efficient and general approach for the joint order batching and picker routing problem. European Journal of Operational Research, 285(2):497–512.

[2] Ratliff, H, & Rosenthal, A. (1983). Order-picking in a rectangular warehouse: A solvable case of the traveling salesman problem. Operations Research, 31(3):507–521.

[3] Valle, C., Beasley, J., & Salles da Cunha, A. (2017). Optimally solving the joint order batching and picker routing problem. European Journal of Operational Research, 262(3):817–834.

# Using cross-learning forecasting methods to improve anticipatory order picking efficiency

Son Tran
Maastricht University, Department of Quantitative Economics
e-mail: t.tran@maastrichtuniversity.nl

Rui Jorge Almeida
Maastricht University, Department of Data Analytics and Digitization
e-mail: rj.almeida@maastrichtuniversity.nl

Christof Defryn
Maastricht University, Department of Quantitative Economics
e-mail: c.defryn@maastrichtuniversity.nl

E-commerce activities put warehouses under high pressure. First, the smaller order sizes result in efficiency loss when retrieving the requested items from storage due to missing economies of scale, while the cost for handling and packaging of these smaller orders tends to be higher [2]. Second, the very short delivery lead times require warehouses to respond to incoming customer orders within hours or even minutes. Because of this short planning horizon, customer demand is highly uncertain up to the moment that a customer places the order. Such a high degree of dynamism makes it more challenging to operate at minimum cost and easily leads to ad-hoc decisions and missed opportunities for optimization [2].

To deal with such a high degree of dynamism and uncertainty while ensuring high efficiency in our picking operations, we make use of *anticipatory order picking* (AOP). The core idea is to explicitly incorporate forecast information, provided by machine learning methods, and allow the picking of SKUs even before they are actually ordered by the customer. As such, products can be picked when their cost (marginal travel time) is low, which results in a reduced total travel time compared to the situation in which products can only be picked once ordered. Our previous research also revealed that AOP leads to a better capacity utilization, a more balanced workload and an earlier completion time for picking all customer orders over the full planning horizon.

In this research, we focus on the generation of the required forecast information to allow for a successful AOP implementation. Many of the popular methods for forecasting retail sales data are limited to individual products. At the same time, however, empirical studies show that small forecast accuracy improvements can result in significant higher service levels and lower stock holding cost [3]. The M forecasting competitions, whose goal is to identify ways to improve the forecasting accuracy by empirically evaluating several forecasting methods, show that machine learning forecasting methods utilizing cross-learning obtained more accurate results than the alternatives when the dataset comprise aligned highlycorrelated series structured in a hierarchical fashion [1].

The accuracy and content of the forecast information are decisive factors on the performance of anticipatory order picking. With reasonably accurate forecasts, warehouse operators can substantially shorten the completion time of the picking process. In this study, we make use of machine learning forecasting methods for demand prediction using real-life instances after which we employ this information to construct order picking plans. We aim to explore how using the demand of related products to forecast the demand of individual product may provide more accurate inputs for AOP than using the demand of that individual product only and how such forecast information can be connected to improve order picking performance.

## References

[1] Makridakis, S., Spiliotis, E. & Assimakopoulos, V. (2022), "M5 accuracy competition: Results, findings, and conclusions", *International Journal of Forecasting.*

[2] Moons, S., Ramaekers, K., Caris, A. & Arda, Y. (2018), "Integration of order picking and vehicle routing in a B2C e-commerce context", *Flexible Services and Manufacturing Journal* 30(4), 813–843.

[3] Pooya, A., Pakdaman, M. & Tadj, L. (2019), "Exact and approximate solution for optimal inventory control of two-stock with reworking and forecasting of demand", *Operational Research* 19(2), 333–346.

# Hierarchical Coordination of Transmission and Distribution System Operations in European Balancing Markets

Ilyès Mezghani
Engie Impact
ilyes.mezghani@engie.com

Nicolas Stevens
UCLouvain
nicolas.stevens@uclouvain.be

Anthony Papavasiliou
National Technical University of Athens
papavasiliou@mail.ntua.gr

Dimitris I. Chatzigiannis
Aristotle University of Thessaloniki
dimitris.chatzigiannis@gmail.com

We propose a hierarchical scheme for the coordination of transmission and distribution system operations, which is inspired by nested decomposition and tailored for integration in upcoming European balancing markets.

The are two principal contributions in our work: (i) to outline a well-defined market design for flexibility platforms that aims at aligning with integrated European balancing market evolutions, and (ii) to propose a computational infrastructure for implementing the proposed market design.

Our proposed hierarchical approach aims at bridging the gap between the physical details of the grid constraints (that include the distribution grid) and the real-time market, operated by the Transmission System Operator (TSO), which typically ignores most of the distribution grid constraints, notably for institutional reasons. In brief, the approach consists of representing *implicitly* the distribution grid constraints in the form of "grid-secure" bids (later called a *residual supply function*) that are submitted to the TSO balancing market. The approach preserves the independence of the Distribution System Operator (DSO) while accounting for its grid constraints in the market operated by the TSO.

The transmission network that we consider is meshed, and we use a standard DC approximation which is considered adequate for high-voltage grids in market clearing models. Distribution networks are assumed to be radial in our model, and we consider AC power flow in order to account for losses, voltage limits, and reactive flows, which are more relevant in medium and low-voltage grids.

We compare our proposal to a number of recent European research projects and pilots. We demonstrate the effectiveness of our proposal in resolving a number of transmission-distribution coordination dilemmas, including conflict resolution between network operators, self-healing, sharing charges in mutually beneficial actions, and the recursive integration of distributed resources in pan-European balancing markets. We implement a proof-of-concept market clearing platform that can match transmission and distribution system market orders, cope with non-convexities in market offers and power flow constraints, scale to systems of realistic size, and respect the decentralization of communications and information sharing which is required in European electricity markets. We demonstrate the effectiveness of our platform on large-scale instances of the Italian and Danish power systems based on data sourced from the EU Horizon 2020 SmartNet project. We demonstrate the superiority of our approach to alternative coordination schemes in terms of economic efficiency, alignment of incentives and system security.

# Locally stable exchanges

M. Baratto
HEC Management School of the University of Liege,
e-mail: marie.baratto@uliege.be

Y. Crama
HEC Management School of the University of Liege,
e-mail: yves.crama@uliege.be

J.P. Pedroso
INESC TEC and Faculty of Sciences of the University of Porto
e-mail: jpp@fc.up.pt

A. Viana
INESC TEC and Polytechnic of Porto, School of Engineering
e-mail: ana.viana@inesctec.pt

While optimal stable matchings and optimal kidney exchanges have been widely studied, the optimization of stable kidney exchanges has not been investigated as much. We extend recent work by [1] on this topic by introducing and underlining the relevance of a new concept of *local stable exchanges* (L-stable exchanges).

A *compatibility digraph* is a directed graph $G = (V, A)$ (representing, for example, feasible kidney transplants) such that for each vertex $i \in V$, a preference order is given on the set of neighbors $N^-(i) = \{j : (j, i) \in A\}$. We introduce several new definitions:

- Given a cycle $c$ in $G$, its set of *potentially L-blocking cycles*, $PBC(c)$, is the set of cycles $c'$ non vertex disjoint with $c$ and such that each vertex of $c'$ is either not in $c$ or *strictly* prefers its follower in $c'$ to its follower in $c$.
- Given an exchange (i.e., a cycle packing) $\mathcal{M}$ in $G$, an *L-blocking cycle $c$* for $\mathcal{M}$ is a cycle that is not included in $\mathcal{M}$ but that has a vertex in common with $\mathcal{M}$, and such that, for every vertex $i$ involved in cycle $c$, either $i$ is not involved in $\mathcal{M}$, or $i$ prefers its follower in $c$ to its follower in $c'$, where $i \in c' \in \mathcal{M}$.
- An *L-stable exchange* is an exchange without L-blocking cycle.

Then, with the compatibility digraph $G = (V, A)$, we associate a *blocking digraph* $G' = (V', A')$ such that:

- The vertex set $V'$ is the set of cycles of $G$.
- An arc $(c_u, c_v)$ is in $A'$ if the cycle $c_u$ intersects the cycle $c_v$ in $G$ and if
  - either $c_u$ is potentially L-blocked by $c_v$,

- or $c_u$ is not potentially L-blocked by $c_v$, and $c_v$ is not potentially L-blocked by $c_u$.

The problem of finding a local stable exchange in a directed graph is equivalent to finding a so-called local kernel (L-kernel) in $G'$ (whereas finding a stable exchange is equivalent to finding a kernel of $G'$). The notion of local kernel has been considered earlier in graph theory (see, e.g., [2]). However, the literature does not seem to mention any algorithmic or numerical contributions.

We prove that it is NP-hard to determine whether a graph has a nonempty local kernel, and hence, to find a local kernel of maximum cardinality.

Next, we propose several integer programming formulations for the maximum L-stable exchange problem. These formulations can simply be viewed as modeling the maximum L-kernel problem in a directed graph. We also numerically compare these formulations. The results show that the most efficient formulation (with regard to computing time) is one where there is a cycle packing constraint and a stability constraint for each cycle. Other formulations have such constraints for each pair of intersecting cycles. We numerically compare our formulation with a formulation of the maximum stable exchange problem proposed in [1]. Even though the two problems are not equivalent, all the instances tested so far are such that, when a maximum stable exchange exists, then its cardinality is equal to the cardinality of a maximum L-stable exchange. On the other hand, tests on the instances that do not have a stable exchange reveal that all these instances have a (nonempty) L-stable exchange. Even more interestingly, for a given instance size, the average cardinality of a maximum L-stable exchange is very close to the average optimal value over the instances that do have a stable exchange. These observations suggest that the concept of local stable exchange is indeed interesting, and provides relevant solutions for more instances than the concept of stable exchanges.

All the above results and observations carry over when the concept of *local stable exchanges* is extended to the concept of *local strong stable exchanges* (in a similar way as what was done earlier for strongly stable exchanges).

## References

[1] Klimentova, X., Biró, P., Viana, A., Costa, V., and Pedroso, J.P. (2022). Integer programming models for the stable kidney exchange problem.

[2] Duchet, P. and Meyniel, H. (1993). Kernels in directed graphs: a poison game. *Discrete Mathematics*, 115(1):273-276.

# Exact route evaluation for node edge and arc routing problems with stochastic customers

O. Tellez
HEC Liège - Management School of the University of Liège – Belgium
e-mail: oatellez@uliege.be

V. François
HEC Liège - Management School of the University of Liège – Belgium
e-mail: veronique.françois@uliege.be

Y. Arda
HEC Liège - Management School of the University of Liège – Belgium
e-mail: yasmin.arda@uliege.be

Arc routing problems arise in applications such as postal delivery services or waste collection where the density of service points is high, and the direction of edge traversal is relevant (Assad and Golden, 1995). We consider a general arc routing setting where the presence of nodes, edges and arcs of the network is stochastic according to a known probability. This feature is known as stochastic customers in the vehicle routing literature (Lagos et al., 2019). While considerable research has been devoted to the arc routing problem with stochastic demands, research about stochastic customers is rather scarce (De Maio et al., 2021).

Evaluating the expected travel time of a given route in the presence of stochastic customers is a complex task because is requires computing a shortest path for each distinct realization of the random variables, i.e., the presence or absence of the stochastic customers. Even when a fixed sequence is used for customer visits, decisions about edge traversal orientation must still be made. In the literature, a fixed traversal orientation of customers has been considered for all realizations of a given route (Mohan et al., 2010). However, this type of approach may be highly sub-optimal when the customer presence probability is low, as this will likely lead to an overestimated travel time.

In this research, we propose a method to accurately estimate the expected travel time of routes containing stochastic customers, assuming a fixed sequence of customer visits. Each time a customer is missing, the path is adapted by skipping that customer and following the shortest path to the next stop. Unlike existing approaches, the proposed procedure optimally re-evaluates the best edge traversal orientation for each possible realization of the random variables. The time complexity of the proposed method is $\mathcal{O}(n^2)$ for $n$ customers, which is lower than the complexity of one known approach using fixed sequences and fixed edge traversal orientations ($\mathcal{O}(n^3)$).

We perform extensive numerical analysis to evaluate the practical performance of the presented method on different instance sizes. We also measure the travel time improvement against a fixed edge traversal strategy. Finally, we present managerial insights related to the impact of the customer presence probability.

## References

Assad, A. A. and Golden, B. L. (1995). Chapter 5 Arc routing methods and applications. *Handbooks in Operations Research and Management Science*, 8:375–483.

De Maio, A., Laganà, D., Musmanno, R., and Vocaturo, F. (2021). Arc routing under uncertainty: Introduction and literature review. *Computers & Operations Research*, 135:105442.

Lagos, F., Klapp, M. A., and Toriello, A. (2019). Branch-and-Price for Routing with Probabilistic Customers. *H. Milton Stewart School of Industrial and Systems Engineering Georgia Institute of Technology*, (1):1–35.

Mohan, S., Gendreau, M., and Rousseau, J.-M. (2010). Heuristics for the Stochastic Eulerian Tour Problem. *European Journal of Operational Research*, 203:107–117.

# The Electric On-Demand Bus Routing Problem with Partial Charging and A Nonlinear Function

Ying Lian
ANT/OR - Operations Research Group, Department of Engineering Management, University of Antwerp
e-mail: ying.lian@uantwerpen.be

Flavien Lucas
IMT Nord Europe, Institut Mines-Télécom, Univ. Lille Centre for Digital Systems, F-59000 Lille, France
e-mail: flavien.lucas@imt-nord-europe.fr

Kenneth Sörensen
ANT/OR - Operations Research Group, Department of Engineering Management, University of Antwerp
e-mail: kenneth.sorensen@uantwerpen.be

The electric vehicle routing problems (EVRPs) with recharging policy consider the limited range of electric vehicles and thus include intermediate visits to charging stations (CSs). In general, minimizing the resultant charging costs such as charging duration or charging amount are also part of the objective of the EVRP. Accordingly, the EVRP have received considerable attention over the past years. Nevertheless, this type of problems in the domain of passenger transportation, a VRP variant, has been rarely studied in the literature, especially with time windows, a realistic nonlinear charging function or partial charging policy. Hence this research extends the existing work on the EVRP to an On-Demand Bus Routing Problem (ODBRP) which transports passengers with the bus station assignment (BSA). The resultant problem is the EODBRP. Specifically, each passenger can have more than one stations to board or alight, and they are assigned to the ones with the smallest increase in the total user ride time (URT). In the EODBRP, frequent intermediate visits to CSs are considered. Moreover, a nonlinear charging function is used and the partial charging strategy is applied. To solve the EODBRP, a greedy insertion method with 'charging first, routing second' strategy is developed, followed by a large neighborhood search (LNS) which consists of local search (LS) operators to further improve the solution quality. Experimental data were generated by a realistic instance generator based on a real city map, and the corresponding results show that the proposed heuristic algorithm performs well in solving the EODBRP. Finally, sensitivity analyses with divergent parameters such as the temporal distributions of passengers and bus ranges may provide practical guidance.

# The TSP with drones: The benefits of retraversing the arcs

N. Morandi[1,⋆] R. Leus[1] J. Matuschke[2] H. Yaman[1]

[1]KU Leuven, Belgium, Research Center for Operations Research and Statistics
[2]KU Leuven, Belgium, Research Center for Operations Management
[⋆]e-mail: `nicola.morandi@kuleuven.be`

Since the pioneering introduction of the Flying Sidekick Traveling Salesman Problem (FSTSP) by Murray and Chu (2015), the scientific literature about applications of drones to routing and parcel delivery has grown at a remarkable pace. At the time of writing, searching the words "truck drone routing" by Google Scholar produced 19 200 results, 20% of which from just 2021. An impressive number of surveys on the topic already appeared, like, among other interesting ones, those by Macrina et al. (2020) and Otto et al. (2018). The interest in drones' applications also comes from public institutions and the private sector. For example, the European Commission forecasts more than 100 000 people employed and an economic impact of over 10 billion Euros per year in the European drone sector by 2035 (European Council, 2019). At the same time, the e-commerce multinational Amazon obtained the approval from the relevant US authority for its Prime Air service "beyond visual line of sight" (CNBC news, 2020). From the algorithmic point of view, the FSTSP and its generalizations can model a wide range of routing applications with cooperating vehicles, which are not necessarily drones; in principle, any vehicle with limited fuel, or traveling person with limited payload capacity, could play the role of the drone in these problems.

In the FSTSP, a truck and a drone cooperate to visit all the customers in a given network in the minimum amount of time. The drone can only serve one customer per sortie (i.e., drone flight). A natural generalization of FSTSP is the TSP with drones (TSP-mD), where multiple drones are allowed to serve multiple customers per sortie, with the length of each sortie bounded by a limited battery capacity. Most studies of FSTSP and TSP-mD impose the additional constraint that the truck cannot visit customers multiple times, with the notable exception of the work by Agatz et al. (2018), Bouman et al. (2018), and Tang et al. (2019). As pointed out by Roberti and Ruthmair (2021), allowing such *revisits* poses additional computational challenges which cannot be easily accommodated by many of the existing approaches in literature. Indeed, by adding sufficiently many copies of the nodes to the underlying graph, one can always reduce revisiting truck routes to cycles; this approach, however, does not appear to be computationally viable in practice.

In this talk, we show that optimal solutions for TSP-mD may not only need to revisit nodes, but also *retraverse* arcs of the underlying directed graph, i.e., in the course of its tour the truck might need to repeatedly travel directly from customer $i$ to customer $j$ for some fixed pair of customers $i, j$. Consequently, excluding such *arc-retraversing* solutions can lead to a significant increase in the optimal value. The necessity of *arc retraversals* does not seem to have been investigated in previous studies, and those studies that allow node revisits seem to operate under the assumption that there always exists an optimal solution without *retraversals*. We prove that this implicit assumption is correct under specific conditions, which are commonly met in the literature. Although, when these conditions are not satisfied (e.g., when allowing multiple customers to be visited by a single sortie), the optimal value might increase significantly when excluding *arc-retraversing* solutions. We provide asymptotically tight upper bounds on such increase. Finally, we provide an approximation algorithm whose approximation guarantee depends on the speed and the number of available drones, and show that unless P=NP, no approximation algorithm can obtain a guarantee that does not depend on these two parameters.

## References

Agatz, N., Bouman, P., and Schmidt, M. (2018). Optimization Approaches for the Traveling Salesman Problem with Drone. *Transportation Science*, 52(4):965–981.

Bouman, P., Agatz, N., and Schmidt, M. (2018). Dynamic programming approaches for the traveling salesman problem with drone. *Networks*, 72(4):528–542.

CNBC news (2020). Amazon wins FAA approval for Prime Air drone delivery fleet. `https://www.cnbc.com/2020/08/31/amazon-prime-now-drone-delivery-fleet-gets-faa-approval.html`. Last accessed: 06-July-2022.

European Council (2019). Drones: reform of EU aviation safety. `https://www.consilium.europa.eu/en/policies/drones/`. Last accessed: 06-July-2022.

Macrina, G., Pugliese, L. D. P., Guerriero, F., and Laporte, G. (2020). Drone-aided routing: A literature review. *Transportation Research Part C: Emerging Technologies*, 120:102762.

Murray, C. C. and Chu, A. G. (2015). The flying sidekick traveling salesman problem: Optimization of drone-assisted parcel delivery. *Transportation Research Part C: Emerging Technologies*, 54:86–109.

Otto, A., Agatz, N., Campbell, J., Golden, B., and Pesch, E. (2018). Optimization approaches for civil applications of unmanned aerial vehicles (UAVs) or aerial drones: A survey. *Networks*, 72(4):411–458.

Roberti, R. and Ruthmair, M. (2021). Exact Methods for the Traveling Salesman Problem with Drone. *Transportation Science*, 55(2):315–335.

Tang, Z., van Hoeve, W.-J., and Shaw, P. (2019). A Study on the Traveling Salesman Problem with a Drone. In Rousseau, L.-M. and Stergiou, K., editors, *Integration of Constraint Programming, Artificial Intelligence, and Operations Research*, pages 557–564, Cham. Springer International Publishing.

# Using Bayesian optimization to solve constrained simulation Optimization problems

Sasan Amini
Department of Quantitative Methods, Hasselt University, 3590 Diepenbeek, Belgium
Department of Information Technology, Ghent University, 9052, Gent, Belgium
Data Science Institute, Hasselt University, 3590 Diepenbeek, Belgium
e-mail: sasan.amini@uhasselt.be

Inneke Van Nieuwenhuyse
Department of Quantitative Methods, Hasselt University, 3590 Diepenbeek, Belgium
Data Science Institute, Hasselt University, 3590 Diepenbeek, Belgium
e-mail: inneke.vannieuwenhuyse@uhasselt.be

Simulation optimization is an often used approach for optimizing black-box functions. Different approaches exist, and have been used successfully in , e.g., engineering design, financial portfolio selection, and hyperparameter tuning of machine learning algorithms. More details on the applications of Simulation optimization could be found in [1]. Most of these approaches (e.g., evolutionary algorithms) require many function evaluations before converging to the optimal solution(s). Yet, this may cause a problem: in some cases, the simulation model is expensive to run (because of long running times, high monetary cost, or hazardous experiments), implying that *data-efficient* optimization algorithms are needed. Moreover, in many simulation optimization problems, the true objective functions (also referred to as outputs) are not observable with perfect accuracy; only *noisy* observations are available [1]. This noise makes it harder to study/predict the effect of the input vector on the output values, and adds to the optimization challenge. While replication is a common strategy to reduce this noise, this again poses a problem in settings with an expensive simulator [5].

In this research, we present some first results on the use of Bayesian optimization (BO) to solve expensive (noisy, constrained) optimization problems. In Bayesian Optimization, a computationally inexpensive meta-model (usually a Gaussian process model [4]) is estimated based on a small set of input-output observations. The algorithm then iteratively decides where to sample next, based on an *infill criterion* or *acquisition function*: in this way, additional observations are added and the new information is incorporated in the metamodel. This sequential sampling continues until a stopping criterion is met.

While there is a vast literature on deterministic unconstrained single-objective BO algorithms, scarce work exists on noisy constrained settings. We use Stochastic Kriging [2] to account for the *inherent noise* in the outputs when constructing the metamodel, and when selecting new points to sample. Furthermore, we adapt the *Expected barrier acquisition function* recently proposed in [3] to noisy settings. The primary experimental results show that the resulting algorithm is capable of finding the optimal solutions in the presence of noisy observations. To conclude, we discuss our future research steps.

## References

[1] S. Amaran, N. V. Sahinidis, B. Sharda, and S. J. Bury. Simulation optimization: a review of algorithms and applications. *Annals of Operations Research*, 240(1):351–380, 2016.

[2] B. Ankenman, B. L. Nelson, and J. Staum. Stochastic kriging for simulation metamodeling. *Operations Research*, 58(2):371–382, 2010.

[3] T. Pourmohamad and H. K. Lee. Bayesian optimization via barrier functions. *Journal of Computational and Graphical Statistics*, (just-accepted):1–23, 2021.

[4] C. E. Rasmussen and C. Williams. Gaussian processes for machine learning the mit press. *Cambridge, MA*, 2006.

[5] D. Zhan and H. Xing. Expected improvement for expensive optimization: a review. *Journal of Global Optimization*, 78(3):507–544, 2020.

# Simulation-based optimization of selective assembly costs in a hybrid production context

A. Mencaroni, D. Claeys, S. De Vuyst
Ghent University, Department of Industrial Systems Engineering and Product Design
Industrial Systems Engineering (ISyE), Flanders Make, www.FlandersMake.be
e-mail: Andrea.Mencaroni;Dieter.Claeys;Stijn.DeVuyst@ugent.be

Selective assembly is a production technique whereby components are smartly instead of randomly matched based on their measured feature values. Although it can reduce scrap while increasing product quality, it is often not adopted due to the high logistic costs related to measuring and matching operations.

Therefore, we propose and investigate a hybrid assembly strategy whereby a number of random assembly rounds are performed before switching to selective assembly. At each round, components are randomly picked from their corresponding piles and the assembly is attempted. At this point, components are either removed from their corresponding piles because matching was possible, or placed into their corresponding unmatched piles after the current round. Those piles will then be the starting point for the random selection of components in the following assembly round. When $N$ such rounds have been performed, the remaining unmatched components will finally be assembled using selective assembly.

Due to these prior random assembly rounds, fewer selective assembly operations will eventually be needed, which considerably reduces logistic costs related to selective assembly. On the other hand, every additional round also implies random assembly trial costs. Hence, an optimal number $N^*$ of random assembly rounds before moving on to selective assembly has to be determined.

Therefore, we develop a cost function together with an optimality condition. This cost function is expressed in terms of the matching probabilities at each round and the relative costs of selective and random assembly. These probabilities express the likelihood of a successful match when components are randomly picked from their corresponding unmatched piles after a certain round.

Since the analytic calculation of the matching probabilities is in most of the practical cases not feasible, we develop a simulation technique to estimate these probabilities. First, we use simulation to obtain a sample of the unmatched components after each assembly round. Then, we use these samples to estimate the matching probabilities. Finally, we evaluate the cost function in order to find the optimal number of random assembly rounds prior to selective assembly.

To assess the performance of this method, we apply it to two study cases from literature: a Shaft and Hole two-components assembly [1] and a four-components Ball Bearing assembly [2]. In order to reduce variability, we investigate and compare three estimation methods for the matching probabilities, and we study their impact on the total variability of the cost function using the Delta Method. Finally, with sensitivity analysis we investigate the behavior of the optimal number of random assembly rounds $N^*$ as the relative cost of selective assembly to random assembly varies.

Despite the differences between the two cases concerning component features' distributions, complexity and assembly matching functions, we observe similar results. Since the matching probabilities rapidly decrease with every additional random assembly round, the highest cost reductions are achieved within the first few rounds. In both cases, the proposed hybrid assembly technique can significantly reduce average assembly costs while still delivering higher quality products.

By addressing the gap in the existing literature about selective assembly around its combination with traditional assembly, this study achieves a double goal. On the one hand, it provides a framework to evaluate the effects on the average assembly cost of performing a number of random assembly rounds prior to apply selective assembly. On the other hand, it allows to estimate assembly costs taking component features' distributions as an input parameter. Moreover, since it provides an efficient way to estimate matching probabilities, this model could also be used to estimate and compare average assembly costs as a function of the quality of incoming components by repeating the analysis with different types of component features' distributions.

## References

[1] Fang, X. and Zhang, Y. (1995). A new algorithm for minimising the surplus parts in selective assembly. *Computers and Industrial Engineering*, 28(2):341–350.

[2] Victor Raj, M., Saravana Sankar, S., and Ponnambalam, S. (2011). Genetic algorithm to optimize manufacturing system efficiency in batch selective assembly. *The International Journal of Advanced Manufacturing Technology*, 57(5):795–810.

# Sample path optimisation methods
# A comparative case study

Harbi Taki Eddine
Royal Military Academy, Department of Maths
e-mail: taki.harbi@mil.be

Lauwence Ben
Royal Military Academy, Department of Maths
e-mail: ben.lauwens@mil.be

Vanden Berghe Greet
KU Leuven, Department of computer science
e-mail: greet.vanden.berghe@kuleuven.be

Many real-world non-deterministic decision-making systems and problems are very difficult, if not impossible, to model using mathematical formulations. This is primarily due to their increasing complexity, which is caused by many stochastic factors. One way to cope with this challenge is to take advantage of the ability of simulation modelling approaches (Discrete event simulation, Agent-based simulations, etc) to mimic real-world systems with integrity while accounting for their stochasticity and complexity.
Overall, the primary objective of analysing these systems through simulations is to study several scenarios and select the best solution, or in other word the possible inputs, for these simulations that improve certain output performance measures. However, these output measures are often non-deterministic in the presence of many stochastic factors embedded in these systems, leading to further problems when trying to select the best set of inputs given their non-deterministic outputs. In fact, dealing with this stochasticity is a major challenge for researchers involved in simulation optimisation field.
A very common tool in the context of simulation optimisation to deal with this uncertainty is the Sample Path (SP) approach or sometimes called the Sample Average Approximation (SAA) approach. It aims to compare several solutions under the same circumstances by evaluating them on the same sample of stochastic factors and then aggregating all the outputs. However, it must evaluate the simulation on each sample and for each set of inputs, which is a considerable computational burden especially where the simulation is expensive in terms of time. To overcome this drawback, the Variable Number Sample Path (VNSP) approach, which is an extension of the sample path approach, attempts to minimise the number of the simulation runs. The VNSP starts solving the problem with a reduced number of samples and iteratively increases it so that in the last iterations a sufficient number of samples is taken into account.
In order to compare the mentioned approaches and to study their ability to handle

the uncertainty inherent in the systems, we conducted an experimental comparison between them. Through this comparison, we aim to select the best approach that offers a good compromise between the computational effort and the quality of the final solution. We have applied both approaches to the vehicle routing problem with soft time windows and stochastic travel time. Each approach is tested with multiple parameters on several instances of the Solomon benchmark. The comparison was based on the quality of the final solution and the number of objective function calls. The results show that the sampling path approach gives better results in terms of final solution quality, however the variable number sampling path makes a very good compromise between the objective function value and the number of simulation runs.

# An OR perspective on conflict-driven cutting plane learning

Jo Devriendt
Nonfiction Software, e-mail: `jo.devriendt@outlook.com`

Linear relaxation lies at the heart of mixed integer programming (MIP) solvers. From an abstract point of view, MIP solvers generate constraints that strengthen some linear relaxation (the LP) of a MIP instance to such an extent that a(n optimal) solution to the LP can be rounded to a(n optimal) solution to the MIP instance. Such constraints are called *cutting planes*, or *cuts* for short. This addition of cuts happens through a myriad of techniques, grouped by categories such as *cut generation*, *row generation* and *pre/inprocessing.* As a result, MIP formulations with a "tight" linear relaxation (where an optimal solution to the LP is relatively close to an optimal solution of the MIP) are favored for a given problem, all else being equal.

However, many interesting problems do not allow tight linear relaxations. A particular class of such problems are those whose natural formulation is *clausal* - the constraints are *clauses* representing disjunctions of Boolean variables or their negations. In MIP parlance, a clause has the following form:

$$\begin{aligned}
&\sum_{i\in I} x_i - \sum_{j\in J} x_j \geq 1 - |J| \\
&x_i \mapsto \{0,1\} \\
&\text{with } I \text{ and } J \text{ disjoint sets of indices}
\end{aligned}$$

After the elimination of all variable fixings due to *unit* constraints (e.g., $-x \geq 0$) each clause contains at least two variables. Hence the LP of *any* set of clauses is satisfied by setting all variables to 0.5, which provides little information on the actual solution to the clausal problem. Instead, state-of-the-art *Boolean satisfiability* (SAT) solvers do not rely on calculating solutions to relaxed clauses. Instead, SAT solvers use a technique called *conflict-driven clause learning* (CDCL). From each conflict encountered during depth-first search, CDCL derives – *learns* – a fresh clause that would have prevented the conflict from happening. Alternatively, CDCL can be seen as a form of cut generation, deriving clausal cuts that hopefully tighten the linear relaxation to resolve a conflict.

In recent years, CDCL has been lifted from clausal problems to *pseudo-Boolean* (PB) problems, or equivalently, binary linear programs. Such problems have the following form:

$$\begin{aligned}&\text{minimize} \sum_{j \in J} c_j x_j\\&\text{such that} \sum_{i \in I, j \in J} a_{ij} x_j \geq b_i\\&a_{ij}, b_i, c_j \in \mathbb{Z}, x_j \mapsto \{0,1\}\\&\text{with } I \text{ and } J \text{ sets of indices}\end{aligned}$$

We call this generalization of CDCL *conflict-driven cutting plane learning* (CDCPL). Similar to CDCL, CDCPL derives a constraint from each search conflict. However, the constraints derived by CDCPL not have to be clausal. They can be binary linear constraints – full-blown cutting planes – which may entail many clauses at once.

In this talk, we explain the basic mechanism behind CDCPL in detail. We show how it works, argue why it works, and provide pointers to state-of-the-art systems implementing CDCPL. We also give a biased impression of the strengths and weaknesses of this approach. The presented techniques are based on [1, 6, 4, 3, 5, 2] and the integer programming solver EXACT[1].

## References

[1] D. L. Berre and A. Parrain. The sat4j library, release 2.2. *J. Satisf. Boolean Model. Comput.*, 7:59–6, 2010.

[2] S. R. Buss and J. Nordström. Proof complexity and SAT solving. In A. Biere, M. J. H. Heule, H. van Maaren, and T. Walsh, editors, *Handbook of Satisfiability*, volume 336 of *Frontiers in Artificial Intelligence and Applications*, chapter 7, pages 233–350. IOS Press, 2nd edition, Feb. 2021.

[3] J. Devriendt. Watched propagation of 0-1 integer linear constraints. In *Proceedings of the 26th International Conference on Principles and Practice of Constraint Programming (CP '20)*, volume 12333 of *Lecture Notes in Computer Science*, pages 160–176. Springer, Sept. 2020.

[4] J. Devriendt, A. Gleixner, and J. Nordström. Learn to relax: Integrating 0-1 integer linear programming with pseudo-Boolean conflict-driven search. In *Proceedings of the 17th International Conference on the Integration of Constraint Programming, Artificial Intelligence, and Operations Research (CPAIOR '20)*, volume 12296 of *Lecture Notes in Computer Science*, pages xxiv–xxv. Springer, Sept. 2020.

[5] J. Devriendt, S. Gocht, E. Demirović, J. Nordström, and P. Stuckey. Cutting to the core of pseudo-Boolean optimization: Combining core-guided search with cutting planes reasoning. volume 35, pages 3750–3758, May 2021.

[6] J. Elffers and J. Nordström. Divide and conquer: Towards faster pseudo-Boolean solving. In *Proceedings of the 27th International Joint Conference on Artificial Intelligence (IJCAI '18)*, pages 1291–1299, July 2018.

[1] `https://gitlab.com/JoD/exact`

# Agent Based Simulation for Synchromodal Transport

Shafagh Alaei Jordehi
Vrije Universiteit Brussels, Department of Business Technology and Operations
E-mail: shafagh.alaeijordehi@vub.be

Javier Durán-Micco
Vrije Universiteit Brussels, Department of Business Technology and Operations
E-mail: javier.vicente.duran.micco@vub.be

Cathy Macharis
Vrije Universiteit Brussels, Department of Business Technology and Operations
E-mail: cathy.macharis@vub.be

The focus of the current EU transport policy is to ensure a sustainable transport sector and meet the European Commission's goal of cutting 55 percent greenhouse gas emissions by 2030 compared to the 1990 level. For long, the so-called "modal shift" has been put forward as the key solution to a multitude of environmental and congestion problems associated with freight transport. Accordingly, different novel concepts have been put forward in the literature, such as intermodality and, more recently, synchromodality [3].

Intermodal transport focuses on transporting goods from origin to destination using a sequence of different transportation modes, often using the same transport unit throughout the journey to minimize material handling at the transfer points [1]. Synchromodal transport is the newest concept in the conceptual evolotion of intermodal transport, in which it offers the possibility to shift freely between transportation modes at particular times/nodes while the shipload is in transit [1]. The modal choice is thus not fixed before the trip but made based on real-time data of the transportation network and the orders in transit, which provides more flexibility [5]. Thus, synchromodal transport is much more complex in comparison with intermodal transport, as the decisions regarding the mode choice, vehicle routing, transshipment terminal selection, etc., are needed to be made in real time. This real-time decision making, entails uncertainties (e.g., regarding the availability of different capacities, delays, services cancellation, etc.). In order to minimize the impact of these uncertainties and promote synchromodality, decision-makers will need a mechanism that provides reliable assessments in a highly dynamic and real-time environment and enables tests on (future) developments. Such a platform would be able to mirror the current real system and simulate how it could evolve. To this end, the notion of the digital twin (DT) comes into play. A DT is a virtual (simulation) model of a system that collects the best available physical models, sensor updates, system history, etc., to mirror the life of the physical system [2]. Such a mechanism would allow decision-makers to plan ahead (accounting for uncertainties involved) while stay-ing robust in their eventual decisions. Having a great level of detail provided by the digital twin, we can capture complex interrelations between different players in the chain, study the current and future state of the system, testing dynamic algorithms and technologies in a risk-free environment, and more importantly study how the disruptions affect on the performance of the system.

Although the concept of DT is receiving more attention, it has never or rarely been considered in the synchromodal context. Our work is intended to fill this gap by building this DT via linking the real-time data to the (virtual) simulation model of the system. The objective of this work (as part of the DISpATch project [4]) is to develop a DT model by integrating simulation techniques, GIS platforms, and optimization algorithms to enhance the synchromodal concept further.

As a first step toward building this DT, we are focusing on developing a conventional simulation model of the transport setting. This simulation model, is being developed using Agent Based Modeling (ABM), Discrete Event Modeling (DEM), as well as the optimization algorithms. The system will be modelled as collection of the autonomous and heterogeneous agents (trains, trucks, barges, terminals, DCs, retailers, etc). Then, the model will be linked with the physical (living) system in order to update itself through real-time data, which is received from sensors that are embedded inside the system. Thus, by combining the physical model, real time data, and the system history, the DT model will be developed and it will continuously forecasts the state of the system.

The outputs of the model would be vehicle selection, mode choice, routing decision, internal costs (time, fuel consumption) and external costs (GHG, air polluting emissions, accidents, noise, congestion and infrastructure costs), etc. For the validation of the developed model, a case study is defined and will be implemented using real data in close collaboration with the partners of the DISpATch project. This model will be the First ever DT for long-distance flows as well as for Synchromodal transportation. The model will be a combination of different models that captures scales and subprocesses from a big picture; such as supply chain and synchromodal transport.

## References

[1] C. Dong, R. Boute, A. McKinnon, and M. Verelst. Investigating synchromodality from a supply chain perspective. *Transportation Research Part D: Transport and Environment*, 61:42–57, 2018.

[2] R. Giusti, D. Manerba, G. Bruno, and R. Tadei. Synchromodal logistics: An overview of critical success factors, enabling technologies, and open research issues. *Transportation Research Part E: Logistics and Transportation Review*, 129:92–110, 2019.

[3] L. Tavasszy, B. Behdani, and R. Konings. Intermodality and synchromodality. In *Ports and Networks*, pages 251–266. Routledge, 2017.

[4] VUB. Digital twin for synchromodal transport. `https://mobi.research.vub.be/en/dispatch-project-about`, 2020.

[5] H. Yee, J. Gijsbrechts, and R. Boute. Synchromodal transportation planning using travel time information. *Computers in Industry*, 125:103367, 2021.

# Author index