VVSOR OPERATIONS RESEARCH

ORBEL / THE BELGIAN OPERATIONAL RESEARCH SOCIETY

Joint ORBEL-NGB Conference

Maastricht, 29–31 January 2025

Booklet of Abstracts

Published by Maastricht University and ORBEL, The Belgian Operations Research Society

https://orbel.be/

January 2025

Contents

I Abstracts

1. Welcome to Maastricht!

The Joint ORBEL-NGB Conference on Operations Research is the 39th conference of ORBEL, the Belgian Operational Research (OR) Society. This year, it is organized together with their sister organization NGB, the Dutch Society on Operations Research. Both societies are member of EURO, the association of European OR Societies, and they are their national representative of IFORS (International Federation of OR Societies).

The conference is intended as a meeting place for researchers, users and potential users of Operational Research, Statistics, Computer Science and related fields. It will provide managers, practitioners, and researchers with a unique opportunity to exchange information on quantitative techniques for decision making.

This year's meeting will take place in Maastricht, the Netherlands, at Maastricht University School of Business and Economics, and will be organized by the department of Quantitative Economics.

2. ORBEL

ORBEL is the Belgian Society for the promotion of operational research and analytics. Its goal is to contribute to the science, the development, and the adoption of advanced tools, methods, and knowledge in these fields, by stimulating multidisciplinary research and scientific collaboration among its members, as well as fostering cooperation between academia, industry, and society in general, both on the national and the international level. Its main fields of interest include optimization, simulation, data science, and other quantitative methods for effective decision making in complex environments. ORBEL is a member of EURO, the Association of European Operational Research Societies, and IFORS, the International Federation of Operational Research Societies.

ORBEL / THE BELGIAN OPERATIONAL RESEARCH SOCIETY

3. NGB

The NGB (Nederlands Genootschap voor Besliskunde) is the Dutch society for Operations Research. It is a member of EURO and IFORS, just as ORBEL. Its goal is to promote the use of Operations Research in the Netherlands and to connect Operations Research professionals, scientist, teachers, and students to achieve this. Every year, it organizes a one-day seminar in cooperation with the Dutch Network on the Mathematics of Operations Research (LNMB). With this seminar, the NGB tries to accommodate both people from academia and industry, by showcasing how Operations Research can be used in several applications. The NGB also organizes the Young OR Day, a one-day event intended for its youngest, and potential new, members. This day includes a visit to a company that applies Operations Research on a daily basis. The NGB also gives financial support to seminars and events in the Netherlands to further promote the use of OR. In particular, we support the Joint ORBEL-NGB Conference on Operations Research in 2025.

VVSOR
OPERATIONS
RESEARCH

4. Committees

4.1 Organizing Committee

- Lissa Melis
- André Berger
- Mandy Dassen
- Dany Simon
- Tjark Vredeveld

4.2 Scientific Committee

- Aghezzaf, El Houssaine
- Arda, Yasemin
- Belien, Jeroen
- Benoit, Dries
- Blondeel, Wouter
- Braekers, Kris
- Caris, An
- Chevalier, Philippe
- Coussement, Kristof
- Crama, Yves
- De Baets, Bernard
- De Causmaecker, Patrick
- De Smet, Yves
- Defryn, Christof
- Fortz, Bernard
- Goossens, Dries
- Janssens, Gerrit
- Kunsch, Pierre
- Labbé, Martine
- Leus, Roel
- Mélot, Hadrien
- Melis, Lissa
- Papadimitriou, Dimitri

- Paquay, Célia
- Pironet, Thierry
- Sartenaer, Annick
- Schauss, Pierre
- Schyns, Michaël
- Sörensen, Kenneth
- Spieksma, Frits
- Tancrez, Jean-Sébastien
- Van Utterbeeck, Filip
- Vandaele, Arnaud
- Vanden Berghe, Greet
- Vansteenwegen, Pieter
- Verbeke, Wouter
- Vredeveld, Tjark
- Wauters, Tony
- Wittevrongel, Sabine

5. Practical Information

5.1 Presenting

Each presentation is scheduled for 20 minutes, including questions and discussion. Please adhere to this time frame.

Rooms are equipped with a computer attached to a beamer. Thanks in advance for using the provided computers as much as possible.

To ensure a timely progression of the session, we ask speakers to put their presentation files on the computers before the start of the session.

5.2 Chairing

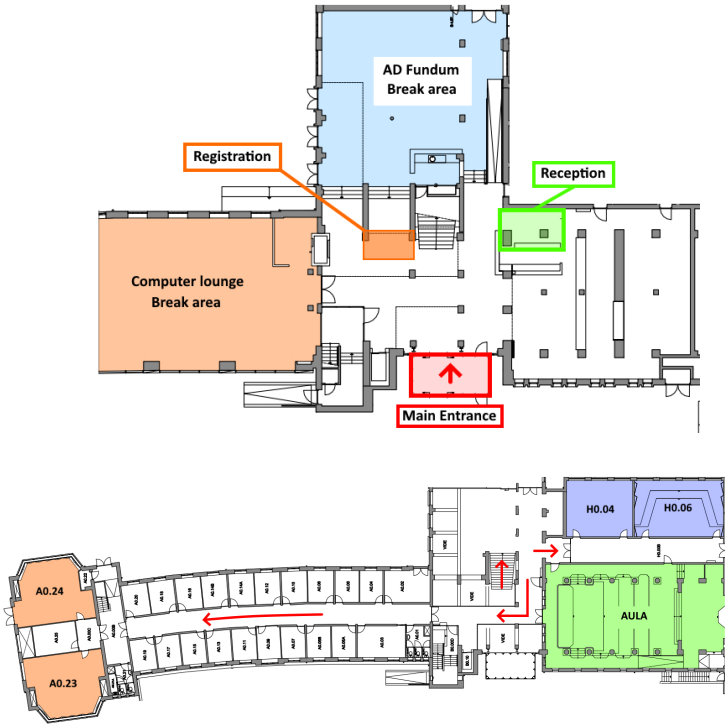
The session chairs are listed in the detailed program. The chair should briefly introduce the session topic as well as each of the speakers.

The chair should also remind the speakers to round off their talk when their time is almost up. If necessary, the chair may interrupt a speaker that exceeds their allotted time slot.

We do not encourage switching rooms between talks in a session. In case of no shows, the chair may distribute the extra time between the speakers.

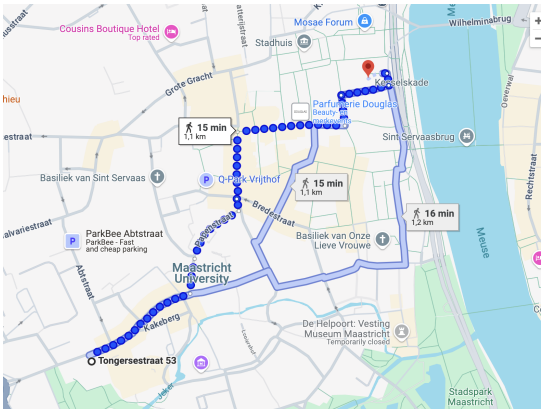
Venue

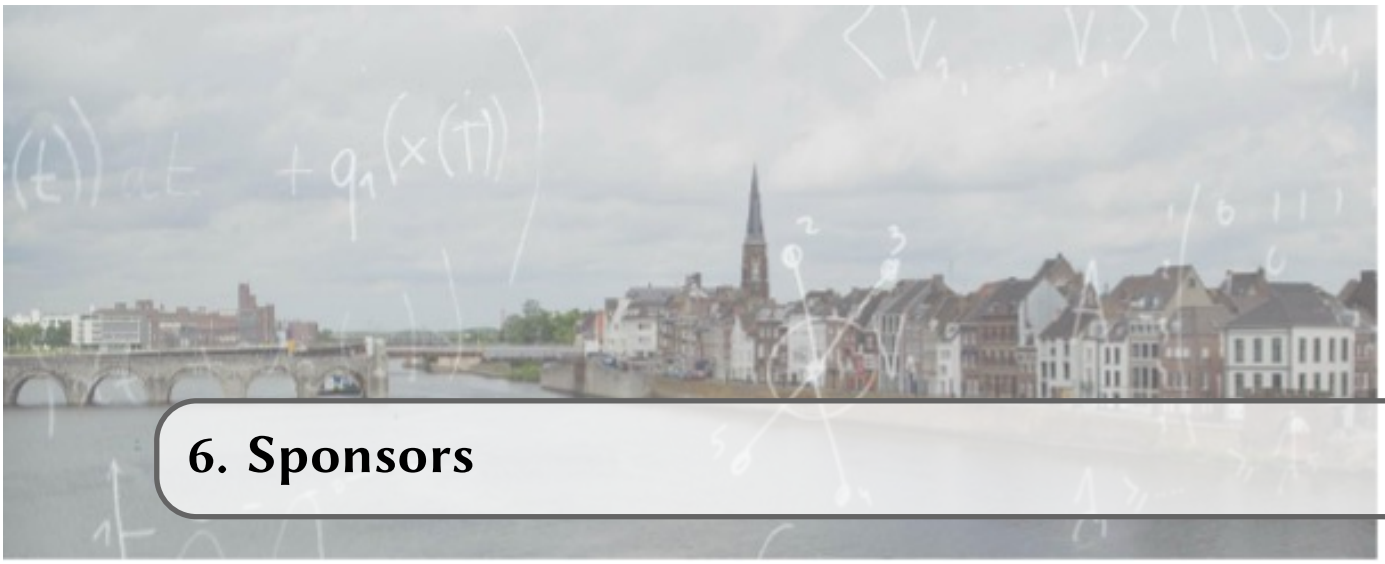
The conference is organized in Tongersestraat 53, the main building of Maastricht University School of Business and Economics. The building is a former Jesuit monastery. The plenary sessions will take place in the Aula of the building, which used to be the chapel of the monastery.



Conference dinner

The conference dinner will take place at Rebelle Maastricht, just a short walk from the conference location. We are welcomed at 18.30h and we are expected to take our seats at 19.00h.





6. Sponsors

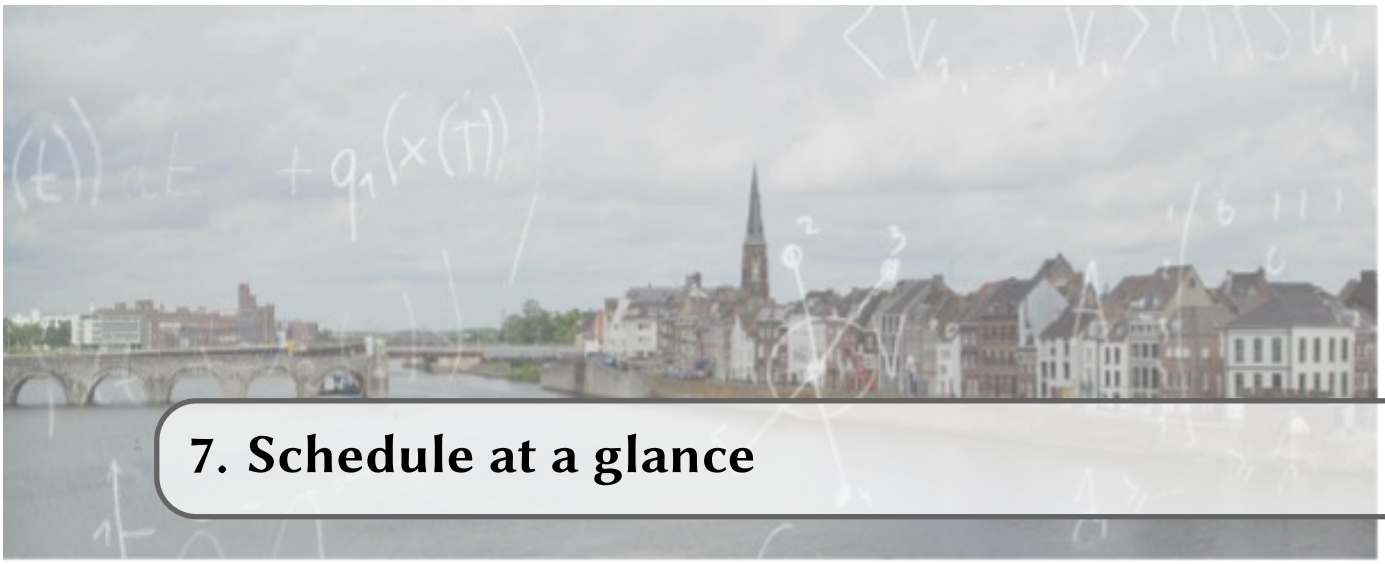
VVSOR OPERATIONS RESEARCH

Universiteitsfonds Limburg | SWOL |

hexaly

OMP.

GUROBI OPTIMIZATION



7. Schedule at a glance

Wednesday, 29 January 2025

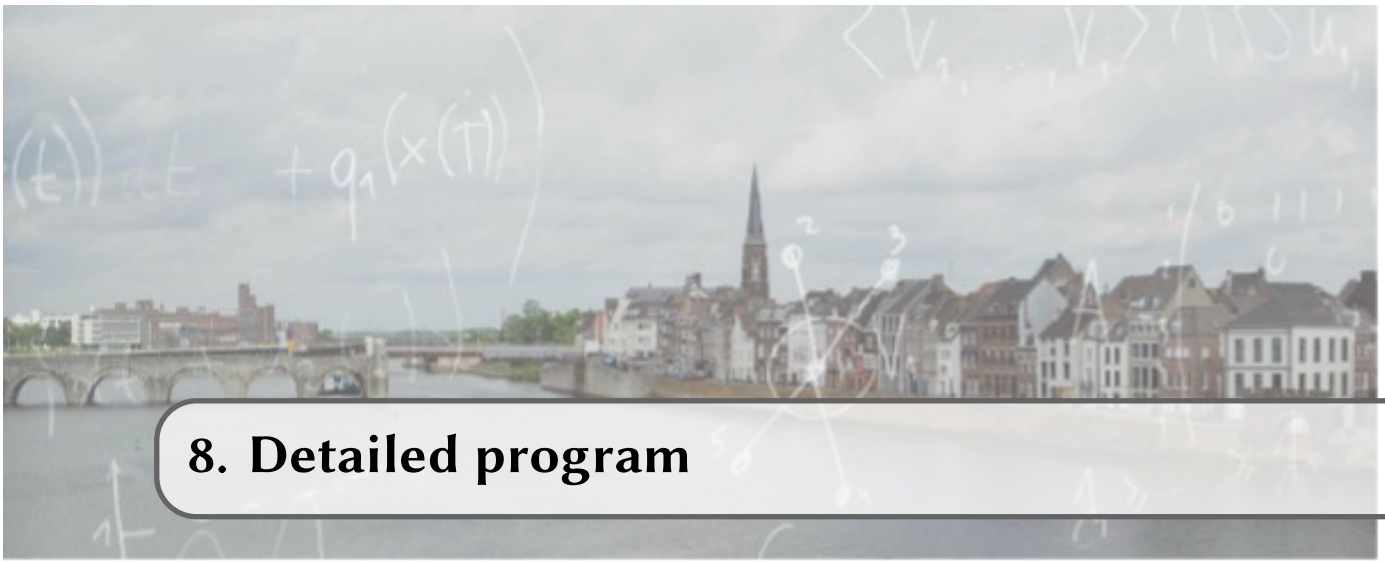
13:00 - 14:00	Registration and welcome coffee	Ad Fundum
14:00 - 15:15	Opening session & Keynote Lübbecke	Aula (H0.01)
15:15 - 15:45	Coffee break	Ad Fundum
15:45 - 16:45	Parallel sessions WA	A0.23, A0.24, H0.04, H0.06

Thursday, 30 January 2025

09:00 - 09:30	Welcome coffee	Ad Fundum
09:30 - 10:30	Opening session & Keynote Brintrup	Aula (H0.01)
10:30 - 11:15	Coffee break	Ad Fundum
11:15 - 12:35	Parallel sessions TA	A0.23, A0.24, H0.04, H0.06
12:35 - 14:00	Lunch	Ad Fundum
	ORBEL board meeting	A0.23
14:00 - 15:20	Parallel sessions TB	A0.23, A0.24, H0.04, H0.06
15:20 - 15:40	Coffee break	Ad Fundum
15:40 - 17:00	Parallel sessions TC	A0.23, A0.24, H0.04, H0.06
17:00 - 18:00	ORBEL general assembly	Aula (H0.01)
18:30 - 22:30	Conference dinner	Rebelle Maastricht

Friday, 31 January 2025

09:00 - 09:30	Welcome coffee	Ad Fundum
09:30 - 10:50	Parallel sessions FA	A0.23, A0.24, H0.04, H0.06
11:00 - 12:00	Parallel sessions FB	A0.23, A0.24, H0.04, H0.06
12:00 - 13:20	Lunch	Ad Fundum
13:20 - 14:20	Parallel sessions FC	A0.23, A0.24, H0.04, H0.06
14:20 - 14:45	Coffee break	Ad Fundum
14:45 - 15:45	Keynote Schöbel]	Aula (H0.01)
15:45 - 16:15	ORBEL award and closing session	Aula (H0.01)



8. Detailed program

8.1 Wednesday, 29 January 2025

13:00 - 14:00 **Registration and welcome coffee**
Ad Fundum

14:00 - 15:15 **Opening session**
Room: Aula (H0.01)

Welcome by Lissa Melis

Keynote Marco Lübbecke
A Dantzig-Wolfe Hierarchy for the Stable Set Problem
Session chair: André Berger

15:15 - 15:45 **Coffee break**
Ad Fundum

15:45 – 16:45 **Parallel session WA1 – Packing and volume maximization**
Session chair: J. Tollenaere
Room: A0.23

Binary linear programming formulations for a two-stage dual bin packing problem for wood reuse
Bessemans, P., Paquay, C., and Dumont, M. 39

Voxel-based, GPU-accelerated approaches for the 3D Irregular Strip Packing Problem
La Greca, A., Tollenaere, J., and Wauters, T. 40

Optimal continuous rotation in volume maximization problems through quaternions
Tollenaere, J., and Wauters, T. 41

15:45 – 16:45 **Parallel session WA2 – Mobility**
Session chair: S. Bayri
Room: H0.06

15:45 – 16:45 **Parallel session WA3 – Optimization**
Session chair: A. Lopez Martinez
Room: H0.04

15:45 – 16:45 **Parallel session WA4 – Healthcare personnel planning**
Session chair: A. Maharani
Room: A0.24

8.2 Thursday, 30 January 2025

9:00 - 9:30 **Welcome coffee**
Ad Fundum

9:30 - 10:30 **Keynote Alexandra Brintrup**
Artificial Intelligence for Nudging Complex Supply Networks
Session chair: Lissa Melis
Room: Aula (H0.01)

10:30 - 11:15 **Coffee break**
Ad Fundum

11:15 – 12:35 **Parallel session TA1 – Decision centric and fairness**
Session chair: J. Peeperkorn
Room: A0.23

11:15 – 12:35 **Parallel session TA2 – Public transport**
Session chair: T. De Munck
Room: H0.06

11:15 – 12:35 **Parallel session TA3 – History of algorithms**
Session chair: R. De Landtsheer
Room: H0.04

11:15 – 12:35 **Parallel session TA4 – Healthcare optimization problems**
Session chair: M. Vollebergh
Room: A0.24

12:35 - 14:00 **Lunch**
Ad Fundum
ORBEL board meeting
Room: A0.23

14:00 – 15:20 **Parallel session TB1 – Data science: Network and process**
Session chair: J. De Moor
Room: A0.23

14:00 – 15:20 **Parallel session TB2 – Logistics**
Session chair: Y. Wang
Room: H0.06

14:00 – 15:20 **Parallel session TB3 – Game theory and decision-making**
Session chair: Y. De Smet
Room: H0.04

14:00 – 15:20 **Parallel session TB4 – Crew scheduling**
Session chair: N. De Walsche
Room: A0.24

15:20 - 15:40 **Break**
Ad Fundum

15:40 – 17:00 **Parallel session TC1 – Data science: NLP**
Session chair: P. Chuor
Room: A0.23

15:40 – 17:00 **Parallel session TC2 – Logistics**
Session chair: O. Elsherif
Room: H0.06

15:40 – 17:00 **Parallel session TC3 – Scheduling**
Session chair: L. Petit-Jean Genat
Room: H0.04

15:40 – 17:00 **Parallel session TC4 – Complexity and algorithm analysis**
Session chair: S. Miltenburg
Room: A0.24

17:00 - 18:00 **ORBEL general assembly**
Room: Aula (H0.01)

18:30 **Conference dinner**
Rebelle Maastricht

8.3 Friday, 31 January 2025

9:00 - 9:30 **Welcome coffee**
Ad Fundum

9:30 – 10:50 **Parallel session FA1 – Data science: Interpretability**
Session chair: S. De Lange
Room: A0.23

9:30 – 10:50 **Parallel session FA2 – Warehousing and inventory**
Session chair: A. Robbes
Room: H0.06

9:30 – 10:50 **Parallel session FA3 – Sports scheduling**
Session chair: K. Devriesere
Room: H0.04

9:30 – 10:50 **Parallel session FA4 – Optimization**
Session chair: C. Renkin
Room: A0.24

11:00 – 12:00 **Parallel session FB1 – Forecasting and uncertainty**
Session chair: A. Leribaux
Room: A0.23

11:00 – 12:00 **Parallel session FB2 – Mobility**
Session chair: B. Coulier
Room: H0.06

11:00 – 12:00 **Parallel session FB3 – Assembly lines**
Session chair: G. Gündüz Mengübaş
Room: H0.04

11:00 – 12:00 **Parallel session FB4 – Data science: NLP**
Session chair: A. Thuy
Room: A0.24

12:00 - 13:20 **Lunch**
Ad Fundum

13:20 – 14:20 **Parallel session FC1 – Machine Learning**
Session chair: Y. De Rocker
Room: A0.23

13:20 – 14:20 **Parallel session FC2 – Orbel Award Nominees**
Session chair: R. Leus
Room: H0.06

Minimizing travel duration in inland shipping through strategic departure time scheduling
R. Annaert

A Synthetic Data Based Simulation Approach to Enhance the Spatial Allocation of Police Resources
N. Prakopetz

Electric vehicle routing problems: hoe omspringen met beperkte batterij-capaciteit
O. El Habti

13:20 – 14:20 **Parallel session FC3 – Matchings**
Session chair: R. Zuidwijk
Room: H0.04

13:20 – 14:20 **Parallel session FC4 – Nurse scheduling**
Session chair: F. Nießen
Room: A0.24

14:45 - 15:45 **Keynote Anita Schöbel**
Sustainable public transport
Session chair: Marjan van den Akker
Room: Aula (H0.01)

15:45 - 16:15 **ORBEL Award and closing session**
Room: Aula (H0.01)

9. Plenary talks

9.1 Artificial Intelligence for Nudging Complex Supply Networks

Alexandra Brintrup, University of Cambridge, UK

This talk will introduce participants to the field of Artificial Intelligence in Operations and Supply Chain Management. We will first talk about the state of affairs and major driving forces shaping supply chain today, to motivate the data driven era we are in. Then AI is introduced with multiple definitions, to cover what is AI and importantly, what is not AI. We introduce sub-fields of AI and data science, and how they are primarily used in supply chain management. We then delve deeper into an "exotic" selection of supply chain AI, deliberately so, in order to emphasise, that which could not have been done before.

This then brings us to state of the art research examples in network analytics, digital supply chain surveillance, collective-learning and distributed decision making and automation. Our aim is to encourage debate on how AI should be evaluated by breaking disciplinary siloes in the OM community. We will then discuss the potential pitfalls and challenges, such as loss of data traceability, complacency, lack of accountability, and cognitive atrophy. The talk concludes with supply chain management needing to become an irrevocably interdisciplinary field with challenges so varied and significant.

Alexandra Brintrup is Professor in Digital Manufacturing at the University of Cambridge's Engineering Department, where she leads the Supply Chain AI Lab. She also leads Digital Manufacturing at the Alan Turing Institute, is external faculty at the Complexity Science Hub Vienna, and is a fellow of Darwin College.

Prof. Brintrup was the first researcher to empirically study large-scale supply chains as complex adaptive networks, examine their emergent properties, and take a data-driven perspective to characterise their resilience, which led to understanding of universal patterns that govern supply chains. She was also the first to develop algorithms to predict supply chain dependencies and disruptions. Over the past decade she advised policy makers, and national and European scientific committees, and worked with both start ups , SMEs and international organisations. She is a member of the All Party Parliamentary Groups in Artificial Intelligence and Data Analytics, and advises policy development in supply chain risk, economic performance and resilience. Her current research includes: Predictive methods for automated detection of supply chain dependencies, especially with collective learning paradigms; complex system approaches to model emergence in supply networks, autonomous and scalable optimisation and distributed decision making technologies, particularly with nature-inspired algorithms and Multi-agent Systems.

9.2 A Dantzig-Wolfe Hierarchy for the Stable Set Problem

Marco Lübbecke, RWTH Aachen, Germany

The stable set problem is an important subproblem in many practical applications. It has been studied extensively also from a theoretical point of view. In particular, the problem is a poster child for the study of the strength of (the relaxations of) many different formulations for the problem. Dantzig-Wolfe reformulation is a general technique for obtaining strong formulations for a problem. We first motivate connecting the two topics by some computational observations about Dantzig-Wolfe reformulations of a classical ("edge") formulation for the stable set problem. For this model, each reformulation can be associated with a subset of the edges of the original graph, thus a subgraph. In fact, we have exponentially many options for selecting such a subgraph and we are able to characterize which reformulations are strongest and weakest possible. Only recently, we fully understood also the reformulations in between the extremes. More precisely, we can show that any two reformulations lead to identical polyhedra if and only if the corresponding reformulated subgraphs contain the same odd induced cycles. This lends itself to the notion of a hierarchy of Dantzig-Wolfe refomulations which is the first such concept in the Dantzig-Wolfe context.

This is joint work with Michasel Bastubbe, Adrian Gallus, and Jonas Witt.

Marco Lübbecke is a full professor and chair of operations research at RWTH Aachen University, Germany. He received his Ph.D. in applied mathematics from TU Braunschweig in 2001 and held positions as assistant professor for combinatorial optimization and graph algorithms at TU Berlin and as visiting professor for discrete optimization at TU Darmstadt. Marco's research and teaching interests are in computational integer programming and discrete optimization, covering the entire spectrum from fundamental research and methods development to industry scale applications. A particular focus of his work is on decomposition approaches to exactly solving large-scale real-world optimization problems. This touches on mathematics, computer science, business, and engineering alike and rings with his appreciation for fascinating interdisciplinary challenges.

9.3 Sustainable public transport

Anita Schöbel, Fraunhofer IWTM and RPTU Kaisserslautern-Landau, Germany

Moving travelers efficiently, with low costs, and respecting environmental goals like CO2 emissions is one of the challenging problems our society faces today. In this talk we sketch how optimization approaches can help to provide sustainable transport which reaches these goals.

A first goal is to make public transport more attractive such that travelers choose bus, tram, metro, or railway transportation instead of using their own cars. Here we focus on minimizing travel times, minimizing transfers and increasing reliability of public transport. In this context, we show that integrating different planning stages may help to further improve efficiency of public transport modes. A second step is to make public transport modes themselves more energy-efficient. This can be done by using electric vehicles, but also by planning energy-efficient lines or using regenerative energy when designing a train timetable. Finally, we argue that for providing sustainable transport for the society, we need also look at other modes of transport besides regular bus or metro transportation. This includes demand-responsive transport, individual cars, sharing modes and active modes like walking and biking. We sketch a first model in which such different transport modes are considered simultaneously.

Anita Schöbel received her PhD in mathematics in 1998 and her habilitation in mathematics in 2003 in Kaiserslautern. From 2004 - 2018 she was professor for Optimization at the Georg-August University Göttingen. Since January 2019, Anita Schöbel has been professor for Applied Mathematics at the University of Kaiserslautern-Landau and director of the Fraunhofer Institute for Industrial Mathematics ITWM. The ITWM has 550 employees organized in 11 departments all dealing with industrial research projects in applied mathematics. At Fraunhofer Society, Anita Schöbel is responsible for the strategic research area Next Generation Computing and vice spokesperson for the quantum computing competence network.

In 2019 and 2020 she has been president of the German Operations Research Society (GOR). Currently, she is president of EURO (Association of European Operational Research Societies). She is also member of the university steering committee (Hochschulrat) Kaiserslautern-Landau, in the senate of the national research data infrastructure (NFDI), and in the council for technology of the state Rhineland-Palatinate.

She has been involved in many industrial and research projects, among them the European projects ARRIVAL, EASIER and OptALI and cooperations with India. She also coordinated a research unit (Forschungsgruppe) on Integrated Transportation funded by the German Research Foundation (DFG) and currently coordinates a ministry-funded project on synchronizing different modes of transport.

In her research interests, Anita focuses on discrete optimization in public transport, robust and multi-objective robust optimization, quantum computing, and on several topics related to continuous facility location.

Abstracts

Binary linear programming formulations for a two-stage dual bin packing problem for wood reuse

Pauline Bessemans
University of Liège, Liège, Belgium, QuantOM, HEC Liège - Management School
e-mail: pauline.bessemans@uliege.be

Célia Paquay
University of Liège, Liège, Belgium, QuantOM, HEC Liège - Management School
e-mail: cpaquay@uliege.be

Morgane Dumont
University of Liège, Liège, Belgium, QuantOM, HEC Liège - Management School
e-mail: morgane.dumont@uliege.be

The increasing demand for raw materials such as wood is undoubtedly contributing to the depletion of natural resources and global warming. To curb this phenomenon, a more sustainable and circular wood management could be developed by intelligently handling wood waste. This wood waste can be in the form of beams or pallets and could be considered as wooden slats. They could be combined, assembled, and glued to build Cross-Laminated Timber (CLT) panels for the construction industry.

We aim to develop optimization techniques to recycle raw wood waste by providing assembly schemes to create CLT panels. The goal is to minimize the waste, which is the wood that could not be reused in the CLT panels.

We conducted a literature review to identify the closest problems in the field of operations research and to name our problem accordingly. The skiving stock problem and the dual bin packing problem, which are not dual versions of the cutting stock/bin packing problems, are the two closest problems. The present work addresses for the very first time an exact case of the two-stage two-dimensional dual bin packing problem (E-2S-2D-DBPP) in the context of wood reuse.

We propose a description of the problem and two mathematical formulations with cuts. Then, we compare them and present the results of several numerical experiments based on realistic instances from the wood industry. We also identify the size limit of the instances for which the problem can still be solved in a reasonable amount of time.

Voxel-based, GPU-accelerated approaches for the 3D Irregular Strip Packing Problem

Alessio La Greca
KU Leuven University, Ghent Campus
e-mail: alessio.lagreca@kuleuven.be

Jonas Tollenaere
KU Leuven University, Ghent Campus
e-mail: jonas.tollenaere@kuleuven.be

Tony Wauters
KU Leuven University, Ghent Campus
e-mail: tony.wauters@kuleuven.be

The 3D Irregular Strip Packing Problem is a generalization of the NP-hard strip packing problem. The problem involves a three-dimensional rectangular container of fixed width, fixed depth and infinite height. There is also a set of three-dimensional irregular items, each with its own set of allowed rotations. The aim is to find a way to pack all the items in the container while minimizing the total height required to do so. Given the complexity of the problem, exact approaches typically fail. This is why various (meta)heuristics have been proposed in the literature, a key metric of which is how fast they can determine the height of a given solution. A solution can be built using a constructive heuristic by defining an ordering of the items and a rule for placing them into a partially-filled container. Most of the complexity of the problem comes from the geometric tests performed, which determine whether or not two items in a certain placement overlap. One way to deal with such irregular, 3D items is to approximate them with voxels, which discretize their representation into a set of unit cubes of predefined size. The main drawback of this approach concerns the granularity of the discretization: the more voxels that are used to approximate a shape, the better. However, more voxels means more geometric tests must be performed to check for overlapping. Our work aims to accelerate the operations performed on a discrete, voxel-based representation of the items. This is done with the aid of the CUDA toolkit, a parallel computing platform and API used to develop general-purpose computing applications on NVIDIA GPUs. By constructing a feasible solution as quickly as possible, our constructive heuristic can then be used as a building block for other heuristics and metaheuristics. Preliminary computational results already demonstrate the potential of this GPU voxel-based approach.

Optimal continuous rotation in volume maximization problems through quaternions

Jonas Tollenaere
KU Leuven, Department of Computer Science
e-mail: jonas.tollenaere@kuleuven.be

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

In a volume maximization problem, the goal is to find the largest item(s) of variable scale that can be extracted from a larger object, the container. The item's continuous translation and rotation within the three-dimensional container is optimized to enable maximal scaling of the item. Naturally, the item should respect containment and non-overlapping constraints as the item and container represent physical objects. After all, this optimization problem is a type of cutting and packing problem. It appears in industries such as gem cutting, where improving the yield is typically highly valuable. While several heuristic solution methods have been proposed in previous literature, exact methods have not been thoroughly explored. The primary cause for this is probably the complexity of mathematically optimizing rotation in three-dimensional space. Euler angles or other angle-based representations lead to complex trigonometric expressions, which are challenging for exact solution methods. Therefore, quaternions are better suited for this purpose. Using quaternions, we can formulate the orientation of an object through quadratic expressions. This allows us to mathematically formulate volume maximization problems with convex containers as quadratically constrained programs, based on quaternions, that can be solved exactly. In this presentation, we will introduce volume maximization problems, give some insights into its applications and heuristic methods, and formulate the aforementioned quadratically constrained models. Our experiments show that, when a single item in a convex container is considered, these models can be effectively solved to optimality within minutes. However, the concave cases prove to be much more challenging.

Approximate Dynamic Programming for the Assignment Problem in Crowdshipping

Emma Innocente
Université catholique de Louvain, Center for Operations Research and Econometrics (CORE)
e-mail: emma.innocente@uclouvain.be

Jean-Sébastien Tancrez
Université catholique de Louvain, Center for Operations Research and Econometrics (CORE)
e-mail: js.tancrez@uclouvain.be

In recent years, crowdshipping has emerged as a potentially cost-effective and sustainable alternative to traditional delivery [6], though it often lacks deeper analysis of its practical implementation and broader impacts. Crowdshipping is a collaborative delivery system that falls within the sharing economy trend in that deliveries are carried out by using existing resources (vehicle capacity and drivers from the crowd, i.e. non-professionals). The so-called crowdshippers deliver shipments by taking a detour from their planned trips [1] in return for a compensation fee. By exploiting excess capacity in existing personal transport for the delivery of freight, crowdshipping is inclined to reduce the number of dedicated freight delivery trucks, reducing delivery costs and supporting sustainability [5].

Crowdshipping can manifest in various forms; however, this work specifically focuses on peer-to-peer crowdshipping [2] since it is less studied than crowdshipping with in-store customers, yet subject to more uncertainty. It employs regular commuters as crowdshippers for the delivery of parcels with sellers, buyers and crowdshippers geographically dispersed. Peer-to-peer crowdshipping is suitable for delivery in C2C platforms (e.g. *Vinted*). The French crowdshipping company *Cocolis* applies this model.

The successful implementation of crowdshipping heavily depends on its operational efficiency, with the matching of crowdshippers and parcels at its core [3]. Parcels need to be assigned to crowdshippers dynamically on a daily basis. Most research does not directly address the challenges of uncertainty in the availability of crowdshippers and parcels, but rather assumes perfect information. Yet, crowdshipping operations are plagued by uncertainty since crowdshippers decides when to be available, introducing uncertainty as it becomes difficult to accurately predict worker availability [4]. The future availability of crowdshippers and parcels being uncertain, it implies that delivering a parcel now may lead to missing a better-suited crowdshipper later. As a result, current assignment decisions influence future ones. In this setting, the dynamic assignment of crowdshippers to parcels over time is challenging as present assignments affect future ones, and as the arrival of crowdshippers and parcels in the future is uncertain. Peer-to-peer crowdshipping has not been studied within the context of matching crowdshippers and parcels while considering the dynamic and uncertain aspects

of the problem.

This work presents an approximate dynamic programming (ADP) algorithm based on value function approximation [7] to assign crowdshippers to parcels over time while considering the uncertainty in crowdshippers and parcels arrivals. Our algorithm learns value functions offline adaptively, capturing the future impact of decisions in an estimated value. This approach provides non-myopic behavior by considering the future impact of current decisions, helping determine whether to postpone deliveries in anticipation of better-suited crowdshippers becoming available later. Yet, it only requires solving sequences of assignment problems no larger than would be required with a myopic algorithm. The characteristics of the crowdshipping model under study required adapting the ADP algorithm to address its specific needs. Unlike prevailing practices, where parcel value estimates depend on the time within the operational horizon, the estimates in this approach rely solely on parcel attributes and have a delivery window, thereby limiting their impact in time.

Through numerical results, we demonstrate our ADP approach's effectiveness compared to a myopic approach, which makes decisions based solely on available information. Our approach increases the total contribution of crowdsourced delivery by 4% on average and up to 20.5% and reduces crowdshippers' detours by 25% on average compared to a myopic approach. Computational experiments also highlight the parameters affecting the relevance of implementing a dynamic assignment method.

References

[1] Archetti, C., Savelsbergh, M., and Speranza, M. G. (2016). The Vehicle Routing Problem with Occasional Drivers. *European Journal of Operational Research*, 254(2):472-480.

[2] Buldeo Rai, H., Verlinde, S., and Macharis, C. (2018). Shipping outside the box. Environmental impact and stakeholder analysis of a crowd logistics platform in Belgium. *Journal of Cleaner Production*, 202:806-816.

[3] Le, T. V., Stathopoulos, A., Van Woensel, T., and Ukkusuri, S. V. (2019). Supply, demand, operations, and management of crowd-shipping services: A review and empirical evidence. *Transportation Research Part C: Emerging Technologies*, 103:83-103.

[4] Mousavi, K., Bodur, M., Cevik, M., and Roorda, M. J. (2024). Approximate dynamic programming for pickup and delivery problem with crowd-shipping. *Transportation Research Part B: Methodological*, 187:103027.

[5] Pourrahmani, E. and Jaller, M. (2021). Crowdshipping in last mile deliveries: Operational challenges and research opportunities. *Socio-Economic Plan- ning Sciences*, 78:101063.

[6] Punel, A., Ermagun, A., and Stathopoulos, A. (2018). Studying determinants of crowd-shipping use. *Travel Behaviour and Society*, 12:30-40.

[7] Spivey, M. Z. and Powell, W. B. (2004). The Dynamic Assignment Problem. *Transportation Science*, 38(4):399-419.

Nash flows over time with tolls

Shaul Rosner
Tel Aviv University

Marc Schröder
Maastricht University

Laura Vargas Koch
RWTH Aachen University

One of the most popular theoretical models in recent years for traffic modelling is the deterministic fluid queuing model. This dynamic model is given by a directed graph in which each edge is endowed with a fixed travel time and a capacity per time unit. When the inflow into an edge exceeds its capacity, a queue starts to build that has a service rate equal to the capacity of the edge. The travel time of a particle on an edge is equal to the fixed travel time plus the waiting time in the queue.

In order to describe traffic congestion in this model realistically, it is essential to take into account the underlying selfish behavior of traffic participants. After all, everyone aims to arrive at their destination as fast as possible given the network state induced by the other traffic participants. It is thus crucial to take a game theoretic perspective, where every infinitesimal particle corresponds to a player minimizing its travel time. A network flow over time in which no particle has an incentive to deviate from its route, given the flow induced by other particles, is called a *dynamic equilibrium*.

This line of research dates back to the work of Vickrey [1], but since then there have been elaborate efforts to theoretically understand congestion dynamics. Koch and Skutella [2] characterized dynamic equilibria by means of concatenations of static thin flows, which connect the dynamic model to the well-known static setting. The existence of dynamic equilibria, even in multicommodity networks, has been proven formally by Cominetti et al. [3], who also showed that equilibria are basically unique. Cominetti et al. [4] proved that after finite time the network reaches a steady state in which queues remain stable under the assumption that the network inflow rate is constant and does not exceed the minimum cut capacity. This latter assumption is relaxed by Olver et al. [5].

We introduce fixed tolls to the deterministic fluid queuing model and investigate whether some of the known properties of dynamic equilibria carry over to the setting with tolls. We assume that particles' costs are equal to the sum of travel time and tolls. Using three examples we show that (1) dynamic equilibria with tolls need not be unique, (2) particles might overtake in dynamic equilibria with tolls, (3) dynamic equilibria with tolls need not reach a steady state, but seem to converge to one. We finally have a more closer look at the characterization of steady states for dynamic equilibria with tolls.

References

[1] Vickrey, W (1969), "Congestion theory and transport investment," *The American Economic Review*.

[2] Koch, R and Skutella, M (2011), "Nash equilibria and the price of anarchy for flows over time," *Theory of Computing Systems*.

[3] Cominetti, R and Correa, J and Larré, O (2015), "Dynamic Equilibria in Fluid Queueing Networks," *Operations Research*.

[4] Cominetti, R and Correa, J and Olver, N (2021), "Long-Term Behavior of Dynamic Equilibria in Fluid Queuing Networks," *Operations Research*.

[5] Olver, N and Sering, L and Vargas Koch, L (2022), "Continuity, Uniqueness and Long-Term Behavior of Nash Flows Over Time," *Annual Symposium on Foundations of Computer Science*.

The Bike-Sharing Station Location Problem in Hybrid Urban Mobility Systems

Soukaina Bayri
Hasselt University, Research Group Logistics
Vrije Universiteit Brussel, Research Group Mobilise
e-mail : soukaina.bayri@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

Bike-sharing systems (BSS) have gained significant popularity in recent years due to their potential contribution to sustainable urban development. They have become the most widespread implementation of the concept of Mobility-as-a-Service (MaaS) (Vallez et al. 2021). BSS offer numerous benefits including emission and congestion reduction, improved public health, and enhanced traffic systems. Furthermore, bike-sharing can also be a good first- or last-mile solution and enhance user's connection with public transport (PT) allowing the creation of hybrid mobility systems. These are systems where users travel through multimodal-trips (Zhang and Zhang 2018 ; Caggiani et al. 2020). Despite the growing interest in hybrid systems, the current literature lacks a methodology to optimize the integration of BSS and PT. One of the most crucial factors for the successful implementation of these systems is the strategic placement and capacity of the bike-sharing stations in relation to the PT-network. Poorly located bike-sharing stations compromise the success of the overall system. Therefore, in this research, we design a BSS given a PT-network in order to fully optimize the benefits of their integration.

In this talk, we will present an optimization model designed to determine the optimal location of bike-sharing stations, taking into account an existing fixed PT-network. This model combines strategic decisions (i.e. location and capacity of BS stations) with operational considerations (i.e. relocation of the bikes). The approach considers an area divided into multiple zones, and demand (i.e. number of requested trips) between these zones. The aim is to design a bike-sharing network that complements the existing public transport system, maximizing the demand coverage. The goal is to fulfill as many requested trips as possible through bike-sharing, PT or a combination of both modes, considering a specified budget.

Interacting with Uncertainty: An integrated simulated annealing algorithm for model-robust choice designs with partial profiles

Yicheng Mao
Maastricht University, Department of Data Analytics and Digitalization
e-mail: yicheng.mao@maastrichtuniversity.nl

Roselinde Kessels
Maastricht University, Department of Data Analytics and Digitalization
University of Antwerp, Department of Economics
e-mail: r.kessels@maastrichtuniversity.nl

Discrete Choice Experiments (DCEs) investigate participants' preferences by observing their choice behavior in hypothetical scenarios and are widely used in numerous domains. To reduce participants' cognitive burden, especially when dealing with a large number of attributes, researchers often employ partial profile designs. In these designs, certain attributes within each choice set are kept constant. Current literature on partial profile designs mainly focuses on main-effects models rather than interaction-effects models, with some partial profile designs even incapable of estimating interaction effects. To address this issue, this paper introduces an integrated Simulated Annealing (SA) algorithm to construct partial profile designs based on an interaction-effects model. During the experimental design phase, however, the existence and magnitude of interaction effects are often unknown. Therefore, this paper proposes a model-robust experimental design strategy. Through extensive computational experiments and a real-life case study, we demonstrate that our SA model-robust partial profile design performs relatively well regardless of the underlying model.

Towards Quantum Annealing for Multi-League Sports Scheduling

Orin Pechler
Maastricht University, School of Business and Economics
e-mail: of.pechler@gmail.com

Frank Phillipson
Maastricht University, School of Business and Economics
TNO, Applied Cryptography and Quantum Algorithms
e-mail: f.phillipson@maastrichtuniversity.nl

1 Introduction

This work introduces the use of quantum annealing for the Multi-League Scheduling Problem (MLSP), assuming that all leagues have the same even number of teams. The goal is to develop a match schedule for multiple leagues comprising various teams and clubs, a particularly relevant challenge in amateur and youth sports. For this scheduling problem, the primary objective is to formulate it as a QUBO (Quadratic Unconstrained Binary Optimization) problem, the key formulation type for a quantum annealer. Scheduling matches for sports leagues is a complex task, especially in amateur competitions that involve hundreds of leagues, clubs, and teams, shared venues and locations, high volume of matches, capacity restrictions, and interdependence between leagues. This work builds on assumptions, results, and formulations from [1,2].

2 Quantum Annealing and QUBO Modeling

Quantum annealing (QA) is an optimization technique that uses quantum computers to find the lowest energy configuration of a system. This principle can be applied to optimization problems by converting them into energy minimization problems. To apply QA to MLSP, the problem must be formulated as a QUBO model. QUBO models consist of binary variables and an objective function to minimize. When converting a linear program into a QUBO, some functions in the QUBO represent the constraints or objectives of the original problem. Violating a constraint increases the objective function, indirectly enforcing constraint compliance.

3 The Multi-League Scheduling Problem (MLSP)

The MLSP focuses on finding a schedule that minimizes the total number of capacity violations at clubs. A capacity violation occurs when more matches are scheduled at a club location in a given round than its capacity allows. The MLSP uses Home and Away Patterns (HAPs) to indicate whether a team plays at home or away in each round. A HAP is a sequence of symbols indicating a teams home ("H") or away ("A") status in each round. A set of HAPs, or HAPset, represents all teams in a league. A HAPset is feasible if a match schedule exists that aligns with it.

4 QUBO Formulations for MLSP

The work develops four distinct QUBO formulations for the MLSP, using various techniques to model the constraints and objective function: [1] One-hot encoding, which uses binary variables to represent discrete values, [2] Domain-wall encoding, a more efficient method requiring fewer qubits and interactions, [3] Binary encoding, which represents numerical values using binary variables, and [4] Unbalanced penalization, which applies an approximate penalty for capacity violations without requiring additional variables.

5 Results and Future Work

The QUBO formulations were implemented and solved using both simulated annealing and quantum annealing on the D-Wave Advantage quantum computer. Their performance was compared based on solution quality and computation time. Domain-wall encoding delivered the best solution quality but required the longest computation time. Unbalanced penalization had the shortest computation time while producing a solution quality only marginally worse than domain-wall encoding. Although quantum annealing has not yet outperformed classical approaches, quantum computers are expected to become a superior alternative in the future.

Future research could focus on extending QUBO formulations to allow for varying league sizes and start times and on optimizing encoding and problem embedding on quantum annealers. The results of this study demonstrate that quantum annealing is a promising technique for solving complex scheduling problems in sports. As technology evolves, quantum annealing is likely to play an increasingly important role in optimizing sports schedules and other challenging optimization problems.

6 References

[1] Li, M., Davari, M., & Goossens, D. (2023). Multi-league sports scheduling with different leagues sizes. European Journal of Operational Research, 307(1), 313-327.

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

Finding Diverse Solutions in Combinatorial Problems with a Distributive Lattice Structure

Mark de Berg
Eindhoven University of Technology, Department of Mathematics & Computer Science
e-mail: m.t.d.berg@tue.nl

Andrés López Martínez
Eindhoven University of Technology, Department of Mathematics & Computer Science
e-mail: a.lopez.martinez@tue.nl

Frits Spieksma
Eindhoven University of Technology, Department of Mathematics & Computer Science
e-mail: f.c.r.spieksma@tue.nl

In combinatorial optimization problems, the objective is typically to identify a single optimal solution. However, this approach may be inadequate or impractical in real-world situations, where some constraints and factors are often overlooked or unknown in advance. This motivates the development of algorithms capable of finding multiple solutions, with *diversity* playing a key role. A growing body of research has focused on finding diverse solutions in classical combinatorial problems, much of it emerging from the field of fixed-parameter tractability [BJM$^+$19, FGJ$^+$24, HKK$^+$23]. These studies show that finding diverse solutions is generally more challenging than finding a single one. For instance, while Matching is solvable in polynomial time, finding two edge-disjoint matchings is NP-hard, even on 3-regular graphs [FGJ$^+$24].

This work aims to design theoretically efficient algorithms for constructing collections of maximally diverse solutions. Using the sum of pairwise Hamming distances between solutions as a diversity measure, we show that a broader class of problems is computationally no harder than finding a single solution in polynomial time. Specifically, we generalize the polynomial-time solvability of k-Diverse Minimum s-t Cuts by De Berg et al. [dBMS23] to a class of combinatorial problems whose solution sets form a distributive lattice.

We state our main result in terms of a unified general problem: Max-Sum k-Diverse Solutions. Let E be a finite set with n elements, and let $\Gamma \subseteq 2^E$ be a set of feasible solutions. For two feasible solutions $X, Y \in \Gamma$, the symmetric difference, or Hamming distance, between them is defined as $X \triangle Y = (X \setminus Y) \cup (Y \setminus X)$. Let $(X_1, X_2, \ldots, X_k)$ be a collection of k subsets of E. We consider the *pairwise-sum* diversity measure: $d_{\text{sum}}(X_1, X_2, \ldots, X_k) = \sum_{1 \le i < j \le k} |X_i \triangle Y_i|$.

Adopting the notation from Hanaka et al. [HKK$^+$23], we define Max-Sum k-Diverse Solutions as follows:

Max-Sum k-Diverse Solutions. *Given a finite set E, a positive integer k, and*

a membership oracle for $\Gamma \subseteq 2^E$*, find a* k*-element multiset* $C = (X_1, X_2, \ldots, X_k)$ *with* $X_1, X_2, \ldots, X_k \in \Gamma$*, such that* $d_{\text{sum}}(C)$ *is maximum.*

Now suppose that all sets in Γ have the same cardinality r, and that $(E, \leq)$ is a poset formed by the disjoint union of r chains $(E_1, E_2, \ldots, E_r)$. The cartesian product $E_{\text{prod}} = E_1 \times E_2 \times \cdots \times E_r$, ordered component-wise, forms a distributive lattice. We prove the following result.

Theorem 1. *If the set of feasible solutions* Γ *under the component-wise order is a sublattice* L *of* E_{prod}*, then* Max-Sum k-Diverse Solutions *can be solved in polynomial time in* n*, given a compact representation of* L*.*

We achieve this result via a reduction to the *submodular function minimization* problem (SFM) on a distributive lattice, which is known to be solvable in polynomial time [GLS12, IFF01, Sch00]. More precisely, we show that the pairwise-sum measure (reformulated as a minimization objective) is a submodular function on a distributive lattice of appropriately ordered k-sized collections of feasible solutions. Using this approach, we obtain polynomial-time algorithms for finding maximally diverse k-sized collections of stable matchings, α-circulations and flows in planar graphs, while also reproducing the findings of De Berg et al. for minimum s-t cuts.

References

[BJM+19] Julien Baste, Lars Jaffke, Tomáš Masařík, Geevarghese Philip, and Günter Rote. Fpt algorithms for diverse collections of hitting sets. *Algorithms*, 12(12):254, 2019.

[dBMS23] Mark de Berg, Andrés López Martínez, and Frits Spieksma. Finding diverse minimum st cuts. In *34th International Symposium on Algorithms and Computation (ISAAC)*, 2023.

[FGJ+24] Fedor V Fomin, Petr A Golovach, Lars Jaffke, Geevarghese Philip, and Danil Sagunov. Diverse pairs of matchings. *Algorithmica*, 86(6):2026–2040, 2024.

[GLS12] Martin Grötschel, László Lovász, and Alexander Schrijver. *Geometric algorithms and combinatorial optimization*, volume 2. Springer Science & Business Media, 2012.

[HKK+23] Tesshu Hanaka, Masashi Kiyomi, Yasuaki Kobayashi, Yusuke Kobayashi, Kazuhiro Kurita, and Yota Otachi. A framework to design approximation algorithms for finding diverse solutions in combinatorial problems. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 37 Issue 4, pages 3968–3976, 2023.

[IFF01] Satoru Iwata, Lisa Fleischer, and Satoru Fujishige. A combinatorial strongly polynomial algorithm for minimizing submodular functions. *Journal of the ACM (JACM)*, 48(4):761–777, 2001.

[Sch00] Alexander Schrijver. A combinatorial algorithm minimizing submodular functions in strongly polynomial time. *Journal of Combinatorial Theory, Series B*, 80(2):346–355, 2000.

Re-planning home healthcare services with roster stability

Arne Delaet
Research Foundation Flanders (FWO),
Hasselt University & Vrije Universiteit Brussel
e-mail: arne.delaet@uhasselt.be

Kris Braekers & Katrien Ramaekers
Hasselt University,
Research Group Logistics
e-mail: {kris.braekers, katrien.ramaekers}@uhasselt.be

Marie-Anne Guerry
Vrije Universiteit Brussel,
Department of Business Technology and Operations
e-mail: marie-anne.guerry@vub.be

Patrick Hirsch
University of Natural Resources and Life Sciences Vienna,
Institute of Production and Logistics
e-mail: patrick.hirsch@boku.ac.at

Home healthcare nurses in Flanders face a persistent challenge: unstable working rosters disrupt their ability to plan personal activities, negatively impacting job satisfaction and contributing to high turnover rates in the industry.

Existing research on nurse satisfaction in HHC predominantly employs single-objective models using weighted sums, often failing to capture the complex trade-offs between competing priorities. In addition, decision-makers struggle to assign appropriate weights to each factor in this function.

This study addresses these gaps by proposing a bi-objective model that balances operational costs and care worker satisfaction. The latter is measured through roster stability when re-planning schedules after disruptions in input data. Our approach utilises a multi-directional local search (MDLS) framework with an embedded large neighbourhood search (LNS) heuristic to approximate the Pareto frontier between these conflicting objectives. This heuristic updates an existing schedule by re-rostering nurses, re-assigning patients to nurses, re-scheduling patient visits and re-optimising routes.

The model provides insights for improving schedule stability and cost efficiency by presenting decision-makers with multiple solutions that quantify trade-offs. The effectiveness of the proposed approach are validated through empirical experiments, resulting in some managerial insights.

Optimizing the training of community first responders

B. Overbeek
Vrije Universiteit Amsterdam, Operations Analytics
e-mail: b.overbeek2@vu.nl

P.L. van den Berg
Erasmus University Rotterdam, Rotterdam School of Management

C.J. Jagtenberg
Vrije Universiteit Amsterdam, Operations Analytics

R.D. van der Mei
Centrum Wiskunde & Informatica, Stochastics Group

Keywords: Health Care, Community First Responders, Emergency Medical Services, Nonlinear Integer Programming

When medical emergencies occur, first aid must be provided in a timely manner. In general, the shorter the emergency response time the better the medical outcomes of patients. As a result, an important objective of the emergency medical services is to minimize response times.

A relatively new but potentially very effective approach to decrease response times is to augment the regular emergency response with community first responder (CFR) systems. In CFR systems, medically trained volunteers are dispatched to incidents occurring in their vicinity by alerting them via a smartphone application. Available volunteers will subsequently accept the alert and travel to the incident as soon as possible to provide first aid until the ambulance services arrive. Existing CFR systems have shown that dispatching volunteers to incidents can significantly reduce response times.

The vast majority of CFR systems dispatch volunteers only to out-of-hospital cardiac arrests. However, when volunteers are sufficiently trained, dispatching volunteers can be an effective approach to decrease the response times for a variety of medical emergencies. Within the UK, for example, numerous CFR systems exist that dispatch volunteers to various emergency types including traumatic emergencies, strokes, and allergic emergencies.

For CFR systems that dispatch volunteers to multiple emergency types, the degree of medical training typically varies considerably across volunteers. As different emergency types have different requirements regarding the medical skills a volunteer must have to be able to provide adequate medical assistance, a volunteer's degree of training directly determines to which emergency types this volunteer can be dispatched. One way in which the effectiveness of a CFR system can be increased is subsequently to train its volunteers as this allows volunteers to be dispatched to more emergency types. The more volunteers there are with the required medical skills to be dispatched to a certain incident, the higher the probability that at least one volunteer will be available to provide medical assistance.

Ideally, all volunteers are fully trained so that all volunteers can be dispatched to all emergency types. However, many CFR systems rely upon financial support from local communities. The budget available to train volunteers is consequently limited, making it financially unfeasible to fully train all volunteers. The problem that then arises is how to decide which volunteers to train and which training to provide to each volunteer to maximize the effectiveness of the CFR system given a certain budget.

Deciding on how to spend the available training budget is oftentimes highly nontrivial due to large

heterogeneity in volunteers' availability (i.e. how often they are available to go to an incident when alerted) and medical skills as well as large geographical differences in the number of volunteers, the number of incidents, and ambulance response times. As a result, in this study we develop an optimization model to decide which volunteers to train and which training to provide to each volunteer to maximize the effectiveness of the CFR system given a certain budget, taking into account the heterogeneous nature of volunteers and potential geographical differences. As volunteers are mainly valuable when they arrive before the ambulance, we moreover introduce a new measure for the effectiveness of CFR systems: the probability that at least one volunteer arrives before the ambulance.

We present a Nonlinear Integer Program (NLIP) formulation of the optimization model and develop a solution approach that can efficiently obtain optimal solutions for instances of realistic size. To demonstrate the application of the introduced optimization model and solution approach, a case study involving the CFR system of the Lincolnshire Integrated Voluntary Emergency Service is performed. Results from this case study and corresponding insights obtained will be presented.

A metaheuristic approach for integrated nurse routing and rerostering in hospital-at-home

Anisha Maharani
QuantOM, HEC Liège - Management School of the University of Liège
e-mail: anisha.maharani@uliege.be

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

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

The global demographic shift toward an aging population has driven a rising demand for medical services. This trend requires adaptation and innovation within the healthcare sector to address the growing demand despite limited resources. One such innovation is the "hospital at home" (HaH), also known as "home hospitalization", which substitutes for inpatient care by delivering short-term treatment for acute illnesses directly at patients' homes. The expected benefits of HaH include: increased capacity of the institutional health care systems, potential cost savings, improved patient quality of life, and shorter length of stay [3].

Having an acute illness, patients in HaH need intensive care for a limited period. As a result, the fluctuations of the patient mix are greater than those observed in classical home health care systems. Since the availability of resources also evolves over time, it becomes a challenging task to match the available resources and the patient needs.

This study focuses on the operational planning decisions in HaH services over a one-week horizon, taking into account the information available at the beginning of the week regarding the patient list and the availability of nurses. Given a baseline roster and the actual availability of nurses, and considering the needs of new and existing patients, several operational decisions are taken simultaneously: select new patients to be admitted, decide whether and to what extent the baseline roster should be updated, and schedule the care visits to build the daily route of each nurse. Generally, the underlying task scheduling, nurse routing, and nurse rerostering subproblems are solved independently in the literature due to computational practicality [2]. However, this sequential decision-making approach may lead to suboptimal or even infeasible solutions, as these subproblems are strongly intertwined.

A lexicographic objective function is employed 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. The routing subproblem addresses additional constraints associated with home healthcare services, including aligning the medical skills of nurses with the needs of patients and respecting time windows for treatment administrations [1]. When rerostering is necessary, compliance with institutional working hour regulations, including limits on working days and forbidden shift sequences, is essential [4]. Additionally, the continuity of rostering constraints between consecutive weeks must be maintained and care for the existing patients must be ensured.

This work proposes a metaheuristic approach to solve the integrated nurse routing and rerostering problem. The approach allows exploration of infeasible solutions during the search by relaxing time windows, rostering, and rerostering constraints. The proposed algorithm iteratively improves incumbent solutions through embedded loops. The outer loop modifies the set of admitted patients. The inner loop improves the nurse roster and routes associated with the current patient set. Destroy-and-repair mechanisms are employed to modify the patient set and solve the routing subproblem. A guided local search is integrated into the metaheuristic to restore the roster feasibility and to refine its performance.

In this talk, new benchmark instances for this integrated problem will be introduced. The effectiveness and performance of the proposed solution approach will be demonstrated by discussing some numerical results. Finally, practical insights will be provided through some managerial analyses.

1 References

[1] Castillo-Salazar, J. A., Landa-Silva, D., & Qu, R. (2016). Workforce scheduling and routing problems: literature survey and computational study. *Annals of Operations Research*, 239(1): 39-67.

[2] Ernst, A., and Jiang, H., Krishnamoorthy, M., & and Sier, D. (2004). Staff scheduling and rostering: a review of applications, methods and models. *European Journal of Operational Research*, 153(1): 3-27.

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

[4] Wickert, T. I., Smet, P., & Berghe, G. V. (2019). The nurse rerostering problem: strategies for reconstructing disrupted schedules. *Computers & Operations Research*, 104: 319–337.

Pessimistic bilevel optimization approach for decision–focused learning

Diego Jiménez
SKEMA Business School - KU Leuven, Faculty of Economics and Business
e-mail: diego.jimenezbustame@skema.edu

Bernardo Pagnoncelli
SKEMA Business School
e-mail: bernardo.pagnoncelli@skema.edu

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

Contextual optimization (CO) is a mathematical framework for formulating problems where decisions ($\mathbf{z} \in \mathcal{Z} \subset \mathbb{R}^{n_z}$) are taken from a decision model (DM) $\min \{c(\mathbf{y}, \mathbf{z}) : \mathbf{z} \in \mathcal{Z}\}$, in which some parameters ($\mathbf{y} \in \mathcal{Y} \subset \mathbb{R}^{n_y}$) are random and must be estimated from other parameters ($\mathbf{x} \in \mathcal{X} \subset \mathbb{R}^{n_x}$) [1]. Parameters $\mathbf{x}$, called features, are not present in the model but are correlated to $\mathbf{y}$. Assuming the existence of a joint probability distribution $\mathbb{P}(\mathbf{y}|\mathbf{x})$, and when the random parameters are present only on the cost function, the expected value of the final decision cost is minimized with respect to this distribution, given the observed features values:

$$\mathbf{z}^{CO}(\mathbf{x}) \in \arg\min_{\mathbf{z}\in\mathcal{Z}} \mathbb{E}_{\mathbb{P}(\mathbf{y}|\mathbf{x})}[c(\mathbf{y}, \mathbf{z})] \tag{1}$$

An *estimator*, represented by a function $\mathbf{f} : \mathcal{X} \to \mathbb{R}^{n_y}$, belonging to a family of functions $\mathbf{f} \in \mathcal{F}$, can be computed to approximate (1) as $\mathbf{z}^\star(\mathbf{f}(\mathbf{x})) \in \arg\min \{c(\mathbf{f}(\mathbf{x}), \mathbf{z}) : \mathbf{z} \in \mathcal{Z}\}$. In the *machine learning* (ML) context, these estimators are calibrated by minimizing a non-negative real-valued *loss function* $\ell : \mathbb{R}^{n_y} \times \mathbb{R}^{n_y} \to \mathbb{R}_+$, over a set of samples $\mathcal{D} = \{(\mathbf{x}_i, \mathbf{y}_i)\}_{i=1}^n$, due to the impossibility of getting the ground truth distribution. Then, the empirical loss minimization (ERM):

$$\mathbf{f}^\star \in \operatorname*{argmin}_{\mathbf{f}\in\mathcal{F}} \frac{1}{n}\sum_{i=1}^{n} \ell\big(\mathbf{f}(\mathbf{x}_i), \mathbf{y}_i\big) \tag{2}$$

searches for the optimal estimator $\mathbf{f}^\star$ using a loss function $\ell(\cdot)$ and the dataset $\mathcal{D}$. Typically in the ML context, the mean squared error (MSE) $\ell_{MSE}(\mathbf{f}(\mathbf{x}_i), \mathbf{y}_i) = ||\mathbf{f}(\mathbf{x}_i) - \mathbf{y}_i||^2$ is used as a loss function. However, this approach disregards the existence of the subsequent decision model on the estimator computation. In contrast, a new methodology referred as *integrated estimation-optimization* (IEO) [2], and *decision-focused learning* by [3], incorporates the optimization structure

of the DM into the definition of the loss function $\ell(\cdot)$, capturing decision errors. To this end, the proposed IEO loss function is the following:

$$\ell_{IEO}\big(\mathbf{f}(\mathbf{x}),\mathbf{y}\big) = \max_{\mathbf{z}\in\mathcal{Z}^\star(\mathbf{f}(\mathbf{x}))} c\big(\mathbf{y},\mathbf{z}\big) - \min_{\mathbf{z}\in\mathcal{Z}} c\big(\mathbf{y},\mathbf{z}\big) \tag{3}$$

where $\mathcal{Z}^\star(\mathbf{y})$ is the set of minimizers of $\min\{c(\mathbf{y},\mathbf{z}) : \mathbf{z}\in\mathcal{Z}\}$. Using (3), ERM (2) becomes a pessimistic bilevel problem, which is known to be hard to solve [4]. To address this, in [5] a surrogate loss function, called SPO+, is used instead, while in [6] major approximations are used. Besides, both proposals only address DMs with continuous variables.

In this work, we propose a method to solve the ERM (2), using the IEO loss (3), with $c(\mathbf{y},\mathbf{z}) = \mathbf{y}^\intercal\mathbf{z}$, for DMs containing 0–1 variables. Our method consists of solving an ε-approximation of the pessimistic bilevel problem using a specialized cut–generation algorithm. The key idea is to solve the following relaxation:

$$\min_{\substack{\mathbf{f}\in\mathcal{F},\theta_i\in\mathbb{R}\\ \bar{\mathbf{z}}_i\in\mathcal{Z}}} \left\{ \frac{1}{n}\sum_{i=1}^{n}\theta_i : \theta_i \geq \max\Big\{\mathbf{y}_i^\intercal\mathbf{z}_i : \mathbf{z}_i\in\hat{\mathcal{Z}}, \mathbf{f}(\mathbf{x}_i)^\intercal(\mathbf{z}_i-\bar{\mathbf{z}}_i)\leq\varepsilon\Big\},\ \forall i\in[n] \right\} \tag{4}$$

with $\hat{\mathcal{Z}}\subseteq\mathcal{Z}$ and $[n]=\{1,...,n\}$. The problem is solved iteratively by adding points to $\hat{\mathcal{Z}}$, i.e., adding cuts, using a branch–and–cut algorithm, in which feasible solutions can be computed. Experiments on instances of the 0–1 knapsack and shortest path problems demonstrate that our method outperforms linear regression and the SPO+ method in out-of-sample performance.

References

[1] Sadana, U., Chenreddy, A., Delage, E., Forel, A., Frejinger, E., & Vidal, T. (2023). A Survey of Contextual Optimization Methods for Decision Making under Uncertainty. *arXiv preprint*.

[2] Elmachtoub, A. N., Lam, H., Zhang, H., & Zhao, Y. (2023). - Estimate-Then-Optimize versus Integrated-Estimation-Optimization versus Sample Average Approximation: A Stochastic Dominance Perspective. *arXiv preprint*.

[3] Mandi, J., Kotary, J., Berden, S., Mulamba, M., Bucarey, V., Guns, T., & Fioretto, F. (2024). Decision-Focused Learning: Foundations, State of the Art, Benchmark and Future Opportunities. *arXiv preprint*.

[4] Wiesemann, W., Tsoukalas, A., Kleniati, P.-M., & Rustem, B. (2013). Pessimistic Bilevel Optimization. *SIAM Journal on Optimization*, 23(1), 353–380.

[5] Elmachtoub, A. N., & Grigas, P. (2022). Smart Predict, then Optimize. *Management Science*, 68(1), 9–26.

[6] Bucarey, V., Calderón, S., Muñoz, G., & Semet, F. (2024). Decision-Focused Predictions via Pessimistic Bilevel Optimization: A Computational Study. In B. Dilkina (Ed.), *Integration of Constraint Programming, Artificial Intelligence, and Operations Research* (pp. 127–135). Cham: Springer Nature Switzerland. ISBN: 978-3-031-60597-0.

Empulse: A unifying framework for profit-driven and cost-sensitive learning in Python

Shimanto Rahman
Ghent University, Faculty of Economics and Business Administration
e-mail: shimanto.rahman@ugent.be

Bram Janssens
Ghent University, Faculty of Economics and Business Administration
e-mail: bram.janssens@ugent.be

Matthias Bogaert
Ghent University, Faculty of Economics and Business Administration
e-mail: matthias.bogaert@ugent.be

Keywords: Cost-Sensitive Learning, Profit-Driven Analytics

1 Extended Abstract

To align machine learning outputs with business objectives, profit-driven analytics and cost-sensitive learning has become increasingly critical in domains such as customer churn prediction, fraud detection and credit scoring. As a result, researchers have proposed several innovative methodologies tailored to optimize business outcomes ranging from evaluation metrics (Verbraken et al., 2013) to algorithms with custom loss functions (Vanderschueren et al., 2022).

While many of these techniques have been open-sourced through GitHub repositories and packages in R or Python, their scattered nature introduces practical challenges for adoption and usability. Techniques often have distinct and incompatible APIs, making integration cumbersome, particularly for practitioners accustomed to standardized frameworks like scikit-learn (Olivier Grisel et al., 2024). Many implementations are not distributed as installable packages and have sparse or missing documentation. This forces users to manually handle dependencies with the risk of having broken environments, runtime issues or even incorrect implementations. In sum, the lack of a robust framework increases the barrier to entry, limits widespread use and undermines trust in these tools, making them less appealing to practitioners.

This fragmented landscape can be observed in a mix of repositories and packages scattered across Python and R ecosystems, each with varying levels of maintenance and accessibility. To address these challenges, we introduce **Empulse**, a Python package designed as a unifying framework for profit-driven and cost-sensitive learning. Built to seamlessly integrate with the scikit-learn ecosystem, Empulse aggregates disparate implementations into a single, cohesive platform. It

offers a comprehensive suite of tools, including Expected Maximum Profit (EMP) metrics (Verbraken et al., 2013), value-driven models like ProfLogit (Höppner et al., 2020) and B2Boost (Janssens et al., 2022), cost-sensitive methods such as CSBoost (Höppner et al., 2022) and R-CSLogit (De Vos et al., 2023), and preprocessing techniques such as over/undersampling and reweighting. Empulse prioritizes rigorous testing and thorough documentation, ensuring both reliability and ease of use.

Empulse aims to standardize APIs, enhance accessibility through a pip installable format, and promote reproducibility in scientific research. The package addresses the need for a centralized, user-friendly, and extensible framework in profit-driven and cost-sensitive learning and can be used as the go-to framework for both researchers and practitioners.

2 References

De Vos, S., Vanderschueren, T., Verdonck, T., & Verbeke, W. (2023). Robust instance-dependent cost-sensitive classification. Advances in Data Analysis and Classification, 17(4), 10571079. https://doi.org/10.1007/s11634-022-00533-3

Höppner, S., Baesens, B., Verbeke, W., & Verdonck, T. (2022). Instance-dependent cost-sensitive learning for detecting transfer fraud. European Journal of Operational Research, 297(1), 291300. https://doi.org/10.1016/j.ejor.2021.05.028

Höppner, S., Stripling, E., Baesens, B., Broucke, S. V., & Verdonck, T. (2020). Profit driven decision trees for churn prediction. European Journal of Operational Research, 284(3), 920933. https://doi.org/10.1016/j.ejor.2018.11.072

Janssens, B., Bogaert, M., Bagué, A., & Van Den Poel, D. (2022). B2Boost: Instance-dependent profit-driven modelling of B2B churn. Annals of Operations Research. https://doi.org/10.1007/s10479-022-04631-5 Olivier Grisel, Andreas Mueller, Lars, Alexandre Gramfort, Gilles Louppe, Thomas J. Fan, Peter Prettenhofer, Mathieu Blondel, Vlad Niculae, Joel Nothman, Guillaume Lemaitre, Arnaud Joly, Loïc Estève, Jérémie du Boisberranger, Jake Vanderplas, manoj kumar, Hanmin Qin, Nicolas Hug, Adrin Jalali, Chiara Marmo. (2024). scikit-learn/scikit-learn: Scikit-learn 1.5.1 (Version 1.5.1) [Computer software]. Zenodo. https://doi.org/10.5281/ZENODO.12634048

Vanderschueren, T., Verbeke, W., Baesens, B., & Verdonck, T. (2022). Instance-dependent cost-sensitive learning: Do we really need it? Hawaii International Conference on System Sciences. https://doi.org/10.24251/HICSS.2022.191

Verbraken, T., Verbeke, W., & Baesens, B. (2013). A Novel Profit Maximizing Metric for Measuring Classification Performance of Customer Churn Prediction Models. IEEE Transactions on Knowledge and Data Engineering, 25(5), 961973. https://doi.org/10.1109/TKDE.2012.50

Decision-Centric Fairness: Evaluation and Optimization for Classification Problems

Simon De Vos
KU Leuven, LIRIS
simon.devos@kuleuven.be

Jente Van Belle
KU Leuven, LIRIS
jente.vanbelle@kuleuven.be

Andres Algaba
VUB, Data Analytics Laboratory
andres.algaba@vub.be

Wouter Verbeke
KU Leuven, LIRIS
wouter.verbeke@kuleuven.be

Sam Verboven
VUB, Data Analytics Laboratory
sam.verboven@vub.be

In an ever-increasing data-driven world, algorithms and machine learning models play a pivotal role in key business decisions, whether it is targeting customers with retention incentives in marketing or allocating investigative resources in fraud detection. Central to the effectiveness and acceptability of these algorithm-informed decisions is the principle of fairness [1]. Predictive models may exhibit discriminatory behavior towards specific demographic groups; hence, resource allocations based on these models' outputs are potentially unfair.

Traditionally, algorithmic fairness has been assessed using output-based metrics, which essentially focus on average model predictions or error rates among different groups [3]. While this approach provides a broad lens, it often fails to capture intricate nuances in algorithmic behavior. Emerging work has sought to expand our understanding of fairness by considering additional statistical moments, such as the variance, and even by evaluating entire output distributions [2]. Our research seeks to extend this landscape further for classification problems. Rather than being bound by the entirety of the output distribution, we advocate for a more pragmatic approach, concentrating on what we term the "decision-making region". This region is essentially the subset of predicted probabilities above a predefined threshold. Specifically, within this region, we aim for demographic parity, i.e., a proportionally equal amount of positive outcomes, across all potential thresholds. This is a departure from the traditional concept of demographic parity, which remains agnostic to thresholds [1].

Our perspective offers several tangible advantages. Firstly, it allows the same predictive model to be versatile, catering to multiple decision thresholds without needing re-trainingthis could be particularly relevant when parameters like marketing budgets or fraud investigation capacity fluctuate. Additionally, making inference decisions based on established models proves to be both cost-effective and environmentally sustainable compared to the resource-intensive nature of model re-training. It is noteworthy that decision thresholds, often determined in

conjunction with business stakeholders post-model development, are contingent on numerous externalities such as strategic objectives, operational limitations, and available resources. Therefore, a model's ability to maintain fairness across different thresholds becomes paramount.

The prevailing discourse in algorithmic fairness has acknowledged the importance of threshold-independent fairness solutions [2]. Nonetheless, there exists an inherent trade-off; mandating fairness across the entire output domain can inadvertently curtail a model's predictive performance. Our contention is straightforward: fairness should be concentrated where decisions manifest in real-world scenarios. For instance, in a customer churn scenario, companies will naturally prioritize engaging high-risk customers while largely overlooking those deemed to be at low risk. Ensuring fairness among the low-risk, non-actionable group may not bring tangible fairness benefits but will invariably exact a performance cost for the actionable segment. Thus, we propose that fairness considerations be focused on regions that matter the most in practice. By doing so, our contributions are the following:

- We propose and formalize the concept of decision-centric fairness for classification problems;
- We propose a methodology to optimize for both decision-centric fairness and predictive performance;
- We empirically evaluate our proposed methodology and show that it outperforms existing fairness solutions which impose fairness across the entire output domain from a decision-centric evaluation perspective.

References

[1] Barocas, S., Hardt, M., and Narayanan, A. (2023). Fairness and machine learning: Limitations and opportunities. *MIT Press*.

[2] Han, X., Jiang, Z., Jin, H., Liu, Z., Zou, N., Wang, Q., and Hu, X. (2023). Retiring ΔDP: New Distribution-Level Metrics for Demographic Parity. *Transactions on Machine Learning Research*.

[3] Makhlouf, K., Zhioua, S., and Palamidessi, C. (2021). On the applicability of machine learning fairness notions. *ACM SIGKDD Explorations Newsletter*, 23(1):14–23.

Achieving Group Fairness through Independence in Predictive Process Monitoring

Jari Peeperkorn and Simon De Vos
Research Center for Information Systems Engineering (LIRIS), KU Leuven
e-mail: jari.peeperkorn@kuleuven.be

Predictive Process Monitoring (PPM) is a branch of process mining that aims to predict the future state of ongoing business processes based on historical event data. A specialized subset of PPM, known as Outcome-Oriented Predictive Process Monitoring (OOPPM), focuses on predicting specific outcomes or labels of process instances. By leveraging historical cases, OOPPM enables organizations to anticipate critical outcomes and intervene at earlier stages of the process, enabling earlier interventions that improve efficiency or reduce risks. Recent advancements in OOPPM have predominantly utilized machine and deep learning models trained on labeled historical data to achieve accurate predictions.

A significant ethical and legal challenge arises when these models are trained on biased data, which may encode systemic inequalities, such as those based on gender or ethnicity. Predictive models can inadvertently reproduce or even exacerbate these disparities if fairness considerations are not addressed. In recent years, different works have focused on using process mining to detect or discover such fairness concerns within a process. In contrast, this work focuses on ensuring OOPPM models produce fair and unbiased predictions. While improving the fairness of the process execution itself is not the primary goal of these approaches, it is assumed that efforts are made in parallel to enhance the fairness of the process. Consequently, our approaches can then be used to measure and improve the extent to which the classifier's outcomes reflect these fairer executions, rather than perpetuating historical biases. From a legal and ethical perspective, this is essential for organizations to align with fairness mandates, such as those outlined in regulations like the EU AI Act.

A key fairness criterion discussed in this project is group fairness through independence, which ensures equal predictive outcomes across protected groups (e.g., defined by gender or ethnicity). Achieving fairness often involves trade-offs, as enforcing fairness constraints may reduce predictive performance. To navigate this tension, this work evaluates the fairness-performance trade-offs, providing a principled framework for balancing these competing objectives. This approach allows stakeholders to make informed policy decisions about the desired level of fairness enforcement, aligning with a risk-based perspective that can vary depending on the application and its regulatory environment.

By addressing these challenges, this project aims to contribute to the development of fair and effective OOPPM models that align predictive capabilities

with ethical and legal fairness standards and to offer tools for practitioners and researchers to do so as well. To this extent, the main contributions of this work are:

- Introducing group fairness into predictive process monitoring, with independence as the fairness criterion.
- Proposing and evaluating metrics for *demographic parity* such as ΔDP, alongside more advanced, threshold-independent alternatives *area between probability density function curves* (ABPC) and *area between cumulative density function curves* (ABCC) [1].
- Incorporating *integral probability metrics* (IPMs), more specifically Wasserstein distance, into a composite loss function, complementing traditional loss functions such as binary cross-entropy (BCE). Experiments demonstrate that balancing IPMs with traditional loss functions enables flexible trade-offs between fairness and predictive accuracy.

Simpler methods such as threshold adjustments for different groups or batch-based groupings, are often insufficient for OOPPM. Processes often involve continuous data streams, and predictions are needed as early indicators for all running cases. Instead, we propose principled approaches to measuring and improving fairness, including the use of propensity scores and modifying the training process directly. Experimental results demonstrate that the continuous version of ΔDP and distribution-based metrics ABPC and ABCC can effectively measure demographic parity violations (independent of the chosen threshold). These metrics can be applied to any predictive model that outputs propensity scores, making them a low-effort addition to existing evaluation pipelines. To achieve models with greater fairness across protected groups, incorporating the IPM loss into composite loss functions offers a practical trade-off between independence and predictive performance. Our experiments reveal that models trained with low weights for the IPM loss achieve high AUC scores, reflecting strong predictive accuracy, but also exhibit significant demographic parity violations. Conversely, increasing the IPM loss weight reduces AUC scores but minimizes demographic parity violations, as evidenced by improvements in ABPC and ABCC metrics. By addressing these challenges, this work contributes to the development of fair and effective OOPPM models. It provides tools for practitioners and researchers to align predictive capabilities with ethical and legal fairness standards while navigating the trade-offs inherent in enforcing fairness constraints.

References

[1] Han, Xiaotian, Zhimeng Jiang, Hongye Jin, Zirui Liu, Na Zou, Qifan Wang, and Xia Hu. "Retiring ΔDP: New Distribution-Level Metrics for Demographic Parity." Transactions on Machine Learning Research, 2023.

Passenger Assignment for Public Transport Services Considering Line Frequencies and Timetabling

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

Jeroen Verstraete
KU Leuven, Institute for Mobility - CIB

Pieter Vansteenwegen
KU Leuven, Institute for Mobility - CIB

In public transport planning, properly modelling how passengers will use a proposed transport service is crucial. This aspect of the planning process, known as the Passenger Assignment Problem (PAP), involves modelling passenger behaviour on a transport network and poses significant computational challenges [1]. Since the PAP must be solved repeatedly during the planning process [2], simplified versions of the problem are often used. A common approach is to calculate all-to-all shortest paths on the network, which are then combined with passenger demand to estimate average travel times-a key metric in transportation planning.

To address this challenge, we propose a novel transit network representation that allows a decomposition of the original networks into smaller networks, decreasing the computational time. This structure enables existing algorithms to run much faster and more efficiently. Moreover, it allows for integrated assignments that consider routing, service frequencies, and time scheduling. The core idea is to decompose the network into two subgraphs, one containing only transfer nodes, where passengers can change lines, and one connecting the non-transfer nodes to the transfer nodes.

Our method works as follows: First, calculate the travel times over the individual lines without transfers. Second, find the transfer nodes. Third, do a full assignment over the transfer nodes. Fourth, find the shortest path from a non-transfer destination to a transfer node. Fifth and final, find the shortest path from the non-transfer origins to the non-transfer destinations. These last two steps work well since only the transfer nodes on the same line as the origin/destination need to be considered as intermediary nodes.

An example of the reduction in paths that need to be checked, and thus calculation time, is provided for the Floyd Warshall (FW) shortest path algorithm [1]. This algorithm has a theoretical time complexity of n^3 with n the number of nodes in the network. Consider a transit network containing only two lines each with five stops, one of which is a transfer. This means that there are a

total of nine nodes in the network, as the transfer node is visited by both lines. The traditional FW algorithm would need 9^3=729 operations to find the shortest path.

In our improved method, we know that each origin-destination pair is either served by a direct line or transfer. In the latter case, the only node where a transfer can happen is the one transfer node. This means only 9(origins) * 9(destinations) * 1(transfer node)= 81 operations are required, which is significantly less compared to the 729 operations of the original algorithm.

A second benefit of this approach is that this decomposition can be performed several times. When lines are modelled in more detail and contain a time schedule, the problem gets more complicated. All transfers now depend on the arriving and departing times of the concerned busses, plus an additional inconvenience penalty if desired. The adaption time, defined as the difference between the departing time of the bus and the desired departing time of the passenger, also becomes relevant.

In this case, our method performs a full assignment over the transfer nodes first. Afterwards, it connects the non-transfer nodes through the transfer nodes. When this step is completed, the time spent in the bus network is completely known, from everywhere to everywhere. In the final step, the shortest path considering the adaption time is considered, this is an easy operation that allows many trip requests to be processed almost instantly.

Initial results show a significant decrease in calculation time when this method is compared with the SOTA. In our largest tested instance, which is created artificially, our method can solve this instance in 51.02 seconds instead of 1336.36 seconds when solved with the state-of-the-art DLN network [1]. This instance contains 1000 lines, serving an average of 15 stops each, with a total of almost 13000 nodes, of which almost 2000 are transfers.

In the case of frequency and time scheduling, no benchmark instances are available. Therefore, we created test instances, which will become available online. On the supply side, all lines have a frequency and each bus has a starting time. For simplification, we consider symmetric and periodic instances of a fixed service. On the demand side, the demand is modelled as individual trip requests. The largest solved instance contains 300 busses, and 3457 bus stops, of which 1101 are transfers. 10000 trip requests are modelled. Our calculation time for this instance is 93.96 seconds, of which 91.18 seconds were spent to find the shortest paths in the network with the transfer nodes.

Future research will indicate how adaptable this method is to include deviations from the fixed schedule, or how the routing over the transfer points can be sped up due to the exploitable structure present in this graph.

1 References

[1] Aktaş, Dilay, Evert Vermeir, and Pieter Vansteenwegen. "Speeding up the Passenger Assignment Problem in transit planning by the Direct Link Network representation." Computers & Operations Research 167 (2024): 106647.
[2] Iliopoulou, Christina, Konstantinos Kepaptsoglou, and Eleni Vlahogianni. "Metaheuristics for the transit route network design problem: a review and comparative analysis." Public Transport 11 (2019): 487-521.

Optimizing passengers' boarding and alighting operations in urban mass transit

Laura Knappik
RWTH Aachen University, Chair for Data and Business Analytics
e-mail: knappik@analytics.rwth-aachen.de

Dr. Lorena S. Reyes Rubiano
RWTH Aachen University, Chair for Data and Business Analytics
e-mail: reyes@analytics.rwth-aachen.de

Prof. Dr. Sven Müller
RWTH Aachen University, Chair for Data and Business Analytics
e-mail: mueller@analytics.rwth-aachen.de

Urban mass transit systems, often struggle with high operational costs and tight schedules. A major source of inefficiency is the dwell time of vehicles at stations, which is significantly influenced by variability in boarding and alighting times Kuipers et al. (2021). Passengers often prefer boarding through specific doors (e.g., those near the station entrance or closer to the exit at their destination), resulting in overcrowding at some doors while others remain underutilized Oliveira et al. (2019). This imbalance prolongs dwell time, as the vehicle must wait for the last passengers to board or alight. This research aims to reduce dwell times during peak hours by optimizing the allocation of passengers to doors and trips, thereby enhancing overall system efficiency.

To achieve this, our methodology integrates three main components:

- *Optimization of on- and off-boarding processes*: This involves developing an optimization model that minimizes dwell time by allocating passengers to doors and trips.
- *Door/trip choice prediction*: We incorporate discrete choice models to predict passengers' door and trip choices. This component aims to influence passenger behavior by incentives (decision variable) to achieve a more balanced distribution across doors and trips, e.g., passengers comply with the optimization.
- *Dwell time prediction*: We predict the dwell time based on the number of boarding and alighting passengers, taking into account pedestrian dynamics and behavioral aspects.

The problem is modeled as a network flow framework representing passenger movements within the transit system. The decision variable X_{oivj} captures the number of passengers starting at station o, traveling along a feasible connection (i, v), and reaching destination j. This formulation accounts for the complexity of passenger flows, including transfers between trains at different stations and

times, enabling a detailed representation of system dynamics. We assume that the origin-destination (OD) demand at each station and for each time is known, and that line planning, timetabling, and train scheduling are predefined and not part of this problem. Furthermore, we assume that passengers board and alight through the same door, as they are unlikely to move to a different door within a crowded train. In the model we minimize the maximum dwell time by nudging passengers to doors and trips. The latter implies that we consider passenger shifting, where passengers arriving at the station of trip t can either board a line in the same trip or wait for a subsequent trip of the line (e.g., $t+1, t+2$, up to $|T|$) due to capacity constraints.

The maximum dwell time for a train at each station is modeled as a function of the number of boarding and alighting passengers. The dwell time function, derived from Puong (2000) and Lam et al. (1998), reflects pedestrian dynamics and integrates insights from previous studies on dwell time modeling. The dwell time is modeled as:

$$\tau \geq \alpha_0 + \alpha_1 \cdot P_{\text{board}} + \alpha_2 \cdot P_{\text{alight}} + \alpha_3 \cdot P_{\text{board}} \cdot P_{\text{alight}},$$

where α indicates the coefficients and P represents the number of passengers boarding or alighting, respectively. The counterflows of boarding and alighting passengers at the doors result in a nonconvex optimization problem.

Since passengers prefer specific door-trips for boarding, a discrete choice model is integrated into the optimization model. This choice model accounts for passengers' preferences for doors, depending on different attributes, such as the distance to the station entrance. We provide economic incentives, such as discounts, to nudge passengers' choices, minimizing dwell time and mitigating overcrowding. The purpose of the optimization model, therefore, is to determine optimal economic incentives for each door and trip and nudge passengers to those while minimizing dwell time. For each boarding and transfer, an own discount level is selected, thereby ensuring an upper bound on the total discount per journey. Overall, the optimization process combines finding the optimal number of passengers per door and trip and determining the price discount level for each origin for each door and trip to reach this optimal number of passengers per door.

References

Kuipers, R. A., Palmqvist, C.-W., Olsson, N. O. and Hiselius, L. W. (2021), 'The passenger's influence on dwell times at station platforms: a literature review', *Transport Reviews* **41**(6), 721–741. https://doi.org/10.1080/01441647.2021.1887960.

Lam, W. H., Cheung, C. Y. and Poon, Y. F. (1998), 'A study of train dwelling time at the hong kong mass transit railway system', *Journal of Advanced Transportation* **32**(3), 285–295. https://doi.org/10.1002/atr.5670320303.

Oliveira, L. C., Fox, C., Birrell, S. and Cain, R. (2019), 'Analysing passengers' behaviours when boarding trains to improve rail infrastructure and technology', *Robotics and Computer-Integrated Manufacturing* **57**, 282–291. https://doi.org/10.1016/j.rcim.2018.12.008.

Puong, A. (2000), 'Dwell time model and analysis for the mbta red line', MIT OpenCourseWare Project.

The divisibility index: a theoretical device to help decide public transport design

Valentina Gómez
University of Chile, Center for Mathematical Modeling (CMM)
e-mail: vagomez@uchile.cl

Sergio Jara-Díaz
University of Chile, Departament of Civil Engineer
Instituto Sistemas Complejos de Ingeniería (ISCI)
e-mail: jaradiaz@uchile.cl

Andrés Fielbaum
The University of Sydney, School of Civil Engineering
e-mail: andres.fielbaum@sydney.edu.au

Consider a public transport system serving a corridor. There is a long history of analytical models to determine the optimal strategic decisions, such as frequencies, stops spacing, or subsidies [4, 3, 2, 1]. However, little has been done to understand the network-related aspects. In particular, two usual networks found on different cities worldwide is either having a line covering the whole corridor, or two lines each serving one segment. The latter is typically observed with the two lines converging at the CBD (center business district), or in a feeder-trunk manner connected at a subcenter. In this paper, we investigate theoretically under which conditions a divided line is better than a single one, minimizing the sum of users' and operators' costs.

We formulate the problem in a generalized linear city, and optimize the frequency and vehicle size of each of the lines involved, to then compare the resulting costs of the full line versus the line divided at a given node. We prove that there are three conditions that favor the partition of a line, namely i) inducing few transfers, ii) a large difference in the maximum flows between the two segments, and iii) the segment with the lower maximum flow being long. We then synthesized these three conditions into a new Divisibility Index (DI) to measure the suitability for division at a given node. The higher the index, the more advantageous it is to split the line at that node. Subsequently, we proposed two algorithms (varying in speed and accuracy) that use the divisibility index to determine whether and where to split a line.

The approach is tested numerically in a linear version of the Parametric City Model, where we compare the optimal set of lines with the ones resulting after dividing the full line following our algorithms with very good results. The faster algorithm achieved an average error of 0.1%, while the more accurate one achieved

an average error of 0.03%, demonstrating the success of both.

References

[1] Nicolas Coulombel and Guillaume Monchambert. Diseconomies of scale and subsidies in urban public transportation. *Journal of Public Economics*, 223:104903, 2023.

[2] Andres Fielbaum. On the relationship between free public transport, stop spacing, and optimal frequencies. *Transportation Research Part B: Methodological*, 183:102924, 2024.

[3] Sergio Jara-Díaz and Antonio Gschwender. Towards a general microeconomic model for the operation of public transport. *Transport Reviews*, 23(4):453–469, 2003.

[4] Herbert Mohring. Optimization and scale economies in urban bus transportation. *The American Economic Review*, 62(4):591–604, 1972.

On the Benefits of Coordinating Ride-Hailing Platforms with Public Transit Systems

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

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

Jorge Vera
Universidad Catolica de Chile, Industrial and Systems Engineering Department
jver@uc.cl

January 2025

The integration of ride-hailing platforms and public transit systems can bring many benefits, including improved public transit utilization, affordable and accessible customer services, and reduced traffic congestion. In this work, we consider the online problem of a ride-hailing platform (e.g., Uber) that coordinates with public transit. Customers are impatient and make requests in real time. For each request, the platform determines whether to provide door-to-door, first-mile, or last-mile service and which driver to dispatch to the associated service. To address this problem, we develop a rolling-horizon approach based on integer and stochastic optimization. The approach relies on Benders decomposition, so as to make it applicable to realistic instances. In numerical experiments, we demonstrate the effectiveness of our approach in comparison with myopic or deterministic methods. Then, we provide managerial insights on the potential benefits of coordinating ride-hailing platforms with public transit systems, looking at various market and network configurations.

50 years of metaheuristics

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

Marc Sevaux
Université Bretagne Sud

Rafael Martí
Universidad de Valencia

We discuss our paper "50 years of metaheuristics" which appeared in a special issue of the European Journal of Operational Research (EJOR) commemorating the 50th anniversary of EURO. The paper provides a historical review of metaheuristic frameworks and their role in heuristic optimization. These frameworks emerged from the adaptation of heuristic reasoning, developed in the 1940s in mathematical psychology, to operations research in the late 1950s. They were driven by the need to address hard optimization problems for which traditional mathematical methods failed to produce practical solutions.

We trace the evolution of metaheuristics from simple combinatorial heuristics of the 1960s and 1970s to the robust frameworks proposed in the 1980s and 1990s—a period known as the "method-centric era." This development coincided with the "OR crisis," during which operations research faced criticism for prioritizing theoretical complexity over practical solutions. Metaheuristics addressed this gap by introducing structured frameworks to guide heuristic design, improving problem-solving efficiency. However, their natural or social behavior-inspired methodologies also led to some confusion in the field, as such analogies often contributed little to solving practical problems.

The review emphasizes the contributions of EJOR, which has consistently published high-quality heuristic research over the past four decades, albeit cautiously entering the metaheuristics field. While the field has matured significantly, enabling the solution of highly complex problems and attracting a growing research community, challenges remain. Fragmentation persists, with research groups often favoring familiar methods over adapting them to diverse problems. Additionally, the No Free Lunch (NFL) theorem underscores the need to tailor metaheuristics to problem-specific characteristics rather than relying on universal approaches.

This paper serves as both a historical overview and a call to action for researchers and practitioners to explore the diverse potential of metaheuristic frameworks. By understanding their evolution and leveraging their strengths, the community can continue developing powerful optimization algorithms. We hope this review will inspire continued progress and collaboration in the field of heuristic optimization.

What many operations researchers have done wrong and what are the remedies

Thomas Stützle
Université libre de Bruxelles and F.R.S.-FNRS
IRIDIA
e-mail: stuetzle@ulb.ac.be

What many operations researchers have done wrong in their research is that they do not differentiate between training instances and an independent test set when testing their newly developed algorithms. Instead, they often train the algorithm on the same (or a subset of) instances and then report the good results of their algorithm in papers [1, 2, 5, 6].

This is a major problem in the operations research community.

A question arises: how strong is the error's role in optimization? The answer depends on how many instances are considered in the training process. To illustrate this, we present simple results for a WalkSAT algorithm with one parameterthe walk probabilityon random 3-SAT problems with varying numbers of clauses. These results show that as we progress from 1 training instance to 10 training instances and then to 100 training instances, the performance generally becomes better and more stable.

Let us focus on the issue involving a finite or potentially infinite set of future instances. Examples include recurring problems like vehicle routing or scheduling for each working day. This is the most common scenario in operations research. We argue that there is little value in an algorithm which is trained and tested on the same instances you have provided.

A completely different situation arises when there is only one instance of interest, such as the wiring of telephone lines in a country. In such cases, there is no training set; instead, the best result for the instance is the optimal result achieved so far. Here, the goal is not to develop an algorithm that performs well across various instances but rather to achieve optimal performance on the specific instance.

Anything between these two scenarios can be discussed in terms of known and unknown instances.

How can this problem be addressed? In the first scenario, the most common one, there are several alternatives. When dealing with hand-crafted instances, random instances, or encodings of other instances, one can separate a training set from a test set. For benchmark sets, the benchmark instances can serve as the test set, while an independent training set is created. How can the independent test set be generated? One has various options regarding instance sizes, distributions

of values, etc., and must choose these appropriately. This choice also depends on whether the goal is to develop an algorithm that challenges existing ones. Alternatively, one might consider different distributions that the algorithm then can utilize.

For individual application instances, such as those from industrial companies, known techniques can be used to generate additional instances. One approach is to take a known instance and produce similar values. For example, if you have a processing time p, you can use a random distribution with processing times in the range $[0, (1-x)p), (1+x)p]$, where $x \in [0, 1]$, and apply this method to all processing times. Similar methods exist for generating both smaller and larger instances.

One question remains: how can we ensure that training instances differ from test instances? First, everyone should be aware of the error and adopt proper practices, such as generating separate training and test sets. Second, journal editors should be made aware of this issue and enforce the requirement that test results are generated independently from training results. Third, referees should recognize this potential error and adopt the same standards as editors. Finally, researchers in automatic algorithm configuration have long understood that testing instances must be different from training instances, so another reason why one should use automated configuration [3, 4].

References

[1] IBM ILOG CPLEX Optimization Studio, 2024.

[2] Gurobi Optimization, 2024.

[3] F. Hutter, H. H. Hoos, K. Leyton-Brown, Sequential model-based optimization for general algorithm configuration, in: C. A. Coello Coello (Ed.), Learning and Intelligent Optimization, 5th International Conference, LION 5, Vol. 6683 of LNCS, Springer, 2011, pp. 507–523.

[4] López-Ibáñez, M., Dubois-Lacoste, J., Pérez Cáceres, L., Stützle, T., and Birattari, M. (2016). The irace package: Iterated racing for automatic algorithm configuration. *Operations Research Perspectives*, 3:43–58.

[5] Y. Nagata and S. Kobayashi. A powerful genetic algorithm using edge assembly crossover for the traveling salesman problem. *INFORMS Journal on Computing*, 25(2):346–363, 2013.

[6] Stützle, T. and Hoos, H. H. (2000). $\mathcal{MAX}$–$\mathcal{MIN}$ Ant System. *Future Generation Computer Systems*, 16(8):889–914.

OscaR.cbls 6.0: Review after 12 years of Continuous Development

Thomas Fayolle, Quentin Meurisse, Renaud De Landtsheer, Fabian Germeau, Stefano Michelini
CETIC, Departement of Combinatorial Algorithmics
e-mail: {thomas.fayolle, quentin.meurisse, renaud.delandtsheer, fabian.germeau, stefano.michelini}@cetic.be

1 Introduction

OscaR.cbls [1] is a library, written in Scala, dedicated to the implementation of Constraint-Based Local Search [2] engines. This library contains a collection of generic constraints and neighborhoods and provides tools to implement new ones.

In 2013, Oscar.cbls 1.0 was presented in Orbel [3]. The library has just released its version 6.0 and has been subject to a complete refactoring. After more than 12 years of continuous development, this is an opportunity to review some of its main components and some of the major changes performed.

2 Variable and Constraints

OscaR.cbls lets you define three types of variables: integers, sets of integers and sequences of integers. The integer sequence is an ordered sequence of integers and is often used to represent routes in vehicle routing problems.

These variables can be used to define constraints: elements that take variables as input and update the value of other variables as output. These output variables can then be used as inputs to other constraints, which allows us to build a problem-modeling graph. When a variable is modified, it notifies all the constraints that use it as input, informing them of its modification. This information enables the constraint's output values to be updated incrementally.

Between each notification, sequences retain changes in the form of symbolic delta [4]. When they send a notification, they inform their constraints which nodes have been inserted, removed or moved since the last notification.

When the decision variables are modified during the search, these modifications are transmitted along the problem modeling graph, using a propagation mechanism. The propagation affects each node in the graph only once.

Thanks to Scala's object paradigm, a constraint abstraction can manage the integration to the propagation structure. So, implementing a new constraint can be achieved by focusing only on the incremental updates of the outputs.

3 Neighborhoods and Combinators

OscaR.cbls also contains several neighborhoods that are standard in the local search literature. On an array of integer variables, it contains, for example, the *Assign* or the *Swap*. It also contains standard neighborhoods for vehicle routing, such as *OnePointMove*, *2-opt* or *3-opt*.

In addition to the standard neighborhoods, OscaR.cbls also contains neighborhood combinators that allow you to create new neighborhoods from existing ones. For example, you can create a Cartesian product of neighborhoods. Let A and B be two neighborhoods, and let S be the set of solutions. For any solution s and any neighborhood V, let $neigh(V, s)$ be the set of solutions reachable from solution s using neighborhood V. The product of A and B ($A \times B$) explores the solutions defined by the set $\cup_{a \in neigh(A,s)} neigh(B, a)$, i.e., the set of solutions that can be reached using neighborhood B from all solutions obtained using neighborhood A from solution s.

4 Conclusion and Future Works

The OscaR.cbls library contains a collection of constraints and neighborhoods. A specific language is also used to define an optimization model and a search procedure to find a solution to this problem using local search. Its modularity makes it possible to build solvers for a wide range of problems (*Warehouse location problem*, *Vehicle Routing Problem*, etc.).

The library's modularity also means that other heuristics (*VLSN* [5], etc.) or meta-heuristics (simulated annealing, etc.) can be added. To speed up performance, it is also planned to include primitives that facilitate the use of distributed computing in the search procedure.

Acknowlegment: This work is supported by FEDER/FTJ Wallonie 2021-2027 - project no 933 - VirtualLab_Cetic.

References

[1] Oscar website. `https://github.com/cetic/oscar-cbls`

[2] Laurent Michel, and Pascal Van Hentenryck. Constraint-based Local Search. *MIT Press*, 2009

[3] Renaud De Landtsheer, and Christophe Ponsard. Oscar. cbls: an open source framework for constraint-based local search. *Proceedings of ORBEL*, 2013

[4] Renaud De Landtsheer, Yoann Guyot, Gustavo Ospina, Fabian Germeau, and Christophe Ponsard. Reasoning on sequences in constraint-based local search frameworks. *CPAIOR Proceedings 15*, 2018

[5] Sébastien Mouthuy, Pascal Van Hentenryck, and Yves Deville. Constraint-based Very Large-Scale Neighborhood search. *Constraints : an international journal*,17(2):87-122, 2012

Strategy-proofness of Credit Mechanisms for Kidney Exchange

Bart Smeulders
Eindhoven University of Technology,
Department of Mathematics and Computer Science, Combinatorial Optimization
e-mail: b.smeulders@tue.nl

For patients suffering from end-stage renal disease, a kidney transplant from a living donor is the preferred treatment option. A donor must be medically compatible with the recipient, which means some recipients can not receive a transplant, even though they have a donor willing to undergo a transplant for them. We refer to such a combination of recipient and donor as a *pair*. Kidney exchange programs (KEPs), aim to match the recipients from a pair with compatible donors from other pairs. Donors in a pair only donate if their paired recipient receives a transplant in return, leading to cycles of transplants. Since many patients in a KEP are only compatible with a small percentage of donors, the size of KEPs is crucial. Since individual hospitals and even small- and medium size countries are below the scale to be truly efficient, there is increasing interest in collaboration between KEPs.

When multiple agents collaborate, there is always the question how the benefits of collaboration will be divided, and what kind of behaviour the collaboration mechanism incentivizes. Ashlagi and Roth [1] have shown that in single-period settings, there can be no mechanism that both maximizes social welfare, and is strategy-proof. In other words, any mechanism for choosing a KEP solution in a single period either does not guarantee an optimal solution, and transplants are lost, or agents will have an incentive to lie about their pool of pairs to improve the outcome for their own patients. The latter often also comes at a cost of lost transplants. Within the literature, there has therefore been a focus on credit mechanisms that function over multiple periods. In each time period, these mechanisms compute how many credits each agent should receive, or lose, and then uses the credit balances as a tie-breaking mechanism between optimal KEP solutions. In this way, agents that receive a bad outcome in a single time period, and would therefore have an incentive to not cooperate, can be compensated through the promise of better outcomes later on.

Hajaj et al. [2] developed the first credit mechanism for kidney exchange, and explicitly consider strategy-proofness. Their mechanism is based on the number of pairs contributed and is strategy-proof, but only under the assumption that the expected pool size of each participant is known. Klimentova et al. [3] propose credit mechanisms where the number of credits function as a target allocation. Their main objective is to more fairly divide the benefits of collaboration.

Benedek et al. [4] use the same general mechanism, and study several rules for allocating credits based on well known game theoretic concepts, such as Shapley values. Neither of these papers consider strategy-proofness.

1 Results

In this talk, we will consider the credit mechanism of [3], with the Shapley value, nucleolus and normalized Banzhaff value rules introduced in [4]. We distinguish several settings, based on 1) whether or not pairs can be present for more than one time period and 2) whether pairs once revealed to the collaboration can be transplanted without consent of the collaboration. We show that.

- If pairs can only be present for a single time period, and revealed pairs can only be transplanted through the collaboration, then the Shapley value mechanism is strategy-proof.
  - Also in this setting, the nucleolus and normalized Banzhaff value mechanisms are not strategy-proof.
- If pairs can only be present for a single time period, and revealed pairs can still be transplanted without consent of the collaboration, then none of the mechanisms are strategy-proof.
- If pairs can be present for multipe time periods, none of the mechanisms considered are strategy-proof.

References

[1] Itai Ashlagi and Alvin E Roth. Free riding and participation in large scale, multi-hospital kidney exchange. *Theoretical Economics*, 9(3):817-863, 2014.

[2] Chen Hajaj, John Dickerson, Avinatan Hassidim, Tuomas Sandholm, and David Sarne. Strategy-proof and efficient kidney exchange using a credit mechanism. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 29, 2015.

[3] Xenia Klimentova, Ana Viana, João Pedro Pedroso and Nicolau Santos. *Fairness models for multi-agent kidney exchange programmes.* Omega, 102:102333, 2021.

[4] Márton Benedek, Péter Biró, Daniel Paulusma and Xin Ye. Computing balanced solutions for large international kidney exchange schemes. *Autonomous Agents and Multi-Agent Systems*, 38(1):1-41, 2024.

Home healthcare routing and scheduling with task-splitting

Loek van Montfort
Vrije Universiteit, Department of Operations Analytics
e-mail: l.g.a.j.van.montfort2@vu.nl

Markus Leitner
Vrije Universiteit, Department of Operations Analytics
e-mail: m.leitner@vu.nl

Wout Dullaert
Vrije Universiteit, Department of Operations Analytics
e-mail: w.e.h.dullaert@vu.nl

Keywords: Home health care, Task-splitting, Temporal dependencies, Vehicle Routing, Caregiver Scheduling, Synchronization, Integer programming

This study introduces the concept of task-splitting into home healthcare routing and scheduling. It focuses on the design of routes and timetables for caregivers providing services at patients' homes. Task-splitting entails the division of a (lengthy) patient visit into two separate visits that can be performed by different caregivers at different times. Splitting a visit can increase the planning flexibility of home healthcare (HHC) providers, as the resulting split tasks may have lower requirements related to the caregiver qualifications, relaxed visiting time-windows, or a shorter combined duration. However, the inclusion of task-splitting together with the temporal dependencies that arise between the split parts also presents a computational challenge.

To incorporate task-splitting decisions into the planning process, we introduce two different mixed integer linear programming formulations, a Miller-Tucker-Zemlin and a time-indexed variant. These formulations aim to minimize operational costs while simultaneously deciding which visits to split and imposing a potentially wide range of temporal dependencies. We also propose pre-processing routines for the time-indexed formulation and two heuristic procedures. These methods are embedded into the branch-and-bound approach as primal and improvement heuristics.

The results of our computational study demonstrate the additional computational difficulty introduced by task-splitting possibilities and the associated additional synchronization, and the usefulness of the proposed heuristic procedures. From a planning perspective, our results indicate that integrating task-splitting decisions into the planning process reduces staff requirements, decreases HHC operational

costs, and allows caregivers to spend relatively more time on tasks aligned with their qualifications. Moreover, we observe that the potential of task-splitting is not specific to the chosen planning objective; it can also be beneficial when minimizing travel time instead.

Integrated location and inventory decisions in healthcare logistics: review, model and first results

I. Puttemans, K. Braekers, A. Caris
Hasselt University, Research Group Logistics
e-mail: inez.puttemans@uhasselt.be, kris.braekers@uhasselt.be, an.caris@uhasselt.be

Healthcare institutions, including hospitals and retirement homes, face various challenges [5]. First, they struggle with tight budgets due to government savings. Second, they experience rising costs due to factors such as the COVID-19 pandemic, energy crisis, and inflation. Third, they deal with a shortage of care staff due to high workloads, forcing them to make an effort to relieve non-care tasks (i.e., logistics tasks). These challenges put significant pressure on healthcare institutions. To address these challenges, healthcare institutions aim to reduce costs and improve healthcare logistics while ensuring high-quality care [3,4].

One key strategy for improving healthcare logistics is consolidating inventory across healthcare institutions [3,4]. This entails healthcare institutions within a network pooling their inventory from individual warehouses into one or a few central care hubs [1]. Decision support for inventory pooling is facilitated by integrated decision-making on location and inventory management, known as the location-inventory problem (LIP) in the academic literature [2]. An LIP simultaneously addresses two main supply chain decisions: (1) location, including determining the number and locations of central care hubs, assigning healthcare institutions to these central care hubs, and (2) inventory decisions, such as determining optimal inventory parameters. By adopting this integrated decision-making, healthcare institutions can cut costs and enhance logistics, all while maintaining or improving quality of service.

This research investigates the location-inventory problem (LIP) within the healthcare context, providing decision support to improve healthcare logistics by integrating location and inventory decisions. To become familiar with the current state of LIP research, we conducted a comprehensive literature review of LIP across various sectors, like manufacturing, healthcare, spare parts, and so on. Building on the insights from this review, we developed a mathematical model tailored to a healthcare setting. The model aim to minimize total costs including location, transportation and inventory-related costs while respecting relevant constraints. Key healthcare-specific features incorporated in the model include multiple product types and a maximum allowable distance between central care hubs and healthcare institutions.

This talk will highlight the key findings from our new literature review, which covers 45 new papers across various contexts, building on the review presented in [2]. Moreover, the mathematical model is discussed, including the objective and constraints. Currently, this model is tested on small instances, and preliminary

results are discussed.

References

[1] Berman, O., Krass, D., & Mahdi Tajbakhsh, M. (2011). On the benefits of risk pooling in inventory management. *Production and Operations Management*, 20(1), 57-71.

[2] Farahani, R. Z., Rashidi Bajgan, H., Fahimnia, B., & Kaviani, M. (2015). Location-inventory problem in supply chains: a modelling review. *International journal of production research*, 53(12):3769-3788.

[3] Iannone, R., Lambiase, A., Miranda, S., Riemma, S., & Sarno, D. (2014). Pulling drugs along the supply chain: Centralization of hospitals' inventory. *International Journal of Engineering Business Management*, 6(1), 6.

[4] Parker, J., & DeLay, D. (2008). The future of the healthcare supply chain. *Healthcare Financial Management*, 62(4), 66-69.

[5] Volland, J., Fügener, A., Schoenfelder, J., & Brunner, J. O. (2017). Material logistics in hospitals: A literature review. *Omega (Oxford)*, 69, 82-101.

Using multi-period bankruptcy rules for medicine allocation

Maaike Vollebergh
Vrije Universiteit Amsterdam, Department of Operations Analytics
e-mail: m.a.j.vollebergh@vu.nl

Wout Dullaert, Arantza Estévez-Fernández & Yousef Ghiami
Vrije Universiteit Amsterdam, Department of Operations Analytics

Allocation problems arise often in practice: ranging from profit allocation in an intermodal transport system (Algaba et al., 2019), to sharing water in a transboundary river (Ansink and Weikard, 2012) and the allocation of vaccines over multiple populations (Duijzer et al., 2018). These problems can be modeled in the framework of bankruptcy problems. A bankruptcy problem arises when a group of agents each claims a portion of an estate that cannot satisfy the sum of all claims, leading to the question of how to fairly divide the available estate among the agents. A solution to bankruptcy problems is given by so-called bankruptcy rules, which specify the part of the estate that each agent should get. In the traditional bankruptcy problem, there is a single decision made at a single point in time. Looking at the above applications, one can see that these are actually problems where allocations have to be made over time.

To the best of our knowledge, bankruptcy rules over time have not been considered in the literature. Last year, we introduced the concept of bankruptcy over time in the Multi-period Bankruptcy Problem and presented different perspectives on bankruptcy over time. In this year's conference presentation, we want to highlight the key features of multi-period bankruptcy rules and report on the research project that we are conducting on allocating scarce medicines in collaboration with various actors from the Dutch pharmaceutical supply chain.

The COVID-19 pandemic exposed and compounded existing issues and vulnerabilities in the Dutch pharmaceuticals supply chain (PSC). The pandemic led to life-threatening product shortages in healthcare and thus brought new attention to the importance of drug availability (Shuman et al., 2020). Under immense time pressure, solutions were derived that saved many lives, and various initiatives were launched to contribute to a better management of the Dutch PSC. Drug shortages were already present before the COVID-19 pandemic and are currently rapidly increasing (NOS, 2024). In 2023 alone, there were 2292 drug shortages longer than two weeks in the Netherlands (KNMP, 2024). These shortages have a great impact on patient treatment and cause many issues in the pharmaceutical supply chain. A new pandemic combined with current drug shortages would greatly exacerbate vulnerabilities and issues in the PSC. Note that drug shortages are not just a Dutch problem: similar drug shortages have been reported in many other countries (COST, 2020). Drug shortages in the Netherlands are

- internationally speaking - the most frequent across Europe (Beck and Buckley, 2022), with the Netherlands accounting for 15% of the annual 22,487 shortage notifications across the EU.

In our approach to bankruptcy problems over time, we assume that the estate and the claims per period are known with certainty. Based on our current contacts with industry professionals, we observed the need to take priorities among agents (pharmacies, patients, etc.) into account. Moreover, different market conditions (e.g. regular conditions versus emergency or pandemic situations) might call for a different perspective on how unsatisfied claims from previous periods need to be taken into account when allocating scarce medicine during the current period. As these allocations may have a significant impact on the financial performance of supply chain actors, success of medical treatments, and patient well-being, the allocation rules should ideally fulfill specific properties to ensure adoption.

In our conference presentation, we will illustrate how real-life requirements for medicine allocation can be embedded in different game-theoretical bankruptcy rules. Moreover, we will discuss how to assess the effectiveness of these bankruptcy rules on real-life data.

References

Algaba, E., Fragnelli, V., Llorca, N., & Sánchez-Soriano, J. (2019). Horizontal cooperation in a multimodal public transport system: The profit allocation problem. European Journal of Operational Research, 275 (2), 659-665.

Ansink, E., & Weikard, H.-P. (2012). Sequential sharing rules for river sharing problems. Social Choice and Welfare, 38 (2), 187-210.

Beck, M., Buckley, J. (2022). Managing pharmaceutical shortages during the COVID pandemic: An exploratory analysis of European collective and national government responses. The Journal of Medicine Access 6: 27550834221123425.

Chen, C.M., & Thomas, D. J. (2018). Inventory allocation in the presence of service-level agreements. Production and Operations Management, 27 (3), 553-577.

COST (2020). CA15105 - European medicines shortages research network - addressing supply problems to patients. www.cost.eu/actions/CA15105

Duijzer, L. E., van Jaarsveld, W. L., Wallinga, J., & Dekker, R. (2018). Dose-optimal vaccine allocation over multiple populations. Production and Operations Management, 27 (1), 143-159.

KNMP (2024). Geneesmiddelentekorten 2023, https://www.knmp.nl/media/2938

NOS (2024). Tekort aan geneesmiddelen in Nederland vorig jaar groter dan ooit. https://nos.nl/artikel/2505961-tekort-aan-geneesmiddelen-in-nederlandvorig-jaar-groter-dan-ooit

Shuman, A. G., Fox, E.R., Unguru, Y. (2020). Covid-19 and drug shortages: A call to action. Journal of Managed Care & Specialty Pharmacy, 26(8), 945-947.

GARG-AML for smurfing detection in transaction networks

Bruno Deprez
KU Leuven, Information Systems Engineering Research Group (LIRIS)
University of Antwerp - imec, Department of Mathematics
e-mail: bruno.deprez@kuleuven.be

Bart Baesens
KU Leuven, Information Systems Engineering Research Group (LIRIS)
University of Southampton, Southampton Business School

Tim Verdonck
University of Antwerp - imec, Department of Mathematics
KU Leuven, Department of Mathematics

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

Money laundering is used to make illegal money, often obtained from criminal activities, appear to originate from legitimate sources [1]. Once the funds appear clean, it can be used to finance further criminal activities. Estimates have put the amount of laundered money at around 2% to 5% of global GDP, amounting to USD 2 trillion annually [4]. Since this money must go through the regular financial system, regulators have put stringent anti-money laundering (AML) rules on banks and other financial institutions.

One of the most popular methods to evade detection is **smurfing**, also called micro structuring [3]. Smurfing is used to get a large monetary amount from one account to another. Since, large transactions between two accounts are often flagged by business rules [2], the money is split into smaller amounts and transferred using money mules. This results in the typical patterns for smurfing [3], i.e., scatter-gather and gather-scatter, as visualised in Figure 1.

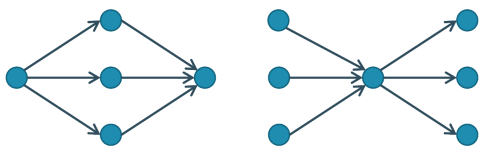


Figure 1: **Smurfing:** scatter-gather (left), and gather-scatter (right) [3]

Given the specific nature of smurfing patterns, we introduce **G**raph-**A**ided **R**isk **G**uarding for **A**nti-**M**oney **L**aundering (GARG-AML), a novel graph-based method for smurfing detection. GARG-AML generates scores based on a node's

second-order neighbourhood, capturing its resemblance to the smurfing patterns. GARG-AML takes a transaction network, groups accounts using node clustering, removes inter-cluster edges, and calculates a score based on high-density and low-density regions in the adjacency matrix.

The main contributions of this work include:

- A novel method for the featurising transaction patterns, scalable to large transaction networks, that can be easily integrated in existing AML workflows;
- Evaluation on realistic open-source synthetic data sets and a labelled proprietary data set of a European bank containing millions of transactions;
- Availability of the full implementation of the methods and experiments at `https://github.com/B-Deprez/GARG-AML`, promoting further adoption and experimentation.

References

[1] Michael Levi and Peter Reuter, Money laundering, *Crime and Justice*, Vol. 34, No. 1, pp. 289–375, 2006. DOI: `https://doi.org/10.1086/501508`.

[2] Richard J. Bolton and David J. Hand, Statistical Fraud Detection: A Review, *Statistical Science*, Vol. 17, No. 3, pp. 235–255, 2002. DOI: `https://doi.org/10.1214/ss/1042727940`.

[3] Michele Starnini, Charalampos E. Tsourakakis, Maryam Zamanipour, André Panisson, Walter Allasia, Marco Fornasiero, Laura Li Puma, Valeria Ricci, Silvia Ronchiadin, Angela Ugrinoska, Marco Varetto, and Dario Moncalvo, Smurf-Based Anti-money Laundering in Time-Evolving Transaction Networks, in *Machine Learning and Knowledge Discovery in Databases. Applied Data Science Track*, Springer International Publishing, Cham, 2021, pp. 171–186. ISBN: 978-3-030-86514-6. URL: `https://doi.org/10.1007/978-3-030-86514-6_11`.

[4] United Nations Office on Drugs and Crime, Money Laundering. URL: `https://www.unodc.org/unodc/en/money-laundering/overview.html`. Accessed: April 7, 2023

Unseen market shares: exploiting CDR data for telecom insights in a prepaid multi-SIM setting

Yaël De Rocker
Ghent University, Department of Marketing, Innovation and Organisation
FlandersMake@UGent - core lab CVAMO, Belgium
Yaël.DeRocker@UGent.be

Matthias Bogaert
Ghent University, Department of Marketing, Innovation and Organisation
FlandersMake@UGent - core lab CVAMO, Belgium
Matthias.Bogaert@UGent.be

In countries where the telecommunications sector is characterized by prepaid mobile subscriptions and widespread multi-SIM usage, estimating market shares is challenging. These markets allow users to frequently switch between providers, driven by short-term promotions or variations in network coverage. This is in contrast to postpaid subscription markets, which have been the focus of existing research [2] because of their predictable consumer behavior and long-term usage patterns compared to the fluctuating and fragmented nature of prepaid and multi-SIM markets [1].

To address the complexities of market share estimation in these environments, our study leverages Call Detail Records (CDR) consisting of logs that capture the specifics of every call in which a particular operator's subscribers are involved. These records offer a valuable yet underutilized source of information in academic studies on market share estimation. We introduce a novel methodology that predicts market shares by estimating the *invisible fraction*the subset of calls exchanged exclusively between subscribers of competing operators, which are structurally absent from an operator's CDR. By analyzing the network characteristics of on-net calls (made among subscribers of the operator) and off-net calls (made between the operators subscribers and subscribers of competing operators), we simulate the behavior of the invisible fraction. This estimated fraction is then used to correct the market share overestimations commonly observed in more naive methods, which assume that all invisible subscribers belong to the operator, thereby inflating market share calculations.

We validate our methodology through a case study involving a major telecommunications provider in Sub-Saharan Africa. At the national level, our estimation method outperforms naive estimates, yielding a close approximation to the provider's traditionally derived market share estimate. Moreover, we demonstrate that segmenting CDR data at the department level and integrating demographic data and Mobile Penetration Rates (MPR) improve the precision of these estimates.

By pioneering the use of CDR data for market share estimation in a prepaid and multi-SIM setting, as well as the combined use of CDR and demographic data, this study fills a crucial gap in existing literature predominantly focused on other application areas within postpaid subscription models. Beyond its academic contributions, our research offers practical value for telecommunications companies by providing a cost-effective and scalable alternative to traditional survey-based market research. Leveraging routinely collected CDR data, companies can gain dynamic, granular insights into market behavior while avoiding the substantial costs associated with conventional methods.

References

[1] Binh, T.V., Thy, N.G., and Phuong, H.T.N. (2021). Measure of clv toward market segmentation approach in the telecommunication sector (vietnam). *SAGE Open*, 11(2):21582440211021584.

[2] Óskarsdóttir, M., Bravo, C., Verbeke, W., Sarraute, C., Baesens, B., and Vanthienen, J. (2017). Social network analytics for churn prediction in telco: Model building, evaluation and network architecture. *Expert Systems with Applications*, 85:204-220.

Estimating treatment effects in networks using domain adversarial training

Daan Caljon
KU Leuven, Faculty of Economics & Business
e-mail: daan.caljon@kuleuven.be

Jente Van Belle
KU Leuven, Faculty of Economics & Business

Wouter Verbeke
KU Leuven, Faculty of Economics & Business

Individualized treatment effect estimation enables data-driven optimization of decision-making. Traditionally, different instances are assumed to be independent, meaning they do not influence one another. However, in many real-world scenarios, spillover, or network, effects are present. For example, a vaccine not only directly benefits its recipient by reducing their risk of severe illness but also benefits people in their social circle (e.g., friends and family) thanks to the individual's improved protection.

Recent advancements in causal machine learning have introduced methods for estimating treatment effects in network settings [4, 3]. These methods often rely on the assumption of a predefined exposure mapping, which specifies how the treatments of connected instances influence the outcome of a given instance. A common approach is to summarize the treatments of direct contacts using the sum or proportion of treated one-hop neighbors. However, this approach is often unrealistic in real-world scenarios where the exact mechanisms of spillover effects are unknown. Moreover, spillover effects may be heterogeneous, i.e., dependent on the features of the instances involved. Accurately estimating these network effects is crucial for optimizing treatment allocation to maximize the desired outcome [2].

Ideally, treatment effects are estimated using data from Randomized Controlled Trials (RCTs), considered the gold standard for causal inference. However, RCTs are often expensive or unethical to conduct. Consequently, it might be necessary to rely on observational data to estimate treatment effects. A major challenge in using observational data is confounding, which creates a covariate shift between the treatment and control groups, potentially biasing treatment effect estimates. To address this issue, previous work has proposed to learn treatment invariant, or balanced, representations of the data [5, 1]. Building on these techniques, we introduce a novel causal machine learning method for estimating heterogeneous treatment effects in network settings. Our approach integrates the representational power of Graph Neural Networks (GNNs) with domain adversarial learning

to construct treatment-invariant representations.

In this work, we empirically examine the importance of using powerful GNN layers and learning balanced representations in improving accuracy of treatment effect estimation. To this end, we perform an extensive analysis on (semi-)synthetic data with varying levels of confounding, homophily, network effect strength, and network effect complexity (e.g., heterogeneous vs. homogeneous effects). The results of this analysis indicate that homophily and confounding interact to create clusters of treated and untreated nodes within the network. These clusters introduce a covariate shift at the network level, as homophily results in connections between nodes with similar features, and confounding increases the likelihood that these connected nodes receive a similar treatment. This poses a significant challenge for graph learning methods that do not address confounding, leading to inaccurate treatment effect estimates. However, we empirically show that our proposed causal method, which employs domain adversarial learning, effectively addresses this covariate shift and outperforms other methods for heterogeneous treatment effect estimation in networks.

References

[1] Ioana Bica, Ahmed M Alaa, James Jordon, and Mihaela van der Schaar. Estimating counterfactual treatment outcomes over time through adversarially balanced representations. In *International Conference on Learning Representations*, 2020.

[2] Daan Caljon, Jente Van Belle, Jeroen Berrevoets, and Wouter Verbeke. Optimizing treatment allocation in the presence of interference. *arXiv preprint arXiv:2410.00075*, 2024.

[3] Song Jiang and Yizhou Sun. Estimating causal effects on networked observational data via representation learning. In *Proceedings of the 31st ACM International Conference on Information & Knowledge Management*, pages 852–861, 2022.

[4] Yunpu Ma and Volker Tresp. Causal inference under networked interference and intervention policy enhancement. In *International Conference on Artificial Intelligence and Statistics*, pages 3700–3708. PMLR, 2021.

[5] Uri Shalit, Fredrik D Johansson, and David Sontag. Estimating individual treatment effect: generalization bounds and algorithms. In *International conference on machine learning*, pages 3076–3085. PMLR, 2017.

ProCause: Realistic Benchmarking in Prescriptive Process Monitoring

Jakob De Moor[1], Alexander Stevens[1], Johannes De Smedt[1], Hans Weytjens[1,2], and Jochen De Weerdt[1]

[1]KU Leuven, FEB, LIRIS, Belgium , {jakob.demoor, alexander.stevens, johannes.desmedt, hans.weytjens, jochen.deweerdt}@kuleuven.be
[2]Technical University of Munich, CIT, Germany, hans.weytjens@tum.de

Prescriptive Process Monitoring (PresPM) is an emerging area within Process Mining, focused on prescribing real-time interventions for effective decision-making during business process executions [2]. PresPM holds significant promise for organizations seeking to improve operational performance. Current methodologies mostly use reinforcement learning and (causal) machine learning techniques [1, 5].

However, accurate method benchmarking remains a persistent challenge in PresPM. Recorded datasets lack counterfactual outcomes, i.e. the results of actions not taken or recorded. This absence makes it difficult to directly evaluate the method performance, since not every prescribed action for a test case has been observed. Fully synthetic data generation addresses this issue by enabling the calculation of outcomes for all possible actions, making it a popular approach in causal machine learning. However, sequential discrete-time business process data used in PresPM are often too challenging to synthetically generate due to exploding state spaces and intricate process-wide correlations. Moreover, methods validated on fully synthetic data may not generalize to real-world scenarios.

A potential workaround is the use of semi-synthetic set-ups, which pair real covariates with simulated intervention selection and outcome mechanisms. This enables the calculation of counterfactual outcomes while partially grounding the data in reality. However, such set-ups remain underexplored in PresPM and carry the notable limitation of potentially unrealistic representations of interventions and outcomes.

Another alternative involves fitting generative models to real data, which aim to approximate the true data distribution without explicitly imposing strong parametric assumptions. Once trained, these models can approximate all outcomes, treating them as ground truth. This method can be advantageous as it relies on actual data covariates and a realistic intervention selection and outcome mechanism. It is used more commonly in PresPM research, as demonstrated in [1, 5]. All PresPM works using this set-up employ RealCause, a causal machine learning benchmarking framework that uses neural networks to parameterize the data distributions [3]. However, case prefixes have to be encoded into one aggregated representation for use with RealCause, originally designed to handle

cross-sectional data. This encoding step may lead to a loss of temporal and order information. Furthermore, since counterfactual outcomes treated as ground truth are still estimates, performance evaluations may be biased for certain datasets or methods.

Subsequently, we present *ProCause*, a benchmarking framework tailored for model selection and evaluation of PresPM methods. Recognizing that the benchmarking approaches mentioned earlier have their own strengths and limitations, ProCause is designed to deliver realistic evaluations while offering users flexibility in customizing their evaluation set-up. Realistic benchmarking is achieved by using real data and fitting models that are able to capture the sequential nature of process data (e.g., an adapted conditional variational autoencoder) and generate reliable ground-truth outcomes in business processes. To ensure valid process sequences, we incorporate sequence-based rules mined from the original event log using the Declare language [4]. ProCause guarantees flexibility by allowing users to choose their preferred evaluation approach (except for the fully synthetic approach, as it can quickly become impractical if not carefully managed). Users have the option to fully define intervention selection and outcomes, creating a semi-synthetic set-up. Alternatively, they can adjust only critical elements, such as confounding factors or heterogeneity, or rely entirely on ProCause to generate all components, following a methodology similar to RealCause. We evaluate ProCause using synthetic data to validate ground-truth estimates and real-life data to assess how well the generated data represents real-world processes. Additionally, we showcase a range of experiments that demonstrate the versatility and practical use of the approach.

References

[1] Bozorgi, Z.D., Dumas, M., Rosa, M.L., Polyvyanyy, A., Shoush, M., Teinemaa, I.: Learning when to treat business processes: Prescriptive process monitoring with causal inference and reinforcement learning (2023), https://arxiv.org/abs/2303.03572

[2] Kubrak, K., Milani, F., Nolte, A., Dumas, M.: Prescriptive process monitoring: Quo vadis? CoRR **abs/2112.01769** (2021), https://arxiv.org/abs/2112.01769

[3] Neal, B., Huang, C., Raghupathi, S.: Realcause: Realistic causal inference benchmarking. CoRR **abs/2011.15007** (2020), https://arxiv.org/abs/2011.15007

[4] Pesic, M., Schonenberg, H., van der Aalst, W.M.: Declare: Full support for loosely-structured processes. In: 11th IEEE International Enterprise Distributed Object Computing Conference (EDOC 2007). pp. 287–287 (2007). https://doi.org/10.1109/EDOC.2007.14

[5] Shoush, M., Dumas, M.: White box specification of intervention policies for prescriptive process monitoring. Data and Knowledge Engineering **155**, 102379 (2025). https://doi.org/https://doi.org/10.1016/j.datak.2024.102379, https://www.sciencedirect.com/science/article/pii/S0169023X24001034

Designing a decision support tool for strategic waste collection

Jérôme De Boeck, Morgane Dumont
HEC Liège, University of Liège, Belgium, e-mail: jerome.deboeck@uliege.be
Meritxell Pacheco
DS & OR, University of Fribourg, Switzerland
Vera Fischer
Swiss postal service, Switzerland

Waste collection management has seen an increasing interest in the OR community these past years. This is due to the wide range of Capacitated Vehicle Routing Problems (CVRP) studied and the growing trend in studying sustainability-related problems.

For waste collection companies, it can be challenging to identify the type of strategy most suitable for a given situation. On the one hand, the complexity of the state-of-the-art algorithms presented in the literature; on the other hand, the data needed for these algorithms can be challenging to obtain and encode.

In partnership with Alpenluft, a Swiss waste collection consulting company, and the Innosuisse agency supporting R&D projects, we developed the WasteLogs application, a user-friendly strategic waste collection decision tool. The application offers interfaces allowing the encoding of the collection points, the amounts of waste to collect, and the collection strategy in different features that can be combined to generate a routing for collection vehicles. There are currently three state-of-the-art collection strategies implemented in the tool. Each algorithm minimizes the CO2 emissions through heuristic methods; the user can then identify what collection strategy is the most suitable and extract the information needed to import them into GPS systems. WasteLogs also allows importing existing collection tours to evaluate whether they can be improved.

The major contribution of this tool is the possibility of testing a wide range of scenarios through the granularity of the data the user can input. We present a case study in Swiss municipalities. Each municipality has very different characteristics such as the topology of the road network, the density of the population, type and amount of waste to collect, some areas with restriction on the type of vehicule that can be used for collection, the location of the waste depots,... We show in numerical experiments that based on these information, the best strategy for waste collection varies strongly from one municipality to another, emphazing the need in having decisions tools allowing to test a wide range of setting for the end user.

A mixed integer program for the smart waste collection problem with partial information

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

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

The primary goal of (municipal) waste collection is to provide a service to public without incurring substantial collection costs. To ensure an economically viable way, waste collection has traditionally been managed by a fleet of heavy vehicles following fixed routes on predetermined frequencies.

Collection routes and frequencies are typically determined based on past data on the expected waste generation of different neighbourhoods for curbside collection or the expected fill rates of different bins for the (underground) trash bins available to the public to dispose their trash. In both cases, due to the unpredictable nature of waste generation-where the amount of municipal solid waste can vary significantly depending on numerous factors-waste collection companies often face high transportation costs and inefficient resource use. In this work, we consider the waste collection via common (underground) trash bins. In this case, if the collection frequencies are decided based on bins with high fill rates, trucks often end up collecting from bins that are only partially or sometimes barely filled at all. Conversely, when collection frequencies are set based on bins with low fill rates, then overflows at the high fill rate bins are unavoidable. Moreover, the stochastic nature of the waste generation, i.e., not knowing exactly how much waste will be collected at which bin causes to include buffers in operations and results in under-utilized vehicle capacities.

With the advent of smart city technologies, the uncertainty surrounding waste generation can be partially addressed. Smart bins equipped with sensors can transmit real-time data on the volume or the weight of waste at any given moment. It is important to note that continuous data transmission requires certain computational effort of processing to extract reliable information. Additionally, waste companies may prefer not to alter routes once vehicles are dispatched. Yet, it could be possible to adjust the collection plan and routes on a weekly or daily basis rather than a fixed route and frequency collection. For this reason, a sensor reading taken at the beginning of each day or just before the collection can be used to make more informed decisions.

Most existing studies on smart waste collection assume that all bins are equipped with sensors, with daily or real-time updates on the volume or weight of trash [1]. However, equipping bins with sensors requires an initial investment

as well as an ongoing maintenance. Moreover, the gain obtained in operational efficiency by equipping certain bins might not be break-even compared to the investment [2]. Thus, the smart waste collection can also address the case where only a portion of the bins are equipped with sensors. This means that while some bins are equiped with sensors that transmit data daily, others rely on predictions based on historical data and the expertise of the waste collection company. In this work, we introduce the Smart Waste Collection Problem with Partial Information (SWCPI) where the aim is to optimize the collection strategy for a given area for a given period, when only a portion of the bins are equipped with sensors.

In *smart* waste collection, the decisions are supported with available sensor data on waste levels in the pursuit of improving efficiency. Yet, there are different ways to make use of this information. Within smart waste collection, *static* systems operate with fixed collection routes and schedules for a specific *planning horizon*, that is not subject to change anymore once decided. An example to this can be optimizing collection schemes for the week ahead using the data available at the time of planning. This means that the collection plan for the week is not subject to change once the week starts. Although the sensor data is used to make decisions in the beginning, it does not allow for mid-period adjustments, even if it is available. In contrast, in *dynamic* systems, updates-such as daily schedule adjustments-based on current waste data are possible. This adaptability allows for more responsive and efficient waste collection given the stochastic nature of waste generation. Within the dynamic collection, dynamic-offline systems refer to the systems where the collection decisions can only be updated if the operational period has not yet started. The dynamic-offline version of the example given previously would then correspond to updating the collection plan daily based on the newly available information while still considering a longer planning horizon, e.g., the next week. In the dynamic-online systems, the collection routes can be updated even while vehicles are performing their routes by inserting a new bin on the route or skipping some of the previously planned bins.

In this work, we consider static and dynamic-offline smart waste collection systems considering a planning horizon of a week. We develop a Mixed Integer Program to solve the SWCPI statically. For the dynamic-offline variant, we solve the MIP model in a rolling horizon principle, i.e., re-optimizing the upcoming weekly collection scheme each day. We present our preliminary results comparing the performance of the two systems and a fixed collection strategy.

References

[1] Hess, Christina, Dragomir, Alina G, Doerner, Karl F, and Vigo, Daniele. *Waste collection routing: a survey on problems and methods*. Central European Journal of Operations Research, vol. 32, no. 2, pp. 399–434, 2024, Springer.

[2] Lopes, Manuel and Ramos, Tânia RP. *Efficient sensor placement and online scheduling of bin collection*. Computers & Operations Research, vol. 151, pp. 106113, 2023, Elsevier.

A Mathematical Model for Optimizing Construction Logistics in Urban Areas

Yunyun Wang
Hasselt University, Research Group Logistics
e-mail: yunyun.wang@uhasselt.be

An Caris
Hasselt University, Research Group Logistics
an.caris@uhasselt.be

Kris Braekers
Hasselt University, Research Group Logistics
kris.braekers@uhasselt.be

This study investigates the integration of a Construction Consolidation Center (CCC) into urban construction logistics. A CCC functions as a consolidation and distribution hub, enabling the bundling of materials and facilitating Just-In-Time (JIT) deliveries [1]. The overall objective of our research is to enhance the efficiency of urban construction logistics while minimizing costs, environmental impacts, and social burdens. Various scenarios and operational parameters will be analyzed to quantify the potential benefits of a CCC for both a city and private parties, including construction companies and material suppliers.

Our problem setting involves multiple construction sites located within an urban area, requiring materials supplied from external locations. The CCC, usually positioned at the outskirts of the city, provides storage facilities and a dedicated vehicle fleet to ensure timely delivery to construction sites. Construction sites, constrained by limited on-site storage, request materials as dictated by their project schedules. Suppliers decide the amount of materials delivered directly to construction sites or to the CCC, which can store and consolidate materials before dispatching them to their final destinations.

In the first step, we review the related literature and we propose a mathematical model to optimize the daily delivery plans for both suppliers and the CCC. Decisions include vehicle assignment, scheduling, and route planning for two types of deliveries: (1) deliveries from suppliers directly to construction sites or to the CCC and (2) deliveries from the CCC to the construction sites. The objective for the model is to minimize total costs, comprising transportation costs for suppliers and the CCC, as well as inventory costs incurred at the CCC and construction sites.

To address dynamic operational challenges, a rolling horizon approach is adopted, inspired by [2]. Daily plans are created for an extended horizon, with the first week's plan being fixed while subsequent plans are updated every week to accommodate disturbances and unforeseen changes.

In later steps, experiments will be conducted under various scenarios to evaluate the CCC's impact on urban construction logistics. Comparative analyses include scenarios with and without a CCC, examining factors such as urban policies (e.g., vehicle restrictions, delivery time windows, and storage fee) and their influence on stakeholder decisions. This study provides insights into the operational and policy implications of CCC adoption, offering a pathway for sustainable and efficient urban construction logistics.

References

[1] C. Guerlain, S. Renault, and F. Ferrero, "Understanding Construction Logistics in Urban Areas and Lowering Its Environmental Impact: A Focus on Construction Consolidation Centres," *Sustainability*, vol. 11, no. 21, p. 6118, Nov. 2019. Available: `https://www.mdpi.com/2071-1050/11/21/6118`. doi: 10.3390/su11216118.

[2] P. C. Nolz, "Optimizing Construction Schedules and Material Deliveries in City Logistics: A Case Study from the Building Industry," *Flexible Services and Manufacturing Journal*, vol. 33, no. 3, pp. 846878, Sep. 2021. Available: `https://link.springer.com/10.1007/s10696-020-09391-7`. doi: 10.1007/s10696-020-09391-7.

Envy-Free Pricing in Concentric Seating Arrangements

Freija van Lent
Department of Data Analytics and Digitalisation, Maastricht University
e-mail: f.vanlent@maastrichtuniversity.nl

Julian Golak
Institute of Operations Management, University of Hamburg
e-mail: julian.golak@uni-hamburg.de

Alexander Grigoriev
Department of Data Analytics and Digitalisation, Maastricht University
e-mail: a.grigoriev@maastrichtuniversity.nl

In this paper, we explore envy-free unit-demand pricing problems that arise in scenarios such as concentric venues (e.g., planetariums), where seats are arranged around a central focal point. We consider a set of customers, each with a preference for one specific seat. Their valuation for other seats diminishes linearly with the Euclidean distance from the central seat. The operator aims to maximize revenue by setting a non-discriminatory price for each seat. The operator also ensures an envy-free outcome: every assigned customer receives their most preferred seat at nonnegative utility and no unassigned customer desires any available seat at the given prices. We refer to this problem as the *concentric envy-free pricing problem*.

Envy-free pricing, an equilibrium conditions in combinatorial markets, has been studied as a fairness and decentralizing mechanism [1, 2, 3]. While prior work often focuses on social welfare or considers relaxed fairness conditions, our focus is on revenue maximization under a geometric valuation structure. Our work is also related to that of Myklebust, Sharpe and Tunçel [4] on segment-based pricing as our methods extend to scenarios where customers form groups with similar preferences.

Key Insights and Contributions. We assume that each customers valuation decreases linearly with distance, a geometric assumption that allows us to identify structural properties of optimal prices. In particular, we prove that there exists an optimal envy-free solution with a piecewise-linear, nonincreasing, convex pricing function that has at most $n + 1$ breakpoints, where n is the number of customers. These breakpoints correspond to intersections of customers linear valuation curves in the graph that plots the maximum willingness-to-pay of all customers to the distance from the central seat.

In the *continuous concentric envy-free pricing problem*, we consider a continuum collection of seats that spans the entire Euclidean plane, implying that customers can be placed without any constraints on proximity or capacity. We present a polynomial-time dynamic programming algorithm that constructs an optimal pricing function. This solution approach leverages the geometric structure and ordering properties of the valuations.

In contrast, the *discrete concentric envy-free pricing problem* features a countable set of seats within the Euclidean plane. This approach aligns with practical seating configurations. The problem is inherently more challenging due to capacity constraints in each distance range. We develop a pseudo-polynomial time algorithm based on dynamic programming to handle this case.

Beyond planetariums or other concentric venues, our methods can be applied to allocation problems in both space and time, such as inventory management or delivery scheduling. Therefore, the approaches presented in this work provide a broadly applicable framework for achieving envy-free revenue maximization.

References

[1] Michele Flammini, Manuel Mauro, and Matteo Tonelli. On social envy-freeness in multi-unit markets. *Artificial Intelligence*, 269:1–26, 2019.

[2] Venkatesan Guruswami, Jason D. Hartline, Anna R. Karlin, David Kempe, Claire Kenyon, and Frank McSherry. On profit-maximizing envy-free pricing. In *ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1164–1173, 2005.

[3] Ning Chen and Atri Rudra. Walrasian equilibrium: Hardness, approximations and tractable instances. *Algorithmica*, 52:44–64, 2008.

[4] Tor G. J. Myklebust, Malcolm Sharpe, and Levent Tunçel. Efficient heuristic algorithms for maximum utility product pricing problems. *Computers & Operations Research*, 69:25–39, 2016.

Quantifying Core Stability Relaxations in Hedonic Games

Tom Demeulemeester
University of Lausanne, Faculty of Business and Economics (HEC), Department of Economics
e-mail: tom.demeulemeester@unil.ch

Jannik Peters
National University of Singapore, School of Computing
e-mail: peters@nus.edu.sg

Abstract

We study relationships between different relaxed notions of core stability in hedonic games, which are a class of coalition formation games. Our unified approach applies to a newly introduced family of hedonic games, called α-hedonic games, which contains previously studied variants such as fractional and additively separable hedonic games. In particular, we derive an upper bound on the maximum factor with which a blocking coalition of a certain size can improve upon an outcome in which no deviating coalition of size at most q exists. Counterintuitively, we show that larger blocking coalitions might sometimes have lower improvement factors. We discuss the tightness conditions of our bound, as well as its implications on the price of anarchy of core relaxations.

1 Introduction

Hedonic games constitute one of the most popular subcases of coalition formation. In a hedonic game, the goal is to divide a set of agents having preferences over subsets of the other agents into disjoint coalitions respecting these preferences. A natural stability concept is that of *core stability*: a solution is core stable if no subset of agents could together form a new coalition in which they are all better off than in the original solution. For various variants of hedonic games, core stable outcomes may fail to exist, even for very simple preference structures.

In response, we study relaxed notions of core stability, introduced by Fanelli et al. [1]. A solution is *q-size core stable* if no subset of at most q agents exists that could form a new coalition in which all agents are better off than in the original solution.

2 Results

Given a q-size core stable solution, we provide an upper bound on the factor with which a coalition of size $m \geq q+1$ can improve their utility. Our result holds for

a highly general family of hedonic games, called α-hedonic games, which contains previously studied variants such as (modified) fractional and additively separable hedonic games.

Counterintuitively, our upper bound is not monotonous for some variants of hedonic games. This implies that coalitions of size m might sometimes only experience a smaller improvement factor than coalitions of size $m' < m$.

We show the tightness of our bound for a large subclass of α-hedonic games, both theoretically and by constructing tight instances that were obtained using integer linear programming. Our bounds have direct implications on the price of anarchy of q-size core stable solutions in several well-studied subclasses of hedonic games. Additionally, our unified approach facilitates the study of core stability in previously unstudied variants of α-hedonic games.

References

[1] Fanelli, A., Monaco, G., and Moscardelli, L. (2021). *Relaxed core stability in fractional hedonic games*. Proceedings of the 30th International Joint Conference on Artificial Intelligence (IJCAI), 182–188.

Are There Arguments In Favor Of Using PROMETHEE To Rank Alternatives Evaluated On Multiple Conflicting Criteria?

Yves De Smet
Université libre de Bruxelles, Ecole polytechnique de Bruxelles,
Computer and Decision Engineering department, CoDE-SMG research unit
e-mail: yves.de.smet@ulb.be

Most strategic decision problems involve the evaluation of alternatives according to multiple conflicting criteria. Over the last 60 years, researchers have developed several approaches to address this question. These are usually classified into three main families [1]: interactive, aggregating and outranking methods. PROMETHEE methods belong the latter category [2].

First introduced by Prof. Jean-Pierre Brans in 1982, they have been widely used in practice. This is probably due to their simplicity and the existence of user-friendly software such as PROMCACL [3], Decision Lab 2000, D-Sight [4] or Visual PROMETHEE [5]. Today, hundreds of applications based on PROMETHEE have been published [6] and extensions to group decision making, robustness and sensitivity analysis, uncertainty management, descriptive problem setting, multicriteria classification and clustering, etc. have been proposed. In this talk, we will summarize arguments in favor or using PROMETHEE methods. This will be articulated around three axes:

- Robustness and sensitivity analysis of preference parameters [7, 8, 9];
- The managament of possible rank reversal occurences [10, 11, 12, 13, 14];
- The computation of net flow scores [15, 16].

This will lead us to highlight remaining research challenges.

References

[1] Philippe Vincke. *Multicriteria decision-aid.* J. Wiley, Chichester, 1992.

[2] Jean-Pierre Brans and Yves De Smet. *PROMETHEE Methods*, pages 187–219. Springer New York, New York, NY, 2016.

[3] Jutta Geldermann and Kejing Zhang. Software review: "decision lab 2000". *Journal of Multi-Criteria Decision Analysis*, 10:317 – 323, 11 2001.

[4] Quantin Hayez, Yves De Smet, and Jimmy Bonney. D-sight: A new decision making software to address multi-criteria problems. *International Journal of Decision Support System Technology*, 4:1–23, 10 2012.

[5] B. Mareschal and Y. De Smet. Visual promethee: Developments of the promethee & gaia multicriteria decision aid methods. In *2009 IEEE International Conference on Industrial Engineering and Engineering Management*, pages 1646–1649, 2009.

[6] Majid Behzadian, R.B. Kazemzadeh, A. Albadvi, and M. Aghdasi. Promethee: A comprehensive literature review on methodologies and applications. *European Journal of Operational Research*, 200(1):198–215, 2010.

[7] Bertrand Mareschal. Weight stability intervals in multicriteria decision aid. *European Journal of Operational Research*, 33(1):54–64, 1988.

[8] Nguyen Anh Vu Doan and Yves De Smet. An alternative weight sensitivity analysis for promethee ii rankings. *Omega*, 80:166–174, 2018.

[9] Xianliang Liu and Yunfei Liu. Sensitivity analysis of the parameters for preference functions and rank reversal analysis in the promethee ii method. *Omega*, 128:103116, 2024.

[10] Wim De Keyser and Peter Peeters. A note on the use of promethee multicriteria methods. *European Journal of Operational Research*, 89(3):457–461, 1996.

[11] Bertrand Mareschal, Yves De Smet, and Philippe Nemery. Rank reversal in the promethee ii method: Some new results. In *2008 IEEE International Conference on Industrial Engineering and Engineering Management*, pages 959–963, 2008.

[12] Julien Roland, Yves De Smet, and Céline Verly. Rank reversal as a source of uncertainty and manipulation in the promethee ii ranking: A first investigation. In *Advances in Computational Intelligence: 14th International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems, IPMU 2012, Catania, Italy, July 9-13, 2012, Proceedings, Part IV 14*, pages 338–346. Springer, 2012.

[13] Gilles Dejaegere and Yves De Smet. A new threshold for the detection of possible rank reversal occurrences in promethee ii rankings. *International Journal of Multicriteria Decision Making*, 09 2021.

[14] Boris Coquelet, Gilles Dejaegere, and Yves De Smet. Analysis of third alternatives' impact on promethee ii ranking. *Journal of Multi-Criteria Decision Analysis*, 31(1-2):e1823, 2024.

[15] Denis Bouyssou. Ranking methods based on valued preference relations: A characterization of the net flow method. *European Journal of Operational Research*, 60(1):61–67, 1992.

[16] Gilles Dejaegere, Mohamed Boujelben, and Yves De Smet. An axiomatic characterization of promethee ii's net flow scores based on a combination of direct comparisons and comparisons with third alternatives. *Journal of Multi-Criteria Decision Analysis*, 03 2022.

Crew Allocation and Scheduling Optimization for Freight Trains: A Case Study in Chile

Javier Santibáñez-Molina
University of Chile, Center for Mathematical Modelling (CMM).
e-mail: jsantibanezm@cmm.uchile.cl

Vicente Acuña
University of Chile, Center for Mathematical Modelling (CMM)
and Department of Mathematical Engineering (DIM).
e-mail: viacuna@dim.uchile.cl

Jorge Amaya
University of Chile, Center for Mathematical Modelling (CMM)
and Department of Mathematical Engineering (DIM).
e-mail: jamaya@dim.uchile.cl

This work presents a model for the crew scheduling problem in the context of train operations for a company in Chile. The model addresses specific requirements arising from regulatory constraints for railway material transport operators. Each week, a fixed set of trips must be completed, while adhering to restrictions on work schedules and mandatory rest periods for employees. In particular, every worker must have at least one day off in any 7-day window and at least two Sundays off per month.

To ensure a fair workload distribution, the proposed model introduces a rotational schedule where crews alternate positions at the start of each week. This setup translates into a large-scale binary linear optimization problem, designed to assign drivers to trips optimally while balancing weekly working hours among crews.

The results demonstrate significant improvements: a 1% increase in trip coverage, a 50% extension in available scheduling weeks, and a reduction in workload imbalance of 78%. Specifically, the difference in weekly working hours between the most and least loaded workers was reduced from 27 hours, in the manual solution, to just 6 hours. Additionally, the model provides faster solution times and greater flexibility for simulating various scenarios.

Multi-Objective Optimization for the Preventive Tasks Assignment in the Chilean Safety Association

Fernando Arenas-Araya
University of Chile, Center for Mathematical Modeling (CMM)
e-mail: farenas@cmm.uchile.cl

Javier Santibáñez-Molina
University of Chile, Center for Mathematical Modeling (CMM)
e-mail: jsantibanezm@cmm.uchile.cl

Jorge Amaya
University of Chile, Center for Mathematical Modeling (CMM)
and Department of Mathematical Engineering (DIM)
e-mail: jamaya@dim.uchile.cl

This work is oriented to explore optimization models of human resources allocation to improve the distribution of monthly tasks by the Chilean Safety Association (ACHS) to its prevention experts. Those experts are distributed in the different ACHS offices around Chile to periodically visit different companies (clients) based on the districts that they have assigned. Because of the number and duration of the tasks and the time spent traveling, ACHS has its relative weight to the companies (clients) and the prevention tasks to prioritize in the monthly assignment.

The organizational prioritization could be connected to the prevention experts' preferences and include their advice based on experience. For this purpose, interviews were conducted with prevention experts to understand their preferences when they receive and perform tasks, and we included this information in the generation of various mathematical formulations [1].

The human resource allocation problems have been solved using exact, heuristic, and metaheuristics approaches [2]. This work achieves the resolution through mixed-integer linear programming (MILP). Specifically, we tested three different formulations, whose purposes are: Model 1 (multi-objective), maximizing the assignment of critical companies (clients) and the most relevant tasks; Model 2 (mono-objective), maximizing the assignment of critical companies with constraints based on task relevance; Model 3 (multi-objective), maximizing the assignment of critical companies, relevant tasks, and head offices (ACHS) and companies coverage based on historical assignment.

The comparison of the different formulations indicates that Model 3 has the best performance according to the current purposes of the ACHS, increasing the number of tasks assigned by up to 12% and of companies visited by up to 14% in the scenarios studied, compared with the current strategy assignment. Precisely, as the implemented and validated model is multi-objective, the relative importance of the different criteria is incorporated as parameters based on the strategic decisions of the ACHS decision-makers, allowing the generalization of multiple scenarios to compare their advantages and disadvantages through different indicators.

REFERENCES:

[1] Center for Mathematical Modeling (CMM), University of Chile. Exploración de modelos de optimización para mejorar la efectividad de la labor preventiva de los expertos en prevención de los OAL (ACHS). Research and Innovation Projects series of Superintendence of Social Security - SUSESO, Chile (2024). https://www.suseso.cl/619/w3-article-732222.html

[2] Bouajaja, S., Dridi, N. A survey on human resource allocation problem and its applications. Oper Res Int J 17, 339-369 (2017). https://doi.org/10.1007/s12351-016-0247-8

An efficient algorithm for the truck driver scheduling problem

Niels De Walsche, Greet Vanden Berghe, Pieter Smet
KU Leuven, Numerical Analysis and Applied Mathematics (NUMA)
e-mail: {niels.dewalsche, greet.vanden.berghe, pieter.smet}@kuleuven.be

In recent years, there has been an increased focus on the wellbeing of truck drivers, who are often on the road for long periods of time. The European Union has imposed regulations to ensure the safety and wellbeing of such drivers, namely (EC) No. 561/2006. These regulations are also important for drivers' employers, who risk significant penalties if their drivers are not compliant. Goel and Vidal [1] highlighted the fact that adopting hours-of-service regulations can more than double the total travel duration. It is therefore crucial for companies to carefully optimize the routes of their drivers.

The truck driver scheduling problem (TDSP), introduced by Goel [2], aims at creating schedules that comply with the aforementioned European Union regulations. Ensuring compliant schedules represents a crucial component of the Vehicle Routing and Truck Driver Scheduling Problem (VRTDSP), where the goal is to find routes for the trucks and schedules for the drivers that minimize both the number of drivers and the total distance traveled.

We propose an algorithm for solving the TDSP within the framework of a Ruin & Recreate vehicle routing heuristic. Making use of the properties of such a heuristic allows us to solve the TDSP in a computationally efficient manner. The algorithm outperforms the current state-of-the-art heuristic approach in terms of both solution quality and computational time. We demonstrate the effectiveness of our algorithm on a set of benchmark instances from the literature.

References

[1] A. Goel and T. Vidal, "Hours of service regulations in road freight transport: An optimization-based international assessment," *Transportation science*, vol. 48, no. 3, pp. 391–412, 2014.

[2] A. Goel, "Vehicle scheduling and routing with drivers' working hours," *Transportation Science*, vol. 43, no. 1, pp. 17–26, 2009.

A Deep Learning Approach for Analyzing Visual and Textual Content in Tintin Comics

Azat Şaşkal
Solventure
e-mail: asaskal@gmail.com

Toon Van Camp
Ghent University, Research Group Data Analytics
FlandersMake@UGent-corelab CVAMO
e-mail: toon.vancamp@ugent.be

Lukas De Kerpel
Ghent University, Research Group Data Analytics
FlandersMake@UGent-corelab CVAMO
e-mail: lukas.dekerpel@ugent.be

Dries F. Benoit
Ghent University, Research Group Data Analytics
FlandersMake@UGent-corelab CVAMO
e-mail: dries.benoit@ugent.be

Comics analysis investigates the interplay of visual and textual elements to uncover narrative structures and storytelling mechanisms. Traditional approaches in comic art studies predominantly rely on manual methods to summarize content through these multimodal cues (Sugishita & Masuda, 2023). However, manual labeling and panel analysis are both time-intensive and laborious, underscoring the need for automated solutions to enable large-scale and efficient analysis. The field of Comic Recognition (CR) has advanced automation in comics analysis by leveraging deep learning techniques to detect panels, characters, speech balloons, and text (Lenadora et al., 2019; Li et al., 2024; Nguyen et al., 2018). Nevertheless, current CR approaches overlook a crucial aspect: they are not explicitly trained to identify the main characters within a comic book series, instead focusing on detecting whether an object is a character in general. Moreover, publicly available datasets in CR literature are typically derived from a mix of comic series and lack annotations specific to individual characters. These limitations hinder the ability to conduct systematic analyses of comics and their underlying narratives.

This study proposes a novel framework for automated extraction and analysis of narrative elements across an entire comic book series, applied to *The Adventures of Tintin*. The contributions of this research are threefold: (1) comprehensive data collection and segmentation of panel images across the entire

series, (2) the development of a deep learning pipeline for detecting main characters, speech balloons, speaker-balloon associations, and textual content, and (3) the use of extracted data in a character network analysis to investigate social and emotional dynamics throughout the series.

The methodology combines edge detection for panel extraction with YOLOv11 models for character and balloon detection. Furthermore, we developed an algorithm to associate speech balloons with the corresponding characters by detecting the tail point, identified as the contour point with the smallest angle. The algorithm then verifies whether the extended lines of this angle intersect the character's bounding box. Transformer-based OCR is employed for text recognition within the balloons. This pipeline generates a structured dataset wherein each panel includes metadata on main characters, associated text, and their interactions. The results of a comparative evaluation against the ground truth show the effectiveness of this framework in accurately extracting visual and textual elements from the panels of the *Tintin* comics. The extracted dataset is used for a character network analysis, leveraging text recognition, speaker-balloon associations, and character identification to examine narrative dynamics, social structures, and emotional interactions across *The Adventures of Tintin*. This interdisciplinary framework bridges the gap between machine learning and sequential art studies, offering new insights into the analysis of visual storytelling.

References

Lenadora, D., Ranathunge, R., Samarawickrama, C., De Silva, Y., Perera, I., & Welivita, A. (2019). Comic digitization through the extraction of semantic content and style analysis. *2019 19th International Conference on Advances in ICT for Emerging Regions (ICTer)*, *250*, 1–9.

Li, Y., Aizawa, K., & Matsui, Y. (2024, April). Manga109Dialog: A Large-scale Dialogue Dataset for Comics Speaker Detection [arXiv:2306.17469].

Nguyen, N.-V., Rigaud, C., & Burie, J.-C. (2018). Digital Comics Image Indexing Based on Deep Learning. *Journal of Imaging*, *4*(7), 89.

Sugishita, K., & Masuda, N. (2023). Social network analysis of manga: Similarities to real-world social networks and trends over decades. *Applied Network Science*, *8*(1), 79.

Unlocking Real Estate Insights with Large Language Models

Margot Geerts
KU Leuven, LIRIS
e-mail: margot.geerts@kuleuven.be

Manon Reusens
KU Leuven, LIRIS

Bart Baesens
KU Leuven, LIRIS

Seppe vanden Broucke
UGent, Department of Business Informatics

Jochen De Weerdt
KU Leuven, LIRIS

With residential real estate valued nearly three times as much as global GDP in 2022[1], the real estate market plays a critical role in the global economy, influencing various stakeholders such as homeowners, investors, and policymakers. However, significant information asymmetry exists, where sellers typically hold more market knowledge than buyers, leading to inefficiencies in pricing and reduced market liquidity due to ill-informed decisions [3]. Data-driven approaches, especially those leveraging Machine Learning (ML), have the potential to address this issue by automating real estate valuations [1]. However, these methods require both advanced expertise and substantial datasets, which may be inaccessible to many stakeholders.

Large Language Models (LLMs), trained on extensive datasets [2] and adopted more and more for a myriad of use cases across the world[2], emerge as potential tools to democratize access to real estate data, offering insights and improving transparency. LLMs, particularly adept at regression tasks with In-Context Learning (ICL) [4], can help bridge knowledge gaps, thereby enabling more precise pricing, improving market liquidity, and aiding informed decision-making for various stakeholders. This study examines LLMs' potential in real estate valuation by exploring different prompting approaches with ICL and evaluating various pre-trained LLMs on international housing datasets, focusing on prediction accuracy, price range delineation, and explanatory capabilities.

We show that, provided with relevant examples and high-quality market reports, LLMs can extract pricing patterns concerning the hedonic variables, resulting in competitive prediction accuracy, surpassing k-Nearest Neighbor (kNN) baselines and approaching the performance of Gradient Boosted Trees (GBT). In addition, we show that LLM-generated explanations are not only consistent with established ML models but also offer a degree of trustworthiness, as they reflect

[1] https://www.savills.com/impacts/market-trends/the-total-value-of-global-real-estate-property-remains-the-worlds-biggest-store-of-wealth.html

[2] https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai

key predictive drivers identified through robust, model-agnostic interpretability methods like SHAP. Despite this, LLMs tend to be overconfident in generating prediction intervals and perform less effectively than Conformal Prediction procedures with ML models. Finally, we explore different research directions for improving LLMs' performance in real estate valuation, offering a user-friendly and accessible solution that could enhance decision-making and market transparency for real estate stakeholders.

References

[1] João A. Bastos and Jeanne Paquette. On the uncertainty of real estate price predictions. *Journal of Property Research*, pages 1–19, 2024.

[2] Abhimanyu Dubey, Abhinav Jauhri, Abhinav Pandey, Abhishek Kadian, Ahmad Al-Dahle, Aiesha Letman, Akhil Mathur, Alan Schelten, Amy Yang, Angela Fan, et al. The llama 3 herd of models, 2024.

[3] Pablo Kurlat and Johannes Stroebel. Testing for information asymmetries in real estate markets. *The Review of Financial Studies*, 28(8):2429–2461, 2015.

[4] Robert Vacareanu, Vlad-Andrei Negru, Vasile Suciu, and Mihai Surdeanu. From words to numbers: Your large language model is secretly a capable regressor when given in-context examples, 2024.

Reimagining Political Alignment: A Novel Framework for Positioning Parties on the Left-Right Spectrum

Dylan Van Mulders
Ghent University, Department of Marketing, Innovation and Organisation
FlandersMake@UGentcorelab CVAMO, Belgium
e-mail: dylan.vanmulders@ugent.be

Matthias Bogaert
Ghent University, Department of Marketing, Innovation and Organisation
FlandersMake@UGentcorelab CVAMO, Belgium
e-mail: matthias.bogaert@ugent.be

Dirk Van den Poel
Ghent University, Department of Marketing, Innovation and Organisation
FlandersMake@UGentcorelab CVAMO, Belgium
e-mail: dirk.vandenpoel@ugent.be

Social media platforms have offered politicians and political parties new opportunities to communicate and interact with citizens. Especially in Belgium, political parties and politicians are spending a lot of money to social network sites to attract voters. Backed by Belgian party financing, Belgian parties have been among the biggest spenders in Europe, even outside of campaign periods [AdLens, 2022].

This study proposes an alternative approach to positioning parties within the Belgian multi-party system, motivated by the observation that different methodologies can yield divergent conclusions about party positions [Ecker, Marcelo, and Muller, 2021]. The complexities involved in estimating party stances have significant implications for understanding electoral behavior and policy preferences, particularly in multiparty systems. By developing a novel framework for positioning parties, this research aims to address these discrepancies and provide a more nuanced understanding of the Belgian party system's dynamics. We'll focus on what politicians and their party are talking about on the social media platform X, formerly known as Twitter, from January 2021 until December 2022. Supported by the Academic Twitter API, we analyzed the posts of 74 Belgian politicians as well as posts from all the major political party profiles for a period of 2 years.

To achieve this, we propose a three-step approach for clustering political parties based on their online communication. First, we leverage state-of-the-art large language models to generate multilingual embeddings that capture the semantic meaning of party messages given the unique Belgian multilingual context. To

mitigate potential language bias, we benchmark these multilingual embeddings against a monolingual counterpart. The most effective representation is then selected through a comparative analysis of the resulting embeddings. In the next step, we employ topic modeling techniques to aggregate closely related parties based on the co-occurrence of specific topics across different parties and/or party families. This approach provides a nuanced understanding of party positions and alignments, offering novel insights into the complex Belgian political landscape that extend beyond traditional social-economic and socio-cultural dimensions.

This study introduces a novel approach to analyzing the online political communication of parties using social media data, enabling a deeper understanding of their positional alignments beyond a two-party system and grounded in the actual content of their messaging. By leveraging social media as a primary source, we overcome the limitations of traditional survey-based methods, which often rely on pre-defined questions and may introduce biases towards specific topics. Our methodology allows for a more objective and dynamic assessment of party positions over time, unbound by the constraints of electoral cycles or static policy programs. This enables us to capture parties' stances on emerging issues and provide a more nuanced understanding of their relative distances to one another. Ultimately, this study seeks to contribute to the ongoing debate on party system analysis and offer valuable insights for researchers and practitioners seeking to interpret electoral behavior and policy preferences in complex multiparty environments.

References

[AdLens, 2022] AdLens. (2022). The AdLens report on political ads on Facebook & Instagram in Belgium, 2021, https://adlens-be.medium.com/the-adlens-report-on-political-ads-on-facebook-instagram-in-belgium-2021-8b275c55895a.

[Ecker, Marcelo, and Muller, 2021] Ecker, A., Jenny, M., Müller, W. C., & Praprotnik, K. (2021). How and why party position estimates from manifestos, expert, and party elite surveys diverge: A comparative analysis of the leftright and the European integration dimensions. Party Politics, 28(3), 528-540. https://doi.org/10.1177/1354068821990298.

Gen AI vs. NLP for Legal Entity Extraction

Porchourng Chuor, Ashwin Ittoo
HEC ULiege, Université de Liège, Liège, Belgium
e-mail: porchourng.chuor@uliege.be, ashwin.ittoo@uliege.be

Policies and regulations, such as those from the EU, are subject to regular updates and amendments. This evolution leads to an increased demand for data in the form of reporting requirements from the concerned parties, resulting in higher regulatory reporting burdens. Urgent needs to streamline regulatory burdens to harmonize reporting requirement. For policy-making institutions, such as the European Commission (EC), the first step towards this goal is to standardize the regulatory reporting metadata of EU legislation. Typically, reporting obligations specify the information that a reporting entity must disclose to a regulatory entity by a given date. Examples are in Figure 1. As can be seen, reporting obligations expressed in natural language tend to have a complex sentential structure. The two reporting obligations presented are very different and do not follow any predefined templates. Different entities, such as the information to be reported and the regulatory entity, can play the same syntactic role. This lack of structure makes it difficult for humans to read and interpret. More importantly, it hinders the development of methods for automating the reporting process. Therefore, in this work, we present a framework for transforming reporting obligations into a structured format. A key step to achieve this objective is the extraction of relevant entities from reporting obligations. This is the problem we aim to address. In our use case, financial reporting regulation, we focus on the following entities: *Addresser*: Who reports what to whom; *Action*: What action that is performed; *ActionResult*: What is done; *Addressee*: To whom the action is directed; *Date*: When the action occurs or deadline. A popular corpus of regulatory reporting obligations is Eur-Lex[1], which is maintained by the EC. However, the majority of such corpora are not labeled, precluding the use of supervised approaches. In our study, we investigate two methods to address this issue. First, we employ a rule-based syntactic dependency parsing approach. Second, we rely on few-shot prompting with a pre-trained Large Language Model (LLM) [2], which has shown much promise in information extraction (IE) [1]. Several LLMs have been proposed recently. However, an important research question, which is yet to be addressed, which LLM achieves better performance in IE regulatory reporting. Another question is whether an LLM can outperform a rule-based syntactic parsing method. In this paper, we address these research questions. Specifically, in our syntactic parsing approach, we first generated the syntactic dependencies in the reporting obligations. Next, we used syntactic roles to identify entities. For example, a noun sequence as the subject of a main verb is identified as the *Addresser* (e.g. *competent authority of the Member States* in Figure.1). Other

[1] https://eur-lex.europa.eu/homepage.html

entities, such as the *Addressee, ActionResult*, and *Action*, are detected in a similar manner. Concerning our LLM approach, we benchmarked the performance of two LLMs, namely, Llama-3-8B-Instruct[2] (we refer to this model as Llama3) and ChatGPT-4o [3], the latest version of the GPT model. We adopted a few-shot prompting strategy composed of three components: ***{context, entity explanation, exemplars}***. Our experiments were conducted on a subset of the Eur-Lex corpus. The best performance achieved by the syntactic parsing approach was an F1-score of 0.62 for detecting the *Action* entity from the reporting obligations, while its worst performance was 0.07 for the *Date* entity. Furthermore, this approach faced several challenges, including errors in syntactic parse trees and complex dependency structures that made the parsing process brittle and difficult to generalize. Different sentence structures required significant customization of the tree traversal process. Conversely, the performance of both LLMs, namely Llama3 and ChatGPT-4o, was significantly better. Llama3 achieved its highest F1-score of 0.98 for detecting the *Action* entity, while its lowest performance was an F1-score of 0.48 for the *ActionResult* entity. ChatGPT-4o achieved its best and worst performances on these same entities, with F1-scores of 0.84 and 0.27, respectively. An interesting observation is that, overall, Llama3 outperforms ChatGPT-4o, despite being a smaller model. Our key contributions are as follows: i) We investigated the performance of a standard NLP approach based on syntactic rules against LLMs for IE from reporting obligations; ii) As LLMs, we focused on Llama3 and ChatGPT-4o, demonstrating that Llama3 outperforms ChatGPT-4o despite its smaller size; iii) we provide a thorough analysis of the results from both the LLMs and the syntactic parsing approach.

Example 1: Individual suspected adverse reaction reports and follow-ups submitted to the Eudravigilance database by marketing authorisation holders shall be transmitted electronically upon receipt to the competent authority of the Member State where the reaction occurred.
Example 2: Within 30 days of receipt of the assessment report, the marketing authorisation holder and the members of the Pharmacovigilance Risk Assessment Committee may submit comments to the Agency and to the rapporteur.

Figure 1: Here are examples of regulatory reporting, where each color represents a different entity: red for the *Addresser entity*, green for the *Action entity*, blue for the *ActionResult entity*, purple for the *Addressee entity*, and orange for the *Date entity*.

References

[1] Monica Agrawal et al. Large language models are few-shot clinical information extractors. In Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing, pages 1998-2022, December 2022.

[2] Tom Brown et al. Language Models are Few-Shot Learners. In Advances in Neural Information Processing Systems, volume 33, pages 1877-1901, 2020.

[2] https://huggingface.co/meta-llama/Meta-Llama-3-8B-Instruct
[3] We used the OpenAI API.

Multi-objective Traveling Salesman Problem for in-orbit spacecraft routing

Pieter Deleye
Universiteit Antwerpen, Engineering Management Department,
Antwerp Operations Research group (ANT/OR)
e-mail: pieter.deleye@uantwerpen.be

Christof Defryn
Universiteit Antwerpen, Engineering Management Department,
Antwerp Operations Research group (ANT/OR)
e-mail: christof.defryn@uantwerpen.be

Currently, more than 40,000 objects larger than 10 cm are being tracked by ESA orbiting Earth, from which more than one-fourth are active satellites. These satellites play a vital role in e.g. global communication and Earth observation. The number of active satellites is expected to rise sharply in the near future, with Euroconsult projecting another 25,000 satellites to be launched by 2032 (Elmasry, 2023). Meanwhile, space debris, currently estimated at over 900,000 objects larger than 1 cm, threatens operational safety as even the slightest impact comes with huge consequences. (ESA Space Debris Office, 2024)

One of the strategies to ensure sustainable space activities for future generations is collision avoidance manoeuvres. These are common practice, with SpaceX reporting on average 137 of such manoeuvrers per day in the first half of 2024 (Pultarov, 2024). Further remediation measures involve satellite nudging with lasers or the use of dedicated spacecraft for refuelling, repair, and transportation. The focus of this research is on the route optimization of these service spacecraft. As decision makers face the constant trade-off between fuel consumption and mission duration, a multi-objective optimization framework is adopted.

The problem of determining a sequence of in-orbit objects to be visited by a servicing spacecraft is closely related to the definition of a Travelling Salesman Problem (Bourjolly et al., 2006). Spacecraft routing, however, differs fundamentally from terrestrial logistics due to the unique dynamics of orbital mechanics. Satellites move at high velocity to retain their orbits and only orbit changes require fuel. Based on the dynamics of the problem, every transfer trajectory between satellites is time-dependent and changes in earlier trajectories require a complete recalculation of all subsequent trajectories, which is highly computationally expensive. Since the 1960s, researchers have explored pairwise optimal spacecraft routing with multiple proven efficient formulas. Specifically, the Lambert transfer problem allows to determine a specific trajectory between a starting point towards a desired position in time (Prussing & Conway, 2013). Our research is focused on the integration of these (highly non-linear) formulas into the definition of the TSP to support decisions for spacecraft routing.

Based on the dynamics of a two-body problem, the spacecraft routing problem is defined as a multi-objective, state- and time-dependent TSP with dual objectives. Each pairwise transfer between satellites entails an infinite set of possible trajectories, which is tackled by calculating feasible pairwise trajectories only within a bounded timeframe. Further refinement reduces the trajectories to only non-dominated ones either in fuel or time. Notably, the feasibility of each trajectory depends dynamically on its initial state, introducing complexity through the coupling of time and state dependencies.

This research aims to integrate spacecraft routing into multi-objective optimization techniques within the TSP framework. The proposed methodology addresses challenges in satellite servicing and debris remediation but also sets the stage for more scalable, sustainable solutions as space becomes an increasingly vital domain for global operations.

References

Elmasry, F. (2023, December 15). Four tons of satellites to be launched daily by 2032, demand concentrates by a handful of players. *Novaspace*. Retrieved from https://nova.space/press-release/four-tons-of-satellites-to-be-launched-daily-by-2032-demand-concentrates-by-a-handful-of-players/

ESA Space Debris Office. (2024). *ESA's Annual Space Environment Report* (Report No. GEN-DB-LOG-00288-OPS-SD). ESA Space Debris Office. Retrieved from `https://www.sdo.esoc.esa.int/environment_report/Space_Environment_Report_latest.pdf`

Pultarova, T. (2024, July 23). SpaceX Starlink satellites made 50,000 collision-avoidance maneuvers in the past 6 months. What does that mean for space safety? *Space.com*. Retrieved from https://www.space.com/spacex-starlink-50000-collision-avoidance-maneuvers-space-safety

Bourjolly, J.-M., Gurtuna, O., & Lyngvi, A. (2006). On-orbit servicing: A time-dependent, moving-target traveling salesman problem. *International Transactions in Operational Research, 13*(3), 461-481. https://doi.org/10.1111/j.1475-3995.2006.00558.x

Prussing, J. E., & Conway, B. A. (2013). *Orbital mechanics*. United Kingdom: Oxford University Press.

Service network design under uncertainty: Simulating and predicting delay costs

Javier Durán-Micco
Delft University of Technology, Department of Maritime and Transport Technology
e-mail: j.duranmicco@tudelft.nl

Bilge Atasoy
Delft University of Technology, Department of Maritime and Transport Technology
e-mail: b.atasoy@tudelft.nl

This work addresses the service network design (SND) problem from the perspective of logistics service providers (LSPs). The objective is to optimize the selection of transport requests and capacity planning in multimodal services. While previous research mostly focus in stochastic programming techniques to handle uncertainty in the SND [1], in our study we use simulation models due to their flexibility in evaluating multiple uncertainty levels and complex interactions among different players in the system.

We designed a simulated annealing (SA) algorithm to solve the SND problem, which optimizes the selection of requests, capacity booking on trains and barges, and produces a transport plan intended to be repeated periodically by the LSP. To better approximate the real-world performance of such plans, we propose a simulation model that accounts for travel time uncertainties and real-time re-planning due to delays or missed connections. Although this simulation provides a more accurate estimate of expected costs, it is computationally intensive and impractical for evaluating every candidate solution during the SA execution. To address this limitation, we aim to develop a deterministic heuristic that can rapidly estimate additional costs arising from operational uncertainties and re-planning. A pool of feasible solutions for a given test network is generated using a deterministic algorithm, which generates solutions of varying quality and under different problem parameters. These solutions are then evaluated using the simulation model to identify relationships between solution characteristics and the additional costs incurred due to the operational uncertainties. Our findings indicate that the magnitude of these costs primarily relate to limited truck fleet availability and travel time uncertainties. To capture these characteristics, we define a variable η that synthesizes key factors: available truck hours, assigned truck hours, and the relative size of time windows for each truck leg compared to the corresponding expected travel time. This variable can be quickly calculated for a given solution and its tactical plan. Figure 1 displays η against the addi-

[1]Funded by the European Union (ERC, ADAPT-OR, No 101117675). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council. Neither the European Union nor the granting authority can be held responsible for them.

tional cost per container determined by the simulation, for each solution. The plot suggests a systematic relationship between these values. Subsequently, we fit various functions to this data to estimate total additional costs. The best results are achieved using a cubic function and a piecewise function with a linear and a quadratic segment. These functions enable rapid evaluation of solutions, estimating additional costs due to uncertainties without needing directly incorporating stochastic elements or solving truck routing, thus significantly reducing computational time.

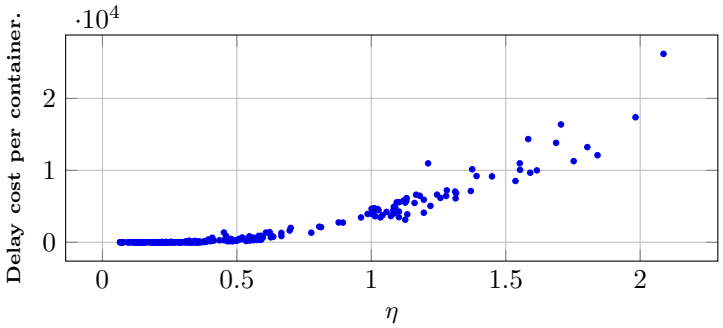


Figure 1: Relationship between η and delay cost predicted by the simulation model.

Table 1 compares average results generated by the SA applied to instances of different sizes. Different SA settings were tested, regarding the evaluation procedure: deterministic method without penalization (SA_{det}); penalization using cubic fitting (SA_{cub}); penalization using piecewise fitting (SA_{pw}); and using simulation for each candidate solution (SA_{sim}). In all cases, the best solution obtained at the end of the SA execution was re-evaluated using the simulation, which is the value compared between settings. While SA_{sim} yields the best results (max profits for LSP), it is computationally expensive. On the other hand, SA_{det} is much faster, but in some cases generates extremely bad solutions due to completely disregarding the potential delay costs. In contrast, the proposed penalization methods (SA_{cub} and SA_{pw}) achieve a balance, delivering high-quality solutions efficiently. Future work will focus on refining penalization calculations and developing adaptive algorithms that dynamically adjust the parameters of the function used in the estimation. Although SA_{pw} performs slightly better, SA_{cub} could be more appropriate for adapting methods due to its simpler function, which makes parameter updates easier.

Instance	**Profits (€ × 10^3)**				**CPU time (s)**			
(n^o requests)	**SA_{det}**	**SA_{cub}**	**SA_{pw}**	**SA_{sim}**	**SA_{det}**	**SA_{cub}**	**SA_{pw}**	**SA_{sim}**
50	181.2	317.6	323.0	348.4	0.13	0.10	0.12	59.86
100	817.2	808.4	822.5	824.5	0.18	0.16	0.19	125.02
200	458.1	1352.9	1375.2	1484.9	0.52	0.35	0.48	748.16
400	-1610.1	2512.1	2648.2	2760.8	1.80	0.91	1.36	2398.19

Table 1: SA average results with different evaluation procedures.

References

1. Delbart, T., Molenbruch, Y., Braekers, K., & Caris, A. (2021). Uncertainty in Intermodal and Synchromodal Transport: Review and Future Research Directions. *Sustainability, 13*(7), 3980. https://doi.org/10.3390/su13073980

Automatic model decomposition in Hexaly

Julien Darlay
Hexaly, 251 Boulevard Pereire, 75017 Paris
e-mail: `jdarlay@hexaly.com`

Hexaly Optimizer is a model-and-run solver based on various heuristics and exact methods. A set-based modeling formalism was introduced to simplify the modeling of some combinatorial problems such as routing, scheduling or packing problems. This talk presents how Hexaly uses automatic model decomposition and branch price and cut to compute lower bounds for several vehicle routing problems.

1 Modeling routing problems with set decisions

Vehicle routing problems can be modeled using `list` decisions in Hexaly's modeling language. A `list(n)` represents an ordered subset of integers between 0 and n (excluded) and can model the sequence of customers visited by one truck. A `partition` constraint on all the lists ensures that each customer is served by exactly one truck. Finally, functional expressions can be used to model capacity constraints and the total traveled distance to model a standard CVRP. Other expressions can be used to model VRP variants, such as time windows, fix costs per truck, etc.

Set-based models are compact with one `list` decisions per truck. This model is well suited for a heuristic search but is much more difficult to solve with a mathematical programming approach to compute lower bounds.

2 Approach

Some set-based models with a `partition` constraint can be expressed as set partition problems with an exponential number of binary variables. Each variable models the selection of a feasible list assignment in the set-based model. A first block of set partition constraints ensures that each client is visited by exactly one selected assignment. A final constraint ensures that at most one assignment per list is selected. Solving the explicit model is not tractable in practice, but efficient techniques such as column generation and branch price and cut [1, 2] can compute a lower bound on this problem. Solutions are computed with a heuristic approach.

Hexaly automatically performs this reformulation on compatible set-based models, discarding incompatible constraints to compute lower bounds on most vehicle routing problems.

3 Results

The method is benchmarked on standard instances from the literature [3] for CVRP and CVRPTW. The solver is launched on all instances for 60s and we measure the gap reported when the time limit is reached. For instances with more than 200 clients the approach is not competitive with a standard arc model and subtour elimination constraints and the solver falls back to the arc model. Table 1 shows for each dataset the percentage of instances with a lower bound computed using automatic model reformulation and the average gap reported by this technique after 60s.

Type	Instances	Compatible inst.	Average gap
CVRP	Set A	100%	0.11%
CVRP	Set B	100%	0.06%
CVRP	Set X ($n \leq 200$)	72.73%	3.74%
CVRPTW	Solomon	59%	0.6%
HVRP	DLP set	32.29%	10.9%

Table 1: Average gap reported by Hexaly with automatic model reformulation in 60s

After 60 seconds on CVRP and CVRPTW instances with less than 200 customers, Hexaly reports near-optimal solutions and bounds. This approach can be generalized to other practical vehicle routing constraints, such as pickup and delivery, site-dependent VRP and other set-based models.

References

[1] Artur Alves Pessoa, Ruslan Sadykov, Eduardo Uchoa, and François Vanderbeck. A Generic Exact Solver for Vehicle Routing and Related Problems. *Mathematical programming*, 183:483–523, 2020.

[2] Jacques Desrosiers, Marco Lübbecke, Guy Desaulniers, Jean Bertrand Gauthier. Branch-and-Price 10.13140/RG.2.2.29888.14088.

[3] http://vrp.galgos.inf.puc-rio.br/index.php/en/updates

Addressing challenges in the vehicle routing problem for grocery delivery

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

Since the Covid pandemic, home delivery of groceries has been on the rise. Vehicle Routing in the context of grocery delivery presents unique logistical and computational challenges due to the perishable nature of the goods and the requirements to keep them fresh or even frozen, strict time windows and the absolute requirement of the customer to be home to accept the delivery, and the complexity of customer preferences. In this talk, we overview the literature on the topic of grocery delivery vehicle routing. We find that — despite significant advancements — substantial research gaps persist in effectively addressing key practical difficulties. This paper highlights these gaps and proposes future directions to enhance research in this area.

Research has primarily focused on minimizing costs and emissions, but critical areas such as stochastic service times, caused by customer availability uncertainties, remain underexplored. Current models inadequately integrate real-world dynamics like fluctuating customer availability within prescribed time windows, a significant factor in ensuring delivery success. Similarly, the handling of undelivered grocery products, which demand specific storage and re-routing considerations, lacks robust solutions. Multi-compartment vehicles, essential for transporting diverse grocery types under varying storage conditions, require more sophisticated algorithms to optimize compartment utilization dynamically in addition to syncronization and scheduling with grocery inventory management.

Future research must also address dynamic adjustments in delivery routes influenced by real-time data, such as traffic and energy constraints for electric vehicles. Advanced techniques integrating machine learning and hybrid metaheuristics can potentially improve grocery delivery vehicle routing by providing faster and more accurate solutions. Incorporating stochastic models and adaptive algorithms to manage uncertainties and constraints dynamically will bridge the gap between theoretical optimization and practical application.

This review underlines the need for comprehensive grocery delivery vehicle routing frameworks that encompass stochastic variables, enhance customer service reliability, and ensure environmental sustainability while addressing the operational challenges unique to grocery logistics. By addressing these gaps, research in this area can significantly contribute to creating resilient and efficient grocery delivery systems.

Investigating the Monte-Carlo Tree Search Approach for the Job Shop Scheduling Problem

Laurie Boveroux
University of Liège, Department of Electrical Engineering and Computer Science
e-mail: laurie.boveroux@uliege.be

Damien Ernst
University of Liège, Department of Electrical Engineering and Computer Science
e-mail: dernst@uliege.be

Quentin Louveaux
University of Liège, Department of Electrical Engineering and Computer Science
e-mail: q.louveaux@uliege.be

The Job Shop Scheduling Problem (JSSP) is a well-known optimization problem in manufacturing, where the goal is to determine the optimal sequence of jobs across different machines to minimize a given objective. In a general JSSP, we are given a set $\mathcal{J}$ of n jobs $J_1, J_2, ..., J_n$ and a set $\mathcal{M}$ of m machines. Each job $J_i \in \mathcal{J}$ has an operation set $\mathcal{O}_i$ which contains n_i operations O_{ij} that must be processed in a specific order (i.e., with precedence constraints). Each operation O_{ij} of job J_i requires a processing time p_{ij} on a specific machine $M_{(ij)}$. A job can have several operations that must be processed on the same machine (i.e., recirculation). Each machine can process at most one operation at a time with no preemption. In this work, we focus on minimizing the weighted sum of job completion times.

We explore the potential of Monte Carlo Tree Search (MCTS), a heuristic-based reinforcement learning technique, to solve large-scale JSSPs. MCTS is a heuristic search algorithm used in decision processes. It combines classic tree search implementations with machine learning principles of reinforcement learning to balance exploration and exploitation. The algorithm is based on the building of a search tree. Specifically the MCTS algorithm can be used to solve Markov Decision Processes (MDP) as it incrementally builds a search tree representing the states and actions of the MDP, using simulation-based techniques to evaluate potential policies.

The first contribution of the work is to explore different environments for the MCTS applied to the JSSP. Each environment is defined by its state space, action space and reward function. These three components can be defined in multiple ways, each influencing the performance and behaviour of the MCTS algorithm differently. We explore a few of possible combinations of these components. In particular, we introduce two state representations for the partial schedule:

1. Absolute Representation: This representation maintains the completion

times of each operation, directly encoding the timing information of the partial schedule. This approach is greedy regarding the order of operations and the completion times.

2. Relative Representation: This representation maintains the order of the operations on each machine, encoding the sequence of operations on each machine rather than their precise timing. This approach is greedy only on the order of operations and it maintains flexibility in terms of scheduling completion time.

Each state representation can be converted in the other. However, each state representation comes naturally with its own set of potential actions and rewards.

We then compare the various environments previously defined with respect to their performance to obtain a good schedule. In order to do so, we introduce a new synthetic benchmark derived from real manufacturing data, which captures the complexity of large, non-rectangular instances often encountered in practice. Our experimental results show that MCTS effectively produces good-quality solutions for large-scale JSSP instances, outperforming our constraint programming approach on those instances.

Fixed Order Scheduling with Deadlines

André Berger, Arman Rouhani (Speaker), and Marc Schröder
Maastricht University, Department of Quantitative Economics
e-mails: {a.berger, a.rouhani, m.schroder}@maastrichtuniversity.nl

1 Introduction and Preliminaries

Scheduling problems are a well-studied field in the domain of operations research. In the literature, various objectives such as *makespan* or *completion time* are examined. Among these, minimizing the number of machines is important in practice. Consider a practical scenario where multiple users utilize a set of parallel processors with shared memory. By scheduling jobs on a certain subset of processors, processors can be put to sleep, thereby reducing energy consumption. In numerous practical applications such as airport gate management, train scheduling, and developing distributed memory architectures, the processing of a set of jobs on a minimum number of machines is desired while the schedule respects the time intervals in which jobs have to be scheduled. These problems are referred to as *SRDM* (scheduling with release times and deadlines on a minimum number of machines). Yu and Zhang [1] provided a 2-approximation algorithm for the common release time variant of SRDM. In various practical applications such as CPU task management, perishable goods handling, airline boarding, and healthcare systems, there is often a requirement to follow an *order* on each machine when scheduling tasks. For example, in CPU scheduling with static priorities, tasks must be executed in a pre-established sequence across processors, following global priority rules. Similarly, in industries dealing with perishable goods, warehouses or delivery systems must follow a predefined order to ensure items are processed and delivered before their expiration. In airline operations, passengers are boarded in a fixed sequence across gates, ensuring smooth and efficient boarding procedures. In healthcare systems, patient appointments are frequently managed in a fixed sequence, with doctors treating patients in the order determined by their schedule, independent of other appointments. These examples demonstrate the importance of efficient scheduling under a fixed global order. In this paper, we address the SRDM problem with common release times and a fixed processing order, denoted as $P|r_j = r, p_j, d_j, \pi^*|m$. We investigate the impact of the imposed processing order on job scheduling. The input to the problem consists of a set of n jobs denoted as $J = \{1, \ldots, n\}$. We also have a sufficiently large supply of identical machines $\{1, \ldots, n\}$. Each job $j \in J$ has a processing time $p_j > 0$, and a deadline $d_j > 0$, where $d_j \geq p_j$. For each job $j \in J$, we define $\lambda_j = d_j - p_j$ as the *slack* representing the difference between the job's deadline and its processing time. The jobs follow a predetermined fixed order for processing, and each machine schedules its assigned jobs according to this order. The objective of the problem is to minimize the number of machines necessary for scheduling the jobs while ensuring that the prescribed order is maintained on

each machine, and every job is completed within its stipulated deadline. We will study two simple greedy algorithms. The *first-fit (FF)* algorithm schedules jobs based on the fixed order and assigns each job to the first feasible machine. An even simpler greedy algorithm *next-fit* (NF) operates similarly with the exception that NF only considers the last open machine for scheduling a job and, if that job fails to fit in the last machine NF uses a new empty machine for that job.

2 Contribution and Conclusion

In this paper we show that next-fit has an unbounded approximation ratio. Next, we show the optimality of first-fit in the case of unit processing times. Then, we prove that first-fit has an approximation factor of 2 in four different classes of instances: (1) if the fixed order aligns with slacks that are in a non-increasing order, (2) if the fixed order aligns with slacks that are in a non-decreasing order, (3) if the fixed order aligns with deadlines that are in a non-increasing order, and (4) if the optimal solution uses at most 3 machines. Lastly, we give a $O(\log n)$ approximation problem for the general problem. Our results seem to suggest that the first-fit algorithm approximates the optimal solution within a factor of two for arbitrary orders, but we have not been able to give a formal proof. Further research could also examine other algorithms as potential approximation methods to determine if they offer better bounds. Our findings are as follows:

Instance 1. Let $\mathcal{I}_n$ be a parametric instance with $n > 2$ jobs. Assume that $(p_1, d_1) = (1, 1)$ and $(p_2, d_2) = (2, 2)$. For $j > 2$, the pair (p_j, d_j) is obtained by $(p_{j-1} + p_{j-2}, p_j + p_{j-1} - 1)$.

Lemma 1. *Next-fit has an unbounded approximation ratio for the instance $\mathcal{I}_n$.*

Theorem 2. *The FF algorithm solves the problem $P|r_j = r, p_j = 1, d_j, \pi^*|m$ optimally in time $O(n^2)$.*

Theorem 3. *First-fit is a 2-approximation algorithm for non-increasing slacks, and this bound is tight.*

Theorem 4. *First-fit is a 2-approximation for non-decreasing slacks, and this bound is tight.*

Theorem 5. *First-fit is a 2-approximation for non-increasing deadlines.*

Theorem 6. *First-fit is a 3/2-approximation when the optimal schedule opens at most two machines.*

Theorem 7. *First-fit is a 2-approximation when the optimal schedule uses at most three machines.*

Theorem 8. *The problem $P|r_j = r, p_j, d_j, \pi^*|m$ can be approximated within a factor of $O(\log n)$.*

References

[1] Yu, G. & Zhang, G. Scheduling with a minimum number of machines. *Operations Research Letters*. **37**, 97-101 (2009)

Scheduling system tests with a common deadline at minimum cost

Emmeline Perneel
KU Leuven, ORSTAT
emmeline.perneel@kuleuven.be

Roel Leus
KU Leuven, ORSTAT
roel.leus@kuleuven.be

The time-critical testing problem is encountered in various domains such as manufacturing, quality control, and project management. It involves determining the operational state of a multi-component system under a fixed deadline with constrained testing resources.

The multi-component system can be in one of two states, namely *up* (indicating, for example, functionality or health) or *down* (signifying, for example, failure or the detection of a disease). Directly observing the system's state is not possible; instead, we rely on testing the individual components comprising the system to reveal the system's state. Each test incurs a specific cost and has a given probability of success. The costs can encompass both monetary expenses and intangible factors and the probabilities of success are assumed to be independent. The relationship between the components and the system state is described by a so-called n-out-of-n system success function. Simply put, the system is operational (up) only if all of its components are working (up). Conversely, if any component is malfunctioning (down), the entire system is considered to be failing (down). Each test takes unit time and the state of the system needs to be determined within a fixed deadline, meaning that some tests might have to be executed simultaneously. The number of tests that can be executed in parallel is limited by the number of testing units available. This paper addresses the challenge of minimizing the expected cost of determining the system's state by the common deadline.

The case without deadline has already been studied more intensively; tests can then be scheduled sequentially in non-decreasing order of their ratio of cost to probability of failure. With a restrictive deadline, however, multiple tests may need to be executed simultaneously and the problem becomes NP-hard. The main contributions of this paper are as follows. We strengthen existing formulations for the time-critical testing problem, leveraging the insight that once a desirable test partition is determined, ordering the subsets of the partition is straightforward; this insight has already been acknowledged in the literature but has not yet been exploited in formulations or algorithms thus far. Based on these formulations we develop and test branch-and-cut-and-price algorithms.

Fast bounds in Hexaly based on single-machine scheduling problems

Léa Petit-Jean Genat
Hexaly, 251 Boulevard Pereire, Paris, France
e-mail: lpetitjean@hexaly.com

1 Introduction

Given a scheduling problem expressed in the Hexaly framework, our goal is to *quickly* calculate a reasonable lower bound on the objective at the beginning of the solution process, supplementing more computationally expensive bounds computed during the search. By *quickly*, we mean within a very short computation time relative to the overall time allocated for solving the problem. Typically under one second for problems involving several thousand activities. To achieve this, we propose using classical algorithms with a complexity of at most $O(n\log(n))$ on single-machine subproblems.

2 Notation et considered polynomial algorithms

Let $\mathcal{A}$ represent the set of interval variables (activities) in the model defined for Hexaly. For $x \in \mathcal{A}$, $C(x)$ denotes the end time of x and $d(x)$ its minimum duration. Let $\mathcal{P}(x) \subset \mathcal{A}$ denote the set of intervals that must end before $C(x)$ due to precedence constraints, and $\mathcal{M} \subset 2^{\mathcal{A}}$ the set of disjunctive constraints of the problem (machines).

The algorithms considered are Moore-Hodgson [2] and Potts-Wassenhove [1] for problems $(1||U_i)$ and $(1||w_iU_i)$, Smith [3] for $(1||\sum w_iC_i)$, and Jackson [4] for $(1||L_{\max})$. These algorithms have a complexity of $O(n\log(n))$ and yield optimal solutions for their specific problem, except for Potts-Wassenhove's, which is a relaxation.

In the following sections, we illustrate our work using Smith's algorithm. The other algorithms are approached in a similar manner.

3 Example of Smith's algorithm

To recall, Smith's algorithm for solving the problem $(1||\sum w_iC_i)$ involves ordering tasks by increasing d_i/w_i ratios.

All model expressions that can be represented as a weighted sum of interval variable end times are identified. Let $r = \sum w_iC(x_i)$ be one such expression. We compute a lower bound for r as described bellow. It is worth noting that

this approach is not limited to expressions r that participate in the objective function; any expression of this form is subject to a lower bound calculation that propagates throughout the rest of the problem.

Given a machine $m \in \mathcal{M}$, we can calculate a bound on $r = \sum w_i C(x_i)$ using Smith's algorithm by distributing the weights w_i of intervals x_i in r across the intervals on m. For example, consider a model expression $r = w_3C(x_3) + w_4C(x_4) + w_5C(x_5)$ and a machine $m = \{x_1, x_2, x_3\}$. Suppose $\{x_1, x_2, x_3\} \subset \mathcal{P}(x_4)$ and $\{x_1, x_2\} \subset \mathcal{P}(x_5)$. We can select weights to use in Smith's algorithm on machine m that yield a valid bound on r, for instance, by evenly distributing the w_i based on precedences: $\{\frac{w_4}{3} + \frac{w_5}{2}, \frac{w_4}{3} + \frac{w_5}{2}, w_3 + \frac{w_4}{3}\}$. Other distributions can also produce valid bounds, with the highest of these bounds being retained.

4 Results

The approach was extended to other algorithms and objective types mentioned above and tested on classic benchmarks (e.g., job shop), either used directly or modified according to the objective. In the table below, the *Gap* columns indicate the average gap between the lower and best upper bounds, before and after the process described here.

Objective	#	Improvement	Gap Before	Gap After	Difference
U_i	640	84%	66%	48%	-18%
w_iU_i	760	77%	63%	54%	-9%
w_iC_i	652	95%	55%	24%	-31%
$L_{\max}$	700	69%	-	-	-35%

Table 1: Average Lower Bound Improvement by Objective Type.

5 Conclusions and Future Work

The identification of single-machine subproblems suitable for polynomial algorithms helps to improve bounds from Hexaly 12.5. Further research is needed to examine different ways to distribute the weights w_i (and, depending on the objectives, due dates) of expressions r among the machine intervals to achieve even better bounds.

References

[1] C. Potts et L. Van Wassenhove. Algorithms for scheduling a single machine to minimize the weighted number of late jobs. *Management Science 1988.*

[2] J. M. Moore. An n job, one machine sequencing algorithm for minimizing the number of late jobs. *Management Science 1968.*

[3] W. Smith. Various s for single-stage production. *Naval Research Logistics Quarterly 1956.*

[4] J.R. Jackson. Scheduling a production line to minimize maximum tardiness. *Office of Technical Services 1955.*

Dynamic Programming and Block-Cut Tree Decompositions for a Maximum Covering Location-Network Design Problem

Felix Rauh, Jannik Matuschke, Hande Yaman
Faculty of Economics and Business, KU Leuven
e-mail: {felix.rauh,jannik.matuschke,hande.yaman}@kuleuven.be

In order to plan accessible public health services and resilient infrastructure, high quality solutions to network design and facility location problems are of utmost importance. We tackle a problem that arises, for instance, when improving health care access in areas that are prone to disruptions such as recurring floods. A given budget can then be used both for building new hospitals as well as making weak links in the network resilient to floods. The goal is to maximize the number of households that can reach a facility via a flood-resilient access of, e.g., at most five kilometers.

More precisely, we define the *Maximum Covering Location-Network Design Problem (MCNDP)* as follows: Given a graph $G = (V, E_w \cup E_r)$ with weak edges E_w and resilient edges E_r, and possible facility locations $F \subset V$, we aim to choose a set of links $X \subset E_w$ and facilities $Y \subset F$ to open such that the total weight of nodes that are within distance $\overline{d}$ to an open facility is maximized (via so-called resilient edges in $X \cup E_r$). The total allocated budget needs to be bounded by $\overline{b}$, i.e., it must hold that $\sum_{i \in Y} c_i^F + \sum_{e \in X} c_i^e \leq \overline{b}$.

While there is extensive work on a wide range of variants of network design as well as maximum covering location problems (see, e.g., [**?**, **?**]), the combined problem RMCP has received little attention (e.g., [**?**]).

In this talk, we present an exact solution framework for the MCNDP based on dynamic programming (DP), and describe its theoretical properties as well as computational performance. First, we show that the problem on trees allows for an optimal solution approach via dynamic programming that is pseudo-polynomial in the given budget. The method is an extension of DP approaches to the maximum covering and p-median problem [**?**, **?**]. Computational comparisons with two different formulations solved by a commercial solver show that the DP is superior in specific situations.

Another class of graphs that are of interest are graphs that have cut vertices (so-called articulation points), thus allowing for a non-trivial decomposition into a block-cut tree (BC tree) [**?**]. The maximal biconnected components of the graph then correspond to vertices in the BC tree. Motivated by the DP on trees, we introduce an exact approach that identifies the optimal way to combine partial solutions on blocks. Our ongoing work improves on that by using primal and dual bounds on the blocks to create a fast heuristic as well as optimal solutions

in a branch and bound fashion.

We present the computational benefits of the approach via preliminary results of an extensive computational study that compares our approach, e.g., with a branch and cut approach that uses the idea of length-bounded cuts.

References

[1] Teodor Gabriel Crainic, Michel Gendreau, and Bernard Gendron, editors. *Network Design with Applications to Transportation and Logistics*. Springer International Publishing, Cham, 2021.

[2] Sergio García and Alfredo Marín. Covering Location Problems. In Gilbert Laporte, Stefan Nickel, and Francisco Saldanha da Gama, editors, *Location Science*, pages 99119. Springer International Publishing, Cham, 2019.

[3] Víctor Bucarey, Bernard Fortz, Natividad González-Blanco, Martine Labbé, and Juan A. Mesa. Benders decomposition for network design covering problems. *Computers & Operations Research*, 137:105417, January 2022.

[4] Nimrod Megiddo, Eitan Zemel, and S. Louis Hakimi. The Maximum Coverage Location Problem. *SIAM Journal on Algebraic Discrete Methods*, June 1983.

[5] Arie Tamir. An $O(pn^2)$ algorithm for the p-median and related problems on tree graphs. *Operations Research Letters*, 19(2):5964, August 1996.

[6] Liang Tian, Amir Bashan, Da-Ning Shi, and Yang-Yu Liu. Articulation points in complex networks. *Nature Communications*, 8(1):14223, January 2017.

Smoothed Analysis of the k-Swap Neighborhood for Makespan Scheduling*

Lars Rohwedder,
Department of Mathematics and Computer Science,
University of Southern Denmark, Denmark
e-mail: rohwedder@sdu.dk

Ashkan Safari, and Tjark Vredeveld
Department of Quantitative Economics, School of Business and Economics,
Maastricht University, The Netherlands
e-mail: [a.safari, t.vredeveld]@maastrichtuniversity.nl

Local search methods are some of the most widely used heuristics to tackle computationally difficult optimization problems [1, 4], due to their simplicity and strong empirical performance. A basic form of local search, known as *iterative improvement*, starts with an initial solution and iteratively searches for a better solution within the neighborhood of the current one. The iterative improvement stops when no further improvement can be achieved, signifying that the current solution has reached a *local optimum*.

This study addresses a fundamental scheduling problem, in which we are given a set of n jobs with positive processing times $p_j (j = 1, 2, ..., n)$ as well as a set of m identical parallel machines. Each of the jobs needs to be processed on one of the machines without preemption, and each machine can process at most one job at a time. The objective is to minimize the makespan $C_{\max} = \max_j C_j$, where C_j denotes the time at which job j is completed. This problem, denoted by $P||C_{max}$, is strongly NP-hard [3]. One way to approach NP-hard problems is by designing heuristic methods such as local search. A recent study [5], investigated the iterative improvement procedure within the *k-swap* neighborhood, offering a detailed analysis of its running time. A k-swap neighbor is obtained by selecting at most k jobs from two machines and interchanging the machine allocations of the selected jobs. In the same study, it has been shown that there exists an initial schedule as well as a sequence of 3-swaps such that the local optimum with respect to the k-swap neighborhood, for $k \geq 3$, is reached after $2^{\Omega(n)}$ iterations. Moreover, according to some computational results, the average number of iterations in converging to a local optimum is not excessively large. This raises the question of whether the provided lower bound reflects a pessimistic scenario which might be rare in practice.

To address such discrepancies between worst-case and practical performance, the notion of *smoothed analysis*, a hybrid of worst-case and average-case, was introduced by Spielman and Teng [6] to provide a more realistic measure of an

*This research is supported by NWO grant OCENW.KLEIN.176.

algorithm's performance. In smoothed analysis, an adversary selects an arbitrary instance, and then the numbers of this instance are (slightly) randomly perturbed. The smoothed performance is defined as the worst expected performance among all adversarial choices; the expected value is taken over the random perturbation of the adversarial instance. In a more general model of the smoothed analysis, introduced by Beier and Vöcking [2], an adversary specifies probability density functions upper-bounded by $\phi \geq 1$ for all the numbers that need to be perturbed.

Motivated by the exponential lower bound of $2^{\Omega(n)}$ [5], we perform smoothed analysis on the number of k-swaps for the problem $P||C_{max}$, and show that it is upper bounded by $O(m^2 \cdot n^{2k+2} \cdot \log m \cdot \phi)$. In our analysis, we perturb the processing times of the jobs. An adversary specifies density functions $f_1, \ldots, f_n$ with $f_i : [0, 1] \to [0, \phi]$, and then, an instance of the problem is obtained by drawing $p_1, \ldots, p_n$ independently according to the densities $f_1, \ldots, f_n$, respectively.

Let $p(S)$ denote the summation of the processing times of a set of jobs S, and let $\delta_{min} = \min |p(A) - p(B)|$, where the minimum is taken over all the sets of jobs A and B with $A \cap B = \varnothing$ and $1 \leq |A| + |B| \leq k$. To perform the analysis, we divide the set of machines in our schedule at the start of iteration t into two sets γ_l and γ_s. The set γ_l includes machines with a load greater than $C_{max} - \delta_{min}$, while γ_s consists of machines with a load less than or equal to this value. We analyze two types of k-swap, categorized according to whether the load of the lowest-load machine i' involved in the k-swap has increased significantly or not. More precisely, in an improving k-swap of *type-1*, machine i' moves to γ_l while in an improving k-swap of *type-2*, machine i' stays in γ_s after the k-swap. The number of consecutive iterations of type-1 is upper-bounded by $O(m \cdot n^k)$, whereas for type-2, a global upper-bound of $O(\frac{m \cdot n}{\delta_{\min}})$ has been provided. In the worst case, δ_{min} can be (close to) 0, yielding a large bound. However, in the smoothed setting, this is highly unlikely, and we derive a bound of $O(m^2 \cdot n^{2k+2} \cdot \log m \cdot \phi)$ for the smoothed number of iterations.

References

[1] E.H.L. Aarts and J.K. Lenstra. *Local search in combinatorial optimization.* Princeton University Press, 2003.

[2] R. Beier and B. Vöcking. Random knapsack in expected polynomial time. In *Proceedings of the thirty-fifth annual ACM symposium on theory of computing (STOC)*, pages 232–241, 2003.

[3] D.S. Johnson and M.R. Garey. *Computers and intractability: A guide to the theory of NP-completeness.* WH Freeman, 1979.

[4] W. Michiels, E.H.L. Aarts, and J. Korst. *Theoretical aspects of local search*, volume 13. Springer, 2007.

[5] L. Rohwedder, A. Safari, and T. Vredeveld. A k-swap local search for makespan scheduling. *arXiv preprint arXiv:2401.05956*, 2024.

[6] D.A. Spielman and S.H. Teng. Smoothed analysis of algorithms: Why the simplex algorithm usually takes polynomial time. *Journal of the ACM*, 51(3):385–463, 2004.

On the complexity of finding central configurations of the graph-generalized $(n^2 - 1)$-puzzle

Martijn van Ee
Netherlands Defence Academy, Faculty of Military Sciences
e-mail: M.v.Ee.01@mindef.nl

1 Introduction

In the $(n^2 - 1)$-puzzle [3], the generalization of the famous 15-puzzle [2], we are given a board of size $n \times n$, and tiles numbered $1, \ldots, n^2 - 1$. In a configuration, each tile occupies one field of the board, which leaves one field uncovered. Two configurations are adjacent if we can obtain one from the other by making a move, i.e., by sliding a tile onto the empty field from an orthogonally adjacent field. In the $(n^2 - 1)$-puzzle, the goal is to find a sequence of moves which will transform an initial configuration into a final configuration. We study whether we can find a configuration that is close to a given set of configurations, called scenarios, i.e., a configuration that can be transformed into any of the scenarios using a small number of moves. More specifically, we want to find the median configuration, minimizing the total length of the sequences to the scenarios, and the center configuration, minimizing the maximum length of a sequence of moves from this configuration to any of the scenarios.

These problems are inspired by operational processes on amphibious warfare ships. At the start of the trip to the location of the mission, vehicles are arranged on the loading dock based on their priority for the mission, e.g., the vehicles that need to go first are at the front. However, during the trip the commanding officer might decide to change strategy, for example when new information has become available. One of the consequences might be that the vehicles in the loading dock need to be reconfigured. Since the reconfiguration of the vehicles can be time-consuming, one should think about the initial configuration of the vehicles. Hence, given a set of strategies, we want to compute central configurations of the vehicles.

We consider finding central configurations for the graph-generalized version of the $(n^2 - 1)$-puzzle. We also allow the possibility for some of the tokens to be indistinguishable, and we consider instances with more than one empty space. Formally, our problems are defined as follows. We are given a graph $G = (V, E)$. Furthermore, we are given t_j tokens of type j, for $j = 1, \ldots, \tau$. Here, tokens of the same type are indistinguishable. Finally, we are given a set of K scenarios, $S_1, \ldots, S_K$, where each scenario is a feasible configuration, i.e., each token is as-

signed to a vertex in the graph, and only one token can be placed on a vertex. Two configurations are adjacent if we can obtain one from the other by making a move, i.e., by sliding a token along an edge to an unoccupied vertex. We define $d(X, Y)$ as the length of the shortest sequence of moves that transforms configuration X into configuration Y. In Median Configuration, the goal is to find a configuration X that minimizes $\sum_{k=1}^{K} d(X, S_k)$. In Center Configuration, the goal is to find a configuration X that minimizes $\max_{k=1\ldots,K} d(X, S_k)$.

1.1 Our contribution

We start by studying the most general case. Observe that computing the objective value of either problem requires computing the shortest transformation between two configurations, which is NP-hard [3]. This does not mean that finding the center and median is hard as well. In fact, we show that Median Configuration is polynomially time solvable when $K = 2$. However, we also show that Median Configuration is NP-hard for $K = 3$, and that Center Configuration is NP-hard for $K = 2$. We actually show that if we have polynomial time algorithms for these cases, we can find the shortest transformation between two configurations. This raises the question what happens if we would be able to compute shortest transformations. Therefore, in all other cases we study it is possible to compute the shortest transformation between configurations.

First, we study a case where we do not use the length of the shortest transformation in our objective function, but we use a natural approximation that ignores the dynamics of the puzzle. Given two configurations, for each token, we compute the shortest path in G between the vertices that this token occupies in the two configurations, respectively. Then, we sum these lengths over all tokens. We show that finding the median configuration with approximated distances can be solved in polynomial time. It turns out that finding the center configuration with approximated distances is NP-hard for arbitrary K. For fixed K, the problem reduces to Exact Matching [1], for which the complexity is open.

Second, we studied the original problem with identical tokens. We show that Center Configuration with identical tokens can be solved in polynomial time when $K = 2$, but that both problems with identical tokens are NP-hard for $K = 3$. Finally, we considered the problems on complete graphs. We show that computing the shortest transformation can be done in polynomial time. We also prove NP-hardness for Center Configuration on complete graphs.

References

[1] Vladimir G. Deineko and Gerhard J. Woeginger. On the robust assignment problem under a fixed number of cost scenarios. *Operations Research Letters*, 34(2):175–179, 2006.

[2] Wm. Woolsey Johnson and William E. Story. Notes on "15" puzzle. *American Journal of Mathematics*, 2(4):397–404, 1879.

[3] Daniel Ratner and Manfred Warmuth. The $(n^2 - 1)$-puzzle and related relocation problems. *Journal of Symbolic Computation*, 10(2):111–137, 1990.

Complexity of fixed order routing

Steven Miltenburg, Tim Oosterwijk, René Sitters
VU Amsterdam, Department of Operations Analytics
e-mail: s.j.g.miltenburg@vu.nl; t.oosterwijk@vu.nl; r.a.sitters@vu.nl

Routing problems are present in modern society in a wide range of applications. A classic variant is the vehicle routing problem (VRP) in which a fleet of servers (or vehicles), initially situated at a depot, needs to transport goods from the depot to a set of requests (or customers). For practical reasons, the order in which these requests need to be served is sometimes not very flexible and reordering may be difficult or expensive. We study this extreme case, where reordering is not allowed, in a parallel setting.

We consider an (offline) *fixed order routing* (FOR) perspective, in which the requests are given a priori in a specified order. After allocating requests to servers, the order in which the assigned requests are served needs to match the fixed order for every server.

We consider the case where the number of subtours is bounded by some given number k, which we call the k-FOR problem. If on the other hand the number of requests in each subtour is bounded by some given number c, the capacity, we call this the *c-capacitated* FOR problem.

We consider these problem settings with three different objective functions. We either minimize the distance traveled by the servers, minimize the sum of the completion times of all the jobs, or minimize the maximum of the completion time of the jobs.

Such a fixed order perspective has interesting connections to many different optimization problems. One of them is the scheduling variant with total completion time objective, which was introduced by Bosman et al. [2]. k-FOR with total completion time objective is a generalization of this scheduling problem: Using a star graph as our metric, each of the servers can be seen as a machine in the scheduling problem, and nodes can be placed on distances to match the completion time in the scheduling problem.

While the classic scheduling problem with total completion time objective is NP-hard though easy to approximate, the fixed order version turns out to be much harder. We see a similar behavior in this paper: the vehicle routing problem without fixed order conditions on the line is trivial, while we prove this to be NP-hard on the line for fixed order routing. We complement this with both positive and matching complexity results on general metrics.

We prove strong NP-hardness for k-FOR on the line for the total completion time and maximum completion time objective and where k is part of the input, by a reduction from the 3-dimensional matching problem. For $k = 2$ both version are weakly NP-hard. Note that when there is no order restriction (as in the standard VRP) then all completion times are minimized simultaneously by letting 2 servers move in opposite direction in a straight line.

Theorem 1 *k-FOR on the line is strongly NP-hard for $k \geq 3$ and weakly NP-hard for $k = 2$, for both of the following objectives:*

(i) Minimizing the sum of weighted completion times.

(ii) Minimizing the maximum completion time.

Thus seemingly easy problems, such as 2-FOR on a line for weighted and maximum completion times, are already hard. However, we still have some positive results. On general metrics, we show that both these problems can be solved in pseudopolynomial time for any constant k by a dynamic program.

Theorem 2 *Let D be an upper bound on the maximum distance that any server will travel. k-FOR can be solved in $O(kn^k D^k)$ time for the following objectives:*

(i) Minimizing the sum of weighted completion times.

(ii) Minimizing the maximum completion time.

Next, we give a $2 - \frac{1}{c}$ approximation algorithm for c-capacitated FOR, where every server has a capacity of c, for any metric space. We use the fact that for $c = \infty$ the problem reduces to a flow problem which is solvable in polynomial time. Then we cut up the solution to the flow problem into subtours of length c, and choose the best of any such cuts. The analysis of our algorithm is tight.

Theorem 3 *c-capacitated FOR for the total length objective can be approximated by a factor $2 - \frac{1}{c}$.*

Note that c-capacitated FOR for $c = \infty$ reduces to a flow problem, and for $c = 2$ to a matching problem, both of which can be solved optimally in polynomial time. However, for $c \geq 3$, we show that the problem is APX-hard, with a reduction from maximum bounded 3-dimensional matching.

Theorem 4 *c-capacitated FOR for the total length objective is APX-hard for $c \geq 3$ for general metric spaces.*

However, the fact that it can be approximated by this factor $2 - \frac{1}{c}$ separates c-capacitated FOR from c-VRP, where the best known approximation guarantee is $2.5 - \epsilon$ (for ϵ very small) from Blauth et al.[1].

References

[1] Jannis Blauth, Vera Traub, and Jens Vygen. Improving the approximation ratio for capacitated vehicle routing. *Mathematical Programming*, 197(2):451-497, 2023.

[2] Thomas Bosman, Daria Frascaria, Neil Olver, René Sitters, and Leen Stougie. Fixed-order scheduling on parallel machines. In *Proceedings of the 20th International Symposium on Integer Programming and Combinatorial Optimization, IPCO 2019*, volume 11480 of *Lecture Notes in Computer Science*, pages 88-100. Springer, 2019.

Robustness Analysis of Counterfactual Explanations from Generative Models: A Survey

Kseniya Sahatova
KU Leuven, LIRIS Research group
SKEMA Business School
kseniya.sahatova@kuleuven.be

Johannes De Smedt
KU Leuven, LIRIS Research group
johannes.desmedt@kuleuven.be

Xuefei Lu
SKEMA Business School,
SKEMA Centre for Analytics and Management Science
xuefei.lu@skema.edu

Counterfactual explanations are a key tool in explainable AI, offering insights into complex machine learning models by addressing *"What if?"* scenarios. While conventional methods for generating counterfactual explanations (CFEs) rely on computationally expensive optimization techniques, generative models such as GANs and VAEs have enabled faster CFE generation. However, their opaque nature raises concerns about trustworthiness, especially in high-stakes domains like healthcare and finance, where transparency and accountability are crucial.

Most research on CFEs has focused on binary classification tasks and tabular data, emphasizing desirable properties such as proximity, sparsity, robustness, feasibility, and actionability. Among these, robustness, ensuring that CFEs are stable and consistent under input or model perturbations, remains underexplored, especially for high-dimensional structured data like images.

In this study, we benchmark existing methods for generating CFEs using generative models and connecting them with a range of established metrics to assess robustness. Our work highlights the challenges of evaluating generative models, particularly in the absence of a community consensus on assessing their quality. Furthermore, we address the lack of empirical studies on how the complexity of the classification task impacts the robustness of CFE generation methods. By systematically applying and analyzing existing techniques and metrics in both binary and multiclass image classification settings, this study provides insights into the reliability and robustness of generative model-based approaches for counterfactual explanations.

The experimental setup for robustness evaluation includes the counterfactual generation methods listed in Table 1, and the metrics outlined in Table 2. To estimate the influence of classification task complexity, the MNIST benchmark dataset is used. Furthermore, the PneumoniaMNIST dataset is used to assess the impact of various perturbations on the robustness of CFEs, accounting for the unique challenges posed by medical imaging data. The experiments are conducted in two settings: (1) misclassification correction, where the target class corresponds to the label originally intended for the input but misclassified by the model; and (2) reclassification, applicable in multiclass scenarios, where the target class is

selected from existing classes. The types of perturbation used to evaluate the robustness of CFEs are detailed in Table 2, with each category addressing distinct sources of uncertainty that can affect the quality of the explanations.

Paper	Year	Dataset	Gen Model	Classification task
REVISE [2]	2019	CelebA	VAE	binary
CounteRGAN [3]	2022	MNIST	GAN	many-to-one mapping
CountGen [1]	2023	ImageNet, MNIST, CelebA, Edges2Shoes, ShapeNet, Style	GAN	multiclass and binary
CX-GAN [6]	2023	Tuberculosis chest X-ray	GAN	binary
COIN [5]	2024	CT scans	GAN	binary
DISCOVER [4]	2024	Historical images of blastocyst-stage embryos	Adversarial Autoencoder	binary

Table 1: Overview of the methods for generating counterfactual explanations

Category	Property	Type of perturbations	Metric
Model Changes	If a model M changes insignificantly to M', the prediction on a CFE x' remains consistent: $M(x') = M'(x')$.	An elimination of a few data points; A change in hyperparameters; Incremental change in the random seed for initialization of the models' weights.	a) Validity after retraining b) Stability
Input Changes	If inputs x_1 and x_2 are similar, their corresponding CFEs x'_1 and x'_2 should also be similar.	Insignificant perturbations of the input changes	a) Local instability b) Cost reduction
Model Multiplicity	A CFE x' for an aggregated model prediction remains valid for a subset of models $M' \subseteq M$.	Arbitrary models trained on different portions of the same training dataset	a) Validity under alternative models b) Counterfactual validity

Table 2: Metrics for evaluating the robustness of counterfactual explanations

References

[1] Javier Del Ser et al. "On generating trustworthy counterfactual explanations". In: *Information Sciences* 655 (Jan. 2024), p. 119898.

[2] Shalmali Joshi et al. *(REVISE) Towards Realistic Individual Recourse and Actionable Explanations in Black-Box Decision Making Systems.* arXiv:1907.09615. July 2019.

[3] Daniel Nemirovsky et al. "CounteRGAN: Generating counterfactuals for real-time recourse and interpretability using residual GANs". en. In: *Proceedings of the Thirty-Eighth Conference on Uncertainty in Artificial Intelligence*. ISSN: 2640-3498. PMLR, Aug. 2022, pp. 1488–1497.

[4] Oded Rotem et al. "Visual interpretability of image-based classification models by generative latent space disentanglement applied to in vitro fertilization". en. In: *Nat Commun* 15.1 (Aug. 2024). Publisher: Nature Publishing Group, p. 7390.

[5] Dmytro Shvetsov et al. *COIN: Counterfactual inpainting for weakly supervised semantic segmentation for medical images.* en. arXiv:2404.12832 [cs]. Apr. 2024.

[6] Tehseen Zia, Zeeshan Nisar, and Shakeeb Murtaza. *Counterfactual Explanation and Instance-Generation using Cycle-Consistent Generative Adversarial Networks.* arXiv:2301.08939 [cs]. Jan. 2023.

Evaluating the stability of model explanations in instance-dependent cost-sensitive credit scoring

Matteo Ballegeer, Matthias Bogaert, Dries F. Benoit

Ghent University, Department of Marketing, Innovation and Organisation
FlandersMake@UGent–Corelab CVAMO, Belgium
{matteo.ballegeer, matthias.bogaert, dries.benoit}@ugent.be

In credit scoring, the use of machine learning techniques to support or automate decision-making has become increasingly prevalent. While many advanced models have been proposed to improve discriminatory power, they often overlook the financial consequences of misclassifying applicants, particularly when the costs of such errors are asymmetrical and vary by loan. To address this, several instance-dependent cost-sensitive (IDCS) classifiers have been proposed, incorporating individual loan costs directly into the model's loss function (Bahnsen et al., 2014; Höppner et al., 2020; Vanderschueren et al., 2022). However, these models are often "black-box" algorithms, making their predictions hard to interpret, raising concerns amid increasing regulatory scrutiny, such as the GDPR and the proposed EU AI Act.

In response, eXplainable AI (XAI) techniques, such as SHAP and LIME, have emerged to provide post-hoc explanations of opaque models' decision-making processes while maintaining predictive performance. For a model to be truly interpretable, however, its explanations must not only clarify its decisions but also exhibit robustness, ensuring that similar inputs yield consistent explanations (Alvarez-Melis and Jaakkola, 2018). Unstable explanations may undermine the utility of these models, exposing financial institutions to legal and reputational risks. While existing studies in credit scoring have explored the stability of SHAP and LIME explanations (Visani et al., 2022), as well as the impact of class imbalance on these explanations (Chen et al., 2024), the behaviour of these methods when applied to IDCS classifiers remains unexplored. This gap is critical because XAI techniques such as SHAP and LIME generate explanations based on the raw predictions of a model. Consequently, alterations in the learning process—such as the integration of instance-dependent costs into the loss function—may profoundly impact both the nature and stability of these explanations.

Our research proposes a two-faceted methodology: first, it evaluates prevalent IDCS classifiers in terms of their cost-efficiency and discriminatory power, and second, it assesses the stability of SHAP and LIME explanations when applied to these classifiers. We conduct a comparative analysis of traditional machine learning models and their IDCS counterparts using four open-source credit scoring datasets. To evaluate model performance, we use a combination of traditional and cost-sensitive metrics, introducing relative Average Expected Cost (relAEC) as a novel, dimensionless adaptation of the AEC metric designed for cross-dataset comparisons. To assess explanation stability, we measure the Coefficient of Vari-

ation (CV) and Sequential Rank Agreement (SRA) (Ekstrøm et al., 2018) of feature importances for both traditional and IDCS models. We also examine the additional impact of class imbalance through a controlled resampling procedure, following Chen et al. (2024).

The results show that while IDCS classifiers improve cost-efficiency, they produce significantly less stable explanations compared to traditional models, especially as class imbalance increases. This highlights a critical trade-off between cost optimization and interpretability in credit scoring. Given the growing regulatory emphasis on explainability, this research underscores the urgent need to address the stability issues in IDCS classifiers to ensure that their cost advantages are not overruled by unreliable or untrustworthy explanations.

References

Alvarez-Melis, D. and Jaakkola, T. S. (2018). On the robustness of interpretability methods.

Bahnsen, A. C., Aouada, D., and Ottersten, B. (2014). Example-dependent cost-sensitive logistic regression for credit scoring. In *2014 13th International conference on machine learning and applications*, pages 263–269. IEEE.

Chen, Y., Calabrese, R., and Martin-Barragan, B. (2024). Interpretable machine learning for imbalanced credit scoring datasets. *European Journal of Operational Research*, 312(1):357–372.

Ekstrøm, C. T., Gerds, T. A., and Jensen, A. K. (2018). Sequential rank agreement methods for comparison of ranked lists. *Biostatistics*, 20(4):582–598.

Höppner, S., Stripling, E., Baesens, B., and Verdonck, T. (2020). Profit driven decision trees for churn prediction. *European Journal of Operational Research*, 284(3):920–933.

Vanderschueren, T., Verdonck, T., Baesens, B., and Verbeke, W. (2022). Instance-dependent cost-sensitive learning: do we really need it? In *55th Hawaii International Conference on System Sciences (HICSS-55)*, pages 1–9.

Visani, G., Bagli, E., Chesani, F., Poluzzi, A., and Capuzzo, D. (2022). Statistical stability indices for LIME: Obtaining reliable explanations for machine learning models. *Journal of the Operational Research Society*, 73(1):91–101.

Enriching Process Discovery with Contextual Hierarchies

Zahra Ahmadi
KU Leuven, LIRIS research group
e-mail: zahra.ahmadi@kuleuven.be

Jochen De Weerdt
KU Leuven, LIRIS research group
e-mail: jochen.deweerdt@kuleuven.be

Estefanía Serral
KU Leuven, LIRIS research group
e-mail: estefania.serralasensio@kuleuven.be

Keywords: Process Mining, Contextualization, Activity Specialization

Today, many organizations use Process Mining (PM) as a data-driven technique to gain insights into the actual behavior of their process to enhance efficiency and operational effectiveness. This technique, traditionally used in business process management, is capable of analyzing process event logs to discover organizational processes and support decision-making in many different domains, such as manufacturing and healthcare to offer new opportunities for advanced process analysis. Process discovery, a core technique within PM, can produce detailed workflows of business processes from event logs, however, they often lack sufficient contextual depth, limiting the interpretability and actionability of the revealed activity process flows. The resulting workflows provide very limited interpretability on the context where a process is executed and how this context may affect its execution.

Context variables can provide essential background information about the process execution, enhancing the understanding of the process variants and enabling a more detailed analysis. Incorporating contexts in the model could improve the understanding of the actual process executions. However, adding contextual details introduces the risk of increasing model complexity, which can hinder the interpretability and usability of the model. Although previous studies have explored types of context in PM and some approaches use it, these methods often do not fully integrate context within process models resulting in models that fail to capture the full variability of process execution under different contextual influences. Furthermore, attempts to incorporate context frequently struggle with the resulting complexity, making the models less interpretable and harder to apply. As a result, there is a lack of practical methodologies for integrating and visualizing contextual information in discovered process models while maintaining interpretability.

This research addresses these challenges by introducing a context-driven activity hierarchy method for context-enriched process discovery and visualization. The proposed method systematically identifies, extracts, and integrates meaningful contextual variables into activity flows, creating layered process models that

balance clarity with analytical depth, enabling more informed decision-making in complex environments. Process activities are specialized using the corresponding variables forming specialization hierarchies, where higher-level nodes correspond to general activities and lower-level nodes represent more specialized/contextualized variants of those activities. Each node in the hierarchy provides a level of abstraction, allowing the process model to be simplistic at higher levels while revealing context-driven variations at lower levels.

For example, in a logistics company, the activity "Package Item" may vary depending on contextual variables such as "Shipping Priority", which could include values like *Standard Delivery* or *Express Delivery*. Using these contextual variables, the activity "Package Item" can be specialized into context-enriched variants. This enriched process model and corresponding visualization can enable the company to allocate resources more effectively for high-priority shipping options like *Express Delivery*.

This contribution advances the field of process mining by offering a structured methodology and visualization framework for integrating context into process models. By maintaining interpretability while enriching models with critical contextual information, the approach supports deeper analyses, actionable insights, and more effective decision-making, thereby supporting improved business process management in complex environments.

The dual quest for interpretability and performance in credit scoring via spline-rule ensembles

Simon De Lange
Ghent University, Department of Marketing, Innovation and Organisation
FlandersMake@UGent-corelab CVAMO, Belgium
e-mail: simondla.delange@ugent.be

Matthias Bogaert
Ghent University, Department of Marketing, Innovation and Organisation
FlandersMake@UGent-corelab CVAMO, Belgium
e-mail: matthias.bogaert@ugent.be

Koen W. De Bock
Audencia Business School, Department of Marketing
e-mail: kdebock@audencia.com

Dirk Van den Poel
Ghent University, Department of Marketing, Innovation and Organisation
FlandersMake@UGent-corelab CVAMO, Belgium
e-mail: dirkvandenpoel@ugent.be

Accurately estimating customers' creditworthiness is important for banks and financial institutions because of its effect on their income statements. Accepting loans that later default results in losses, while rejecting good loans results in missed profitable opportunities. Traditionally, these decisions were made using white box statistical models such as linear discriminant analysis or logistic regression. As computational power increased, more complex black box models, such as tree-based ensembles (e.g., random forest and XGBoost) became more popular. This creates a trade-off for financial institutions, as enhancing model performance often comes at the expense of interpretability. In the case of credit scoring, explainability is essential because financial regulators require creditors to provide a statement of specific reasons to denied applicants in the interest of consumer protection (Equal Credit Opportunity Act; Report on big data and advanced analytics, EBA). To stimulate both interpretability and predictive performance two approaches can be distinguished. First, predictive performance of an interpretable white box model, such as logistic regression, can be improved through feature engineering. Second, post hoc explainability methods such as Shapley values can be applied to more complex black box models [1]. The latter is, however, less desirable from a consumer protection perspective, as these techniques are not perfectly faithful to the original black box model [4]. This

study explores the first method by implementing spline-rule ensembles (SRE) as a new model category in credit scoring, which has already demonstrated its performance in areas such as business failure prediction [3] and customer churn modeling [2]. SRE combines the strengths in performance of tree-based ensembles such as random forest with interpretability close to a linear model such as logistic regression because of its additive character.

A traditional rule ensemble (RE) extracts rules from a tree-based algorithm, incorporates them as binary features alongside the original predictors into the feature space, and applies regularized logistic regression to this augmented feature set. The combination of rules and linear terms allows for modeling complex relationships such as thresholds and interactions. SRE extends RE by adding smooth functions as an additional base learner, which enables the ensemble to more directly accommodate non-linear effects, positioning it as a strong candidate classifier. In this manner, SRE occupies a middle ground between white box and black box models, enhancing performance while keeping the classifier interpretable. This two-fold goal is especially important in credit scoring where financial regulators are advocating inherently interpretable models.

Our research has multiple contributions. First, SRE and conventional RE are introduced as new model categories in the credit scoring literature. Second, multiple variants of RE and SRE with different rule generation algorithms are compared to find a high-performing, white box algorithm. Third, a benchmark experiment is conducted against well-established classifiers such as logistic regression and random forest. Finally, a case study exhibits the advantages of SRE in model- and instance-level interpretation. The results indicate that SRE serves as a high performing and interpretable alternative to these well-established algorithms. Compared to traditional white box models, SRE achieves superior performance due to its enhanced flexibility. Furthermore, in comparison with black-box models, SRE offers greater interpretability without a significant reduction in performance.

References

[1] Michael Bücker, Gero Szepannek, Alicja Gosiewska, and Przemyslaw Biecek. Transparency, auditability, and explainability of machine learning models in credit scoring. *Journal of the Operational Research Society*, 73(1):70–90, 2022.

[2] Koen De Bock and Arno De Caigny. Spline-rule ensemble classifiers with structured sparsity regularization for interpretable customer churn modeling. *Decision Support Systems*, 150:113523, 02 2021.

[3] Koen W. De Bock. The best of two worlds: Balancing model strength and comprehensibility in business failure prediction using spline-rule ensembles. *Expert Systems with Applications*, 90:23–39, 2017.

[4] Cynthia Rudin. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence*, 1(5):206–215, May 2019.

Comparative study on the effects of granting more decision autonomy to human operators in warehousing

T. De Lombaert, K. Braekers, K. Ramaekers
Hasselt University, Research Group Logistics
e-mail: thomas.delombaert@uhasselt.be,
kris.braekers@uhasselt.be,
katrien.ramaekers@uhasselt.be

R. De Koster
Erasmus University Rotterdam, Rotterdam School of Management
e-mail: rkoster@rsm.nl

Warehouses play an important role in a company's supply chain and contribute to its failure or success. There are many activities taking place within a warehouse, although it has been shown that order picking (OP) is by far the costliest. Order picking can be defined as the process of retrieving products from their storage locations in a warehouse to prepare them for shipping. As a result of the large share of OP in total warehousing costs, warehouse managers aim to achieve high efficiency levels in their OP system. They can resort to several planning problems to respond to this aspiration, for instance batching, routing, or job assignment [1]. However, in many of today's warehouses, a central planning system coordinates these planning problems, setting out directives for the human workers. This practice erodes the perceived autonomy of order pickers, although autonomy is one of the three basic psychological needs and has been found to affect worker well-being. This comparative study presents the development, testing, and post-hoc evaluation of an autonomy-increasing intervention in two warehouse environments. We show that the intervention is successful in both environments, which further favours granting autonomy to order pickers, as this has a beneficial effect on their well-being as well as organisational outcomes.

One of the planning problems which contributes to efficient warehouse operations is the job assignment planning problem. This planning problem coordinates the allocation of orders/batches/tasks (henceforward: orders) to pickers, as well as the sequence in which these should be handled. In many warehouses, those assignments are completely random, apart from possible considerations regarding due dates (or priorities). This situation leads to the underutilisation of pickers' skills, as picker-order matches may occur. For instance, a relatively short order picker may be assigned to top-shelf picks, or a picker who has a fear of heights is required to pick boxes at great height during his/her pick tour. Those mismatches can be avoided by repeatedly assigning pickers to very specific orders which are

in line with their respective skills. However, previous research has already highlighted the aversion of pickers from being repeatedly assigned to products with the same characteristics [2]. In addition, this form of repetition may physically overstrain pickers and could also lead to boredom as a result of monotony. This setting raises the question of whether it is possible to configure a system in which skills can be leveraged without compromising the physical and mental well-being of pickers.

The aim of this study is to configure such a system and assess its impact. In particular, we developed an order assignment mechanism (OAM) in which order pickers get the opportunity to choose their next order from a set of orders that is presented to them at the depot. This innovative OAM was tested in different warehouse environments and under different conditions. For example, we implemented the system for three weeks in a real-world warehouse and adopted a within-subjects study design. To evaluate the impact of this newly-developed working system, a holistic evaluation approach was used, as psychosocial-, physical-, and performance-related outcome measurements were collected. Psychosocial variables include measures such as perceived autonomy, task variety, job satisfaction, and motivation, while physical outcomes pertain to metrics like heart rate monitoring. Last, performance is conceptualised by number of order lines and number of pallets picked. In total, 18 order pickers voluntarily participated in our study. Another experiment was conducted in a laboratory setting to have more control over the warehouse conditions and to discern the true effect of our intervention. Both studies indicate a significant increase in psychosocial worker well-being, as well as positive, albeit non-significant, enhancement in physical well-being when working in a system with more decision autonomy for workers. Productivity measurements remained stable. Post-experimental semi-structured interviews provide a deeper understanding of the underlying mechanisms that brought about these positive outcomes.

This comparative study shows how high efficiency can coincide with increased worker involvement, contingent upon well-thought-through design choices. Using a holistic evaluation approach, we show the beneficial impacts of an autonomy-increasing intervention in two different settings. The insights derived from this study can be translated to other warehouse planning problems and give rise to several subproblems which can be encountered with an OR perspective.

References

[1] Van Gils, T., Ramaekers, K., Caris, A., & de Koster, R. B. (2018). Designing efficient order picking systems by combining planning problems: State-of-the-art classification and review. European Journal of Operational Research, 267(1), 1-15.

[2] De Lombaert, T., Braekers, K., De Koster, R., & Ramaekers, K. (2023). In pursuit of humanised order picking planning: methodological review, literature classification and input from practice. International Journal of Production Research, 61(10), 3300-3330.

Let Customers Scatter the Inventory: Multi-Objective Storage Location Assignment in Warehouses

Mostafa Bahadornia
Hasselt University, Research Group Logistics
e-mail: mostafa.bahadornia@uhasselt.be

Katrien Ramaekers
Hasselt University, Research Group Logistics
katrien.ramaekers@uhasselt.be

Kris Braekers
Hasselt University, Research Group Logistics
kris.braekers@uhasselt.be

Trijntje Cornelissens
University of Antwerp, ANT/OR Research Group
trijntje.cornelissens@uantwerpen.be

Ruben D'Haen
Hasselt University, Research Group Logistics
ruben.dhaen@uhasselt.be

The rapid growth of online retailing necessitates flexible warehouse management strategies to adapt to this evolving landscape [1]. One of the critical challenges in this area is to reduce the order-picking travel distance [2]. This travel distance is highly affected by the Storage Location Assignment (SLA) decision, which determines how products are allocated to locations in the warehouse [3].

Prevalent SLA strategies are: (a) Scattered storage assignment (SSA): increases the accessibility of each Stock Keeping Unit (SKU) by spreading its units through the storage locations [4]; (b) Correlated storage assignment (CSA): stores correlated SKUs close to each other [5]; (c) Turnover class-based (TCB): stores high-demand SKUs near the depot(s) [6].

This study proposes a mixed SLA approach that adopts each prevalent SLA strategy to some degree, tailored to the customer order pattern. To do so, three analytical measures are defined to assess the realization degree of each prevailing SLA strategy. In order to address the dynamic nature of business needs, a data-driven approach is introduced to weigh each criterion. Then, proceeding from a descriptive phase to a prescriptive one, a novel multi-objective mathematical model for SLA optimization is proposed. This model incorporates the mentioned weighted measures and the contextual constraints.

As the main goal of SLA optimization is to reduce the order-picking travel distance, the proposed model is tested in the collaborative human-robot configuration in a mixed-shelves layout. In this configuration, robots and pickers are paired and pick all items on a pick list; once all items are picked, the picker sends the robot to the depot and starts picking with another robot [7]. Finally, post hoc analyses are being performed to validate the reliability of the proposed approach.

References:

1. Lone, S., N. Harboul, and J. Weltevreden, 2021 European e-commerce report. 2021.
2. Boysen, N., R. De Koster, and F. Weidinger, Warehousing in the e-commerce era: A survey. European Journal of Operational Research, 2019. 277(2): p. 396-411.
3. Van Gils, T., et al., Increasing order picking efficiency by integrating storage, batching, zone picking, and routing policy decisions. International Journal of Production Economics, 2018. 197: p. 243-261.
4. Weidinger, F. A precious mess: on the scattered storage assignment problem. in Operations Research Proceedings 2016: Selected Papers of the Annual International Conference of the German Operations Research Society (GOR), Helmut Schmidt University Hamburg, Germany, August 30-September 2, 2016. 2018. Springer.
5. Xiao, J. and L. Zheng, A correlated storage location assignment problem in a single-block-multi-aisles warehouse considering BOM information. International Journal of Production Research, 2010. 48(5): p. 1321-1338.
6. Petersen, C.G. and R.W. Schmenner, An evaluation of routing and volumebased storage policies in an order picking operation. Decision Sciences, 1999. 30(2): p. 481-501.
7. Azadeh, K., et al., Zoning strategies for human-robot collaborative picking. Decision Sciences, 2023. 00, p. 1-21.

Multi Product Inventory Routing Problem with Pickup and Delivery

Lukas Seeholzer
KU Leuven, Institute for mobility - CIB
e-mail: lukas.seeholzer@kuleuven.be

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

The hospitality sector faces unique and challenging problems in logistics. Often, products have to be delivered under challenging circumstances. For example, some products need to be cooled during storage and transport, other products have strict time constraints due to expiration dates. Often, a variety of different products from different suppliers needs to be delivered to a set of clients on a regular basis.

Based on a case study in the hospitality industry, a new variant of the inventory routing problem (IRP) is defined, modelled and analysed. We propose a mathematical model, show its limitations due to the exponential growth and propose a large neighbourhood search to find good solutions in a reasonable time, even for large instances.

The typical IRP [2] combines the vehicle routing problem (VRP) [3] with the inventory management on the client's side. The goal of the IRP is to find the optimal routing of a vehicle fleet, such that the demands of all clients are fulfilled. An extension of the IRP allows the optimization over multiple time periods, which may lead to better optimal solutions due to the possibility of combining deliveries not only in the two-dimensional space but also in the time dimension [1].

These problems normally only consider one product to be delivered from one supplier. In reality, however, there are often multiple suppliers that deliver products to a set of clients, especially in the hospitality industry. Generally, there is no cooperation, neither between the suppliers nor between the customers, leading to a situation where each supplier operates its own vehicle fleet. Cooperation between the suppliers offers new optimization possibilities. Hence, we define the multi product IRP with pickup and delivery in the following way:

There is a set of customers and a set of suppliers. Each customer demands products from a subset of the suppliers. This demand may vary from day to day. The clients have the possibility to store goods from any supplier, however, storage occurs costs. There is a set of vehicle types, each of which has specific loading capacities and operation, opportunity and amortization costs. The goal is to choose the optimal vehicle fleet, operated centrally, and to optimize routing and delivery schedules such that all demands are satisfied and total costs are minimized.

Preliminary experiments with solving the mathematical model show a significant cost saving potential due to an increase in vehicle efficiency. However, it

is also clear that finding the optimal solution becomes practically infeasible very quickly. Our heuristic approach, currently being developed, aims to find good solutions in a reasonable time.

References

[1] Bertazzi, L., & Speranza, M. G. (2012). Inventory routing problems: an introduction. EURO Journal on Transportation and Logistics, 1, 307-326.
[2] Campbell, A., Clarke, L., Kleywegt, A., & Savelsbergh, M. (1998). The inventory routing problem. In Fleet management and logistics (pp. 95-113). Boston, MA: Springer US.
[3] Toth, P., & Vigo, D. (Eds.). (2002). The vehicle routing problem. Society for Industrial and Applied Mathematics.

The joint stochastic multi-period lot sizing and two-dimensional variable-sized cutting stock problem

Alexis Robbes
Université de Technologie de Troyes, LIST3N, 12 rue Marie Curie, 10000 Troyes, France
e-mail: alexis.robbese@utt.fr

Khadija Hadj Salem
INESC TEC, Porto, 4200-465, Portugal
e-mail: hadj.salem@inesctec.pt

1 Problem definition

This paper deals with a stochastic multi-period lot-sizing problem combined with a variant of the two-dimensional cutting stock problem (2D-CSP), first introduced by [2], in the context of home textiles. The key decisions are to determine the length of each type of fabric roll and to specify the cutting patterns for the items, while minimizing the total area required to cut the demanded pieces (the amount of fabric used), also known as waste. We call this variant the multi-period *two-dimensional Variable-Sized Cutting Stock Problem with stocks and under stochastic defects* (2D-VSCSP-S-SD). Several studies have explored uncertainty in cutting problems, which are summarized and classified in [1]. However, the specific nature of defects encountered in the home textile context remains unexplored.

More formally, we are given a set of rectangular items I, where each item i is characterized by its width w_i, height h_i, and demand d_i. These items must be cut in a two-stage guillotine manner within the fabric rolls J, where each fabric roll j has a predefined width W_j. The planning horizon consists of a set of periods T. A manual trimming step is allowed to separate items from waste parts. The lot-sizing decision involves determining the weaving length L_j^t for each fabric roll j in period t. The production cost of a fabric roll consists of a fixed cost c_p and a variable cost proportional to the area produced, c_{m^2}. In each period, the produced fabric roll is fully used to satisfy the demand for items. Stored items incur a holding cost of c_h per period, while shortages are allowed but penalized by a shortage cost c_s. The objective is to minimize the average total cost, which includes the total production cost, the total holding cost, and the total shortage cost.

However, during the weaving process, fabric rolls may exhibit unexpected defects extending across dozens or even hundreds of meters. These defects can arise from a variety of factors, including mechanical malfunctions in the loom, , irregularities in the thread supply, or external contaminants disrupting the weaving process. This study focuses on defects that affect a specific cutting pattern, consisting of small items arranged through two-stage guillotine cuts. As illustrated in Figure 1, these defects fall into two categories: (a) horizontally missing threads, and (b) vertically missing threads.

(a) Horizontally missing thread (b) Vertically missing thread

Figure 1: How to represent the defects appearance?

2 Modeling approaches

To model thread defects, we propose the following approaches:

- **Horizontally missing thread:** We define a function $\mathcal{H} : (p, y) \mapsto \{\overline{a_{ip}}\}_{(i \in I}$ which, for a given pattern p and a defect position y, returns the number of defective items. For each pattern on the affected fabric roll, this function calculates the number of defective items for each item type.
- **Vertically missing thread:** We define a function $\mathcal{V} : (p, x) \mapsto \{\overline{a_{ip}}\}_{(i \in I}$, which, for a given pattern p and defect position x, returns the number of defective items. This requires identifying the relative position of the cutting pattern with respect to the defect: whether it precedes the defect (before), follows it (after), or overlaps with it.

We propose a stochastic Mixed Integer Linear Programming (MILP) model based on scenarios, where the functions $\mathcal{H}$ and $\mathcal{V}$ are replaced by a set of binary variables and the number of defective items $\{\overline{a^k_{ip}}\}_{(i,p,k) \in I \times J \times K}$ where K represents the set of position classes. We assume that a set of predefined patterns is generated in advance. The model optimizes the expected total cost over all scenarios.

We implemented the proposed approaches in Python and solved them using the Gurobi Optimizer 11.0.3 solver (https://www.gurobi.com/). To evaluate their efficiency, we performed our experiments on realistic randomly generated instances. Numerical results will be presented and analyzed during the talk.

References

[1] Khadija Hadj Salem, Elsa Silva, and José Fernando Oliveira. Cutting and packing problems under uncertainty: literature review and classification framework. *International Transactions in Operational Research*, 30(6):3329 - 3360, 2023.

[2] Khadija Hadj Salem, Elsa Silva, José Fernando Oliveira, and Maria Antónia Carravilla. Mathematical models for the two-dimensional variable-sized cutting stock problem in the home textile industry. *European Journal of Operational Research*, 306(2):549 - 566,2023.

Designing a Fair Orienteering Contest Using Bilevel Optimization

David Van Bulck
Ghent University, Department of Business Informatics and Operations Management
Core lab CVAMO, FlandersMake@UGent, Belgium
e-mail: david.vanbulck@ugent.be

Joonas Pääkkönen
Dalarna University, Sweden Department of Data and Information Management

Benjamin Jacquet
ENSTA Paris, France

Dries Goossens
Ghent University, Department of Business Informatics and Operations Management
Core lab CVAMO, FlandersMake@UGent, Belgium

Keywords— Orienteering problem, Human behaviour in problem solving, Bilevel optimization with inexact followers

Orienteering is a group of cross-country running sports where participants use a map and compass to navigate through unfamiliar terrain to find locations marked on the map, known as control points. In this talk, we discuss a variant of orienteering called rogaining, where each control point is assigned a score, and competitive performance is measured by the cumulative score the competitor gathers by visiting as many control points as possible within a given time limit. In the operations research literature, this problem is known as the 'Orienteering Problem' and has attracted a significant amount of attention (see e.g., [2, 3]).

This talk, based on [1], studies the (sports) orienteering problem from an organizers' point of view. In particular, organizers of a rogaining event need to decide what control points to include and what scores to assign to them. A good score assignment has the following three properties:

1. participants not only need excellent physical (running) but also cognitive (navigational and planning) skills in order to obtain high scores,
2. final scores among the teams exhibit a high variance so that stronger teams can be separated from the weaker ones, and
3. the manpower needed for setting out the control points in the terrain does not exceed a given resource budget.

We coin this problem the score assignment problem, and solve it via a heuristic bilevel optimization approach where the upper level represents the organizers who assign scores to control points and the lower level represents the participants who react by solving the classic orienteering problem. Since humans are unlikely to come up with optimal courses and routes and since sports performance is highly unpredictable, we rely on simulation based heuristics for the lower level. Given the outcome of the lower

level, a provisional ranking can be constructed which allows the upper level to evaluate the quality of the score assignment in terms of the cognitive skill importance and the separation of the participants. As such, this paper provides an interesting application of soâĂŘcalled bilevel optimization problems with inexact followers. The result is a diverse set of score assignments that we offer to the organizers, visualizing the expected behaviour of the participants and highlighting the impact of the various rogaining skills. Our method's efficacy is shown in its application to the 2023 Rogaining World Championships.

References

[1] Van Bulck, Pääkkönen, Jacquet, Goossens (2024). Designing a sports orienteering contest: physical versus cognitive skills in rogaining. International Transactions in Operational Research, In Press, 2024.

[2] Vansteenwegen, Souffriau, Van Oudheusden (2011). The orienteering problem: A survey. European Journal of Operational Research, 209, 1-10.

[3] Gunawan, Lau, Vansteenwegen (2016). Orienteering Problem: A survey of recent variants, solution approaches and applications. European Journal of Operational Research, 2016, 255, 315-332.

Minimal and fair waiting times for single-day sports tournaments with multiple fields

Lisa Garcia Tercero[1,2] David Van Bulck[1,3] Fabien Nießen[2]
Dries Goossens[1,3]

[1]Ghent University, Faculty of Economics and Business Administration
e-mail: {lisa.garciatercero,david.vanbulck,dries.goossens}@ugent.be
[2]KU Leuven, Department of Computer Science
e-mail: fabien.niessen@kuleuven.be
[3]FlandersMake@UGent - core lab CVAMO

We consider a single-day sports tournament where all teams gather at a location featuring multiple fields. The tournament is organized as a single round-robin tournament, meaning all teams play each other exactly once. Each team needs a resting time of at least one time slot between every two games they play, and after this resting period they are considered fully rested. We will refer to all time slots beyond the cut-off point, but before a team's next game, as waiting time for that team. In the context of amateur tournaments, teams typically prefer to return home without experiencing any waiting times. Moreover, having to wait longer than other teams may lead to feelings of unfair treatment and complaints to the tournament's organizers. We therefore aim to create timetables that are both efficient, by minimizing the total waiting time, and fair, by minimizing the maximum waiting time experienced by any one team.

Previous studies [4, 2, 1] have focused on balancing resting times, as differences may impact the outcome of the games. Suksompong [4] provides optimal timetables for a location with one field, whereas Atan and Çavdaroğlu [1], Çavdaroğlu and Atan [2], and Tufaha et al. [5] propose algorithms for a location featuring multiple fields. None of the aforementioned studies consider waiting times. By contrast, Knust [3] investigates the cases where teams are fully rested either 0 or 1 time slots since their previous game for a setting with one field. The tournament spans multiple days and is divided into rounds such that each team plays twice per round. Knust [3] shows that within any round, maximum efficiency and fairness can be obtained simultaneously.

This paper's setting with multiple fields with the objective to optimize waiting times has not yet been studied. Moreover, some papers only consider the number of teams to be even or odd, since each of these cases requires a different approach. The standard setting for a round-robin tournament involves an even number of teams and simply adds a dummy team when the number of teams happens to be odd. However, since a game against a dummy team is not an actual game, this introduces additional waiting times. We therefore provide polynomial-time constructive algorithms for both even and odd numbers of teams.

For an odd number of teams n, we prove that the total waiting time is at least $2n - 4$ and the maximum waiting time is at least 2. Our constructive algorithm generates timetables with a total waiting time of $2n - 3$ and a maximum waiting time of 3, therefore having an absolute performance guarantee of 1 for both objectives. Moreover, we were able to verify with an IP model that this solution is optimal for 5 teams.

For an even number of teams n, we prove that the total waiting time is at least $n - 1$ when n is a multiple of 4 and at least $n - 3$ when n is not a multiple of 4. The results of the constructive algorithms are shown in Table 1.

		Total waiting time	Maximum waiting time
$n \equiv 0 \pmod 8$		n	2
$n \equiv 2 \pmod 8$	$(n \geq 18)$	$n + 2$	3
$n \equiv 4 \pmod 8$		$n + 4$	3
$n \equiv 6 \pmod 8$	$(n \geq 14)$	$n - 2$	2

Table 1: Total waiting time and maximum waiting time obtained by constructive algorithms for different cases of n.

Comparing with the lower bound, our algorithms have an absolute performance guarantee of at most 5. Moreover, as verified by our IP model, these results are optimal for $n \in \{4, 8, 14, 16\}$. Furthermore it is worth noting that an even number of teams is preferred over an odd number of teams, since the total waiting time is almost halved.

References

[1] T. Atan and B. Çavdaroğlu. Minimization of rest mismatches in round robin tournaments. *Computers & Operations Research*, 99:78–89, 2018.

[2] B. Çavdaroğlu and T. Atan. Determining matchdays in sports league schedules to minimize rest differences. *Operations Research Letters*, 48(3):209–216, 2020.

[3] S. Knust. Scheduling sports tournaments on a single court minimizing waiting times. *Operations Research Letters*, 36(4):471–476, 2008.

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

[5] T. Tufaha, B. Çavdaroğlu, and T. Atan. Round-robin scheduling with regard to rest diferences. *Top*, 31(2):269–301, 2023.

Fair Schedules for Single Round Robin Tournaments with Ranked Participants

Sten Wessel
Eindhoven University of Technology, Department of Mathematics and Computer Science
e-mail: s.wessel@tue.nl

Cor Hurkens
Eindhoven University of Technology, Department of Mathematics and Computer Science
e-mail: c.a.j.hurkens@tue.nl

Frits Spieksma
Eindhoven University of Technology, Department of Mathematics and Computer Science
e-mail: f.c.r.spieksma@tue.nl

We consider schedules for single round-robin (SRR) tournaments, where each of the $n = 2k$ players plays every other player exactly once, either at *home* or *away*, where $k \in \mathbb{N}$, $k \geq 2$. We assume that the players $1, \ldots, n$ are sorted based on a predetermined *ranking*, where player 1 is the strongest player and player n is the weakest player. In many cases such a ranking is explicitly known; indeed, for most sports, there is a ranking of the best players or teams (e.g., the Elo rating in chess, the ATP/WTA ranking in tennis, the FIFA World ranking in football, etc.). Finally, we assume the existence of an "asymmetry" in any particular game; by an asymmetry, we mean the existence of a (dis)advantage in a game for one player compared to the other player. For instance, in the game of chess, playing with white is considered advantageous compared to playing with black. In fact, more generally, when a match is played in either of the team's venues, there is the so-called "home advantage"; an effect that is well-studied in literature.

When considering a schedule, focus in literature has always been on balancing home and away games over the rounds, i.e., to balance the asymmetry over time. Indeed, ideally, a team plays alternatingly home and away in a so-called single-break schedule [2]. We argue that it is also of fundamental importance to balance the home advantage over the opponents *ranked by strength*. Indeed, we claim that balancing the home advantage evenly over the ranked opponents is needed to ensure that (i) no team is disadvantaged compared to other teams (achieving *fairness*), and (ii) no distortion of the final result is present (achieving *efficacy*).

In this work, we introduce the concept of *ranking fairness* of a schedule. The *ranking home-away pattern (ranking HAP)* of a team is a sequence of H and A symbols of length $n - 1$, specifying whether the team plays home or away against its j-th strongest opponent. In other words, the ranking HAP of a team is a home-away sequence, where the games are ordered by decreasing strength of its opponents. Our aim is to balance the home advantage over the ranked opponents. We consider the asymmetry to be perfectly balanced when home and

away games alternate in the ranking HAP, i.e., the ranking HAP of a team is either $HAHAH\ldots$ or $AHAHA\ldots$. When a schedule is such that the ranking HAP is perfectly alternating for every player, we call the schedule *ranking-fair*.

1 Measuring Ranking Fairness

When a schedule is not ranking-fair, it is relevant to have a measure that represents the amount of imbalance. We express the ranking fairness of a schedule by measuring how much the ranking HAP deviates from the HAP that is perfectly alternating. For a team t, we consider all sub-patterns of the ranking HAP, i.e., each consecutive sub-interval of the ranking HAP characterized by the leftmost and rightmost positions $[i, j]$, where $1 \leq i < j \leq n-1$. Let $H_{i,j}^t$ denote the number of home games of team t that occur in this sub-pattern. In a ranking-fair schedule, $H_{i,j}^t \in \left\{ \left\lfloor \frac{j-i+1}{2} \right\rfloor, \left\lceil \frac{j-i+1}{2} \right\rceil \right\}$, i.e., (approximately) half of the games within each sub-pattern are home games. Our fairness measure F_t for the ranking HAP of team t is defined as

$$F_t = \frac{\Delta_t - \frac{1}{8}(n-2)^2}{\frac{1}{24}n(n-1)(n-2)}\,, \tag{1}$$

with Δ_t denoting the sum of the deviations between the actual and ideal number of home games for every sub-pattern,

$$\Delta_t = \sum_{i=1}^{n-2} \sum_{j=i+1}^{n-1} \left| H_{i,j}^t - \frac{j-i+1}{2} \right|. \tag{2}$$

The normalization ensures that the fairness measure always assumes values between zero and one. The ranking fairness of a schedule equals $F = \frac{1}{n} \sum_{t \in T} F_t$.

2 Results

Our main result is the following theorem.

Theorem 1. *For $n = 4k$ there exists a single-break, ranking-fair schedule.*

For $n = 4k + 2$, we instead provide an integer programming model that generates a single-break, ranking-fair schedule, if it exists. For $n = 6$, $n = 10$, and $n = 14$ no ranking-fair schedule exist, but for $n = 18, 22, 26, \ldots, 98$, we find single-break, ranking-fair schedules.

In practice, a widely used and popular way of creating the schedule of SRR tournaments is using a particular HAP set, the so-called canonical pattern set (CPS) [1], for which we have the following result.

Theorem 2. *The CPS does not allow a ranking-fair schedule for $n \geq 10$.*

References

[1] de Werra, D. (1980), *Geography, games, and graphs*, Discrete Applied Mathematics **2**, 327-337.

[2] de Werra, D. (1981), *Scheduling in sports*, in: Studies on graphs and discrete programming, edited by P. Hansen (pp.381-395), Springer, Berlin.

Maximizing suspense in sports competitions by dynamic scheduling

Karel Devriesere
Ghent University, Ghent, Belgium
Department of Business Informatics and Operations Management
e-mail: karel.devriesere@ugent.be

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

Sports organizers face the challenge of making competitions as exciting as possible. Various theoretical and empirical studies suggest that fan interest is enhanced the later the winner is known [4]. Moreover, fans typically lose interest as soon as their favorite team is eliminated from contention for any of the available prizes [8]. On the other hand, if a team's prize is determined before the competition ends, it is more likely to reduce its effort [3]. This suggests that one way of measuring the suspense of a competition is how late teams clinch one of the possible prizes.

As far as we are aware, this consideration has not received much attention in the sports scheduling community. Indeed, scientific work about sports scheduling is mostly focused on operational aspects, such as minimizing travel times, rest differences, and consecutive number of home or away games [5]. One possible reason for this is the unpredictable nature of sports outcomes: it is hard to accurately predict what the standing will be after a set of matches has been played. Therefore, when restricted to the case in which the schedule is made only once at the beginning for the whole period, there is little information one can use to construct a "suspense-optimal" schedule. In addition, one of the most prevalent competition formats in Europe is the k-round robin format, in which each team faces each opponent k times [6]. When restricted to this format, sports schedulers can only change the sequence in which matches are played, rather than affecting the set of opponents of each team.

In this work, we deviate from the traditional literature and consider the construction of suspense-optimal schedules in an incomplete round robin with dynamic scheduling. In contrast to a conventional round robin, in an incomplete round robin teams only face a subset of their eligible opponents. This format is popular when there are way more participants than time slots in which games can be organized. Dynamic scheduling refers to the process of scheduling parts of the competition gradually, as the season progresses. This seems to be a promising

avenue for constructing suspense-optimal schedules since now, we can use information about the current standing to choose both the opponents of teams and the sequence of matches. As our measure for suspense, we consider the number of teams whose prize is fixed in round r, summed over all rounds in the competition.

Determining whether a team is sure to obtain prize p, and only prize p (assuming prizes do not overlap), where p is tied to positions k to l, amounts to solving whether this team is eliminated from positions $1, \ldots, k-1$ and whether this team is guaranteed to finish in at least position l. These are two separate problems, known as the elimination [1, 7, 9] and guaranteed points placement problem [2]. At present, these problems have only been studied from the perspective of a static schedule, i.e. when the set of future matches is known. In the case of a dynamic schedule, however, a team is only eliminated if there exists no way to schedule the remaining games, such that any possible assignment of outcomes to those games ensures that the team finishes in at least position $k-1$. A similar reasoning applies to determining whether a team is guaranteed at least position k.

In this presentation, we give complexity results for the elimination and guaranteed points placement problems in a dynamic setting. Moreover, we give complexity results for finding a suspense-optimal schedule. Finally, we present an approach based on integer programming to tackle this problem and apply it to the new Champions League format.

References

[1] Thorsten Bernholt, Alexander Gülich, Thomas Hofmeister, and Niels Schmitt. Football elimination is hard to decide under the 3-point-rule. In International Symposium on Mathematical Foundations of Computer Science, pages 410-418. Springer, 1999.

[2] Jan Christensen, Anders Nicolai Knudsen, and Kim S Larsen. Soccer is harder than football. International Journal of Foundations of Computer Science, 26(4):477-486, 2015.

[3] Pascal Courty and Jeffrey Cisyk. Sports injuries and game stakes: Concussions in the National Football League. Economic Inquiry, 62(1):430-448, 2024.

[4] Adam Cox. Spectator demand, uncertainty of results, and public interest: Evidence from the English Premier League. Journal of Sports Economics, 19(1):3-30, 2018.

[5] Karel Devriesere, László Csató, and Dries Goossens. Tournament design: A review from an operational research perspective. European Journal of Operational Research, In press, 2024.

[6] Dries R Goossens and Frits CR Spieksma. Soccer schedules in Europe: an overview. Journal of scheduling, 15:641-651, 2012.

[7] Walter Kern and Daniël Paulusma. The computational complexity of the elimination problem in generalized sports competitions. Discrete Optimization, 1(2):205-214, 2004.

[8] Jeremy M Losak and Shane A Halpin. Does every game matter? A new perspective on the league standing effect in Major League Baseball. Sports Economics Review, 5:100028, 2024

[9] Ildikó Schlotter and Katarína Cechlárová. A connection between sports and matroids: How many teams can we beat? Algorithmica, 80:258-278, 2018

Optimising Dark Corridors for Biodiversity Conservation through Public Lighting Management

Elodie Bebronne
HEC Liège, School of Management of the University of Liège,
HEC Liège Research: Business Analytics and Supply Chain Management
e-mail: Elodie.Bebronne@Uliege.be

Sabine Limbourg
HEC Liège, School of Management of the University of Liège,
HEC Liège Research: Business Analytics and Supply Chain Management
e-mail: Sabine.Limbourg@Uliege.be

Public lighting has become a fundamental element of urban infrastructure, ensuring safety, comfort, accessibility, and usability in public spaces such as streets, parks, and public squares at night (1). However, artificial light at night (ALAN) alters the natural light cycle, negatively impacting human health and biodiversity (2; 3). While energy-efficient, the advent of light-emitting diode technology exacerbates these biological responses due to its blue light emissions. To address this issue, the nocturnal environment can be viewed as a social-ecological-technical system (4). This lens could strategically support a balance between the ecological and social components of nightscape management. This research aims to reconcile these factors by minimising ALAN's environmental impact while maintaining critical social functions.

Among its many impacts, ALAN interferes with the movement and orientation of species (3), which can generate significant population declines. Also, ALAN interacts with other pressures, including habitat loss and fragmentation (3). Strategies such as green corridors are deployed to support the movement of diurnal species. Yet, ALAN may undermine them. Hence, tactical urban planning decisions that consider activating and deactivating street lighting are needed to create ecological linkages that respect nocturnal species.

Therefore, our research aims to answer the question: Which streetlights should be turned off to create a dark corridor connecting core biodiversity areas and forming a connected reserve?

Our research builds on Billionnet's binary linear programming model for designing connected and compact nature reserves (5). We adapt this model to support the creation of dark corridors, incorporating novel constraints to minimise artificial lighting and maintain areas of natural darkness. Our contributions include a connectivity constraint to ensure the functional interconnection of dark zones, the consideration of social acceptance of mitigation policies, and the introduction of buffer zones to mitigate disturbances from urban settlements, roads,

and lighting.

The adapted model is tested on randomly generated instances. Future work will apply the model to real-world data, refining its practical implementation to support biodiversity conservation in urban and peri-urban environments.

References

[1] Tamar Trop, Sharon Shoshany Tavory, and Boris A. Portnov. Factors affecting pedestrians' perceptions of safety, comfort, and pleasantness induced by public space lighting: A systematic literature review. *Environment and Behavior*, 2023, 55(1-2), 3–46. doi:10.1177/00139165231163550.

[2] Ron Chepesiuk. Missing the dark: health effects of light pollution. *Environmental Health Perspectives*, 2009, 117:A20–A27. doi.org/10.1289/ehp.117-a20.

[3] Kevin Gaston, and Alejandro Sánchez de Miguel. Environmental impacts of artificial light at night. *Annual Review of Environment and Resources*, 2022, 47:373–398. doi.org/10.1146/annurev-environ-112420-014438.

[4] Elodie Bebronne, Samedi Heng, and Sabine Limbourg. Towards sustainable nocturnal environment management: a social-ecological-technical system analysis in Wallonia (Belgium). *Discover Environment*, 2024, 2(1):98. doi.org/10.1007/s44274-024-00128-z.

[5] Alain Billionnet. Designing connected and compact nature reserves. *Environmental Modeling and Assessment*, 2016, 21:211–219. doi.org/10.1007/s10666-015-9465-3.

Applications of CHANAkYA for policy decisions in Leuven.

Gaurav Malik
KU Leuven, Department of Mechanical Engineering
gaurav.malik@kuleuven.be

Chris M. J. Tampère
KU Leuven, Department of Mechanical Engineering
chris.tampere@kuleuven.be

This abstract demonstrates the potential applications of CHANAkYA, a comprehensive game-theoretical framework for pricing and tradable credit charging in complex transportation systems. CHANAkYA enables the modelling of strategic interactions among diverse decision makers, including government authorities and private mobility service providers, while incorporating multi-modal travel options and heterogeneous traveller demographics. Road pricing instruments including entry-based tolls and per-kilometre tolls which can be link, or path based alongside equivalent tradable credit schemes can be modelled within this framework. Further, these instruments can be differentiated between different demographics of travellers. Additionally, the issue of non-uniqueness of user-equilibrium path flows is addressed via the introduction of players perspectives within the framework. This allows the decision maker to make a rational choice to be optimistic or pessimistic. We discuss a series of five case studies based on the city of Leuven to demonstrate the range of applications of CHANAkYA. Each of these case studies investigates a specific policy question. The main intention is to demonstrate how the comprehensiveness and versatility of CHANAkYA can be utilized to explore specific policy questions, offering crucial insights for transport planning.

We model a pseudo-real city based on the city of Leuven and use different components of our framework to investigate distinct policy questions. The modelled links and nodes can be seen in Figure 1 and Figure 2. We model the modes of car (electric and conventional) and bus, and the user-classes of electric and conventional car owners with each subdivided into younger and older adults.

Subsequently, we present the findings from a series of targeted analyses. The policy questions explored and the short conclusions are as follows:

1. **Case study 1: Q:** Should there be a cordon toll? If yes, should there be a differential toll for electric cars? **A:** There should definitely be a cordon toll as there is an opportunity to make gains up to 3000 Euros/hour in terms of overall society's surplus. The gains for introducing a different toll for electric tolls are relatively small i.e., about 60 Euros/hour so it may not justify the additional implementation and political convincing.

2. **Case study 2: Q:** Should young adults be charged different cordon toll

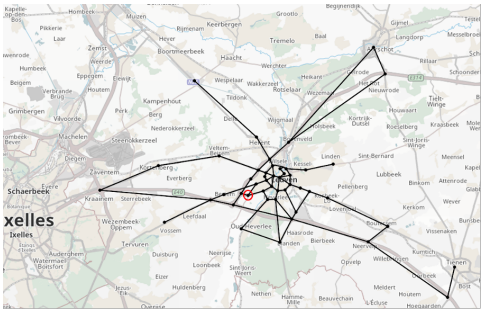


Figure 1: Physical Network based on Leuven. The red circle highlights the park and ride facility.

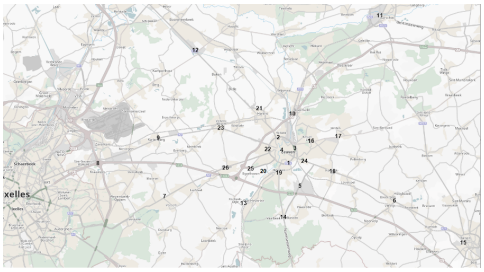


Figure 2: Centroids in the base model.

than the older adults? **A:** It can be a feasible option. Potential gain of 150 Euros/hour can be realized. The model sets a higher car cordon toll for younger adults as compared to older adults and vice-versa for the bus ticket price. This means that car usage by older adults (with a higher Value of Time) is prioritized while the young people are incentivized to use buses.

3. **Case study 3: Q:** Should a tradable credit scheme replace the cordon toll? **A:** The model shows that an equivalent tradable credit scheme leads to exactly the same overall surplus as well as flow distributions. The only difference is that in a tradable credit scheme the collected revenue is inherently distributed into the general public whereas in a toll scheme, it is integrated and added to the government's revenue.

4. **Case study 4: Q:** Should the bus service be provided by a private bus provider? **A:** There is a potential loss of 1000 Euros when the bus service is provided by a private entity solely concerned about its own profit. A private bus provider may only be advisable when the government acts as a Stackelberg leader. The loss of surplus in that case is very much limited and is less than 150 Euros.

5. **Case study 5: Q:** Do the perspectives of the government matter? **A:** Yes the perspectives of the government do matter. When the government is optimistic, it sets lower tolls and optimistically assumes the travellers to choose between the equally attractive options in a way that is good for the society. On the other hand, when the government assumes the opposite i.e., when it is pessimistic, it sets higher tolls to have stricter enforcement.

Towards a generic circular supply chain optimisation model: a systematic review and methodological outlook

Alix Langenaeker
University of Antwerp, Department of Engineering Management
Flanders Make@UAntwerp
e-mail: alix.langenaeker@uantwerpen.be

Philippe Nimmegeers
University of Antwerp, Department of Engineering Management
Flanders Make@UAntwerp, NANOlight Centre of Excellence
e-mail: philippe.nimmegeers@uantwerpen.be

Christof Defryn
University of Antwerp, Department of Engineering Management
e-mail: christof.defryn@uantwerpen.be

The growing demand for food, energy, water, and raw materials, coupled with population growth, climate change, and resource scarcity, highlights the need for smarter, sustainable systems. Circular supply chains play a crucial role in addressing these challenges by enhancing efficiency, reducing costs, and ensuring environmental resilience. Circularity aims to maintain the value of resources at their highest functionality, minimizing waste through strategies outlined in, e.g., the 9R framework (Refuse, Rethink, Reduce, Reuse, Repair, Refurbish, Remanufacture, Repurpose, Recycle, Recover) (Grow Circular, n.d.; Potocnik, 2013). However, the optimisation of sustainable circular supply chains remains a challenge due to the lack of a generic quantification framework to evaluate and enhance circularity and sustainability at the same time.

This study addresses the need for economically viable, environmentally friendly, and socially accepted circular supply chains. By conducting a systematic literature review with defined inclusion and exclusion criteria, we focus on analytical models and operations research within the domain of circular supply chains. Altogether, 118 papers were used for analysis, with 68 featuring analytical models relevant to the study. The meta-analysis of selected studies highlights key research streams, identifies methodological gaps, and proposes a generic optimisation framework applicable across all 9R strategies.

The literature reveals three research streams: (i) conceptual frameworks for transitioning to sustainable circular supply chains, (ii) a shift from economic perspectives to three-pillar (environmental, economic and social) sustainability, (iii) and generic frameworks addressing diverse supply chain challenges. Furthermore,

we find that although equally integrating economic, environmental, and social objectives in decision-making would be ideal, the environmental and social factors are often incompletely addressed — if at all. Also, the reliance on a simplistic set of variables and constraints, as well as oversimplified assumptions, such as substitutability of reprocessed products on the primary market, hinder real-world applicability. Additionally, frameworks often treat R strategies as fixed — omitting case-specific factors — and fail to address a variety of challenges simultaneously (Gao et al., 2021). Another limitation impacting the utility and adoption of the developed frameworks, can be found in the exclusion of resource- and geographic-based barriers.

In addition to the variety in optimisation variables, constraints and objectives, the modelling approach also varies greatly (Burinskienė, 2019). Only 19.6% of studies incorporate decentralised decision-making and stakeholder interactions, often using Stackelberg game theory, of which the real-world dynamics reflection might be questioned. Alternative approaches, like Nash-based coordination, are less prominently addressed.

These multi-stakeholder and multi-objective aspects, combined with the large set of possible model structures (e.g., accommodating the different R strategies) and constraint functions to describe the interactions, complicate the modelling and optimisation of sustainable circular supply chains.

Hence, the quest for the most effective modelling approach, balancing accuracy, complexity, scalability, and applicability, persists.

This systematic review of the literature underscores the critical gaps found in the existing frameworks designed, with the aim of guiding companies in transitioning from linear to sustainable circular supply chains. By bridging these methodological gaps, related to the generic and comprehensive character of the framework, future research has the potential to better equip organizations in this multifaceted transition towards sustainable circularity. Furthermore, a real-world validation through empirical data and case studies is essential to evaluate and enhance the practical applicability of these models.

References

Burinskienė, A. (2019). Review of methods for circular supply chain analysis: the application for optimisation studies. *DOKBAT 2019 - 15th International Bata Conference for Ph.D. Students and Young Researchers*, 169-180.

Gao, Y., Lou, S., Zheng, H., & Tan, J. (2023). A data-driven method of selective disassembly planning at end-of-life under uncertainty. *Journal of Intelligent Manufacturing*, 1-21.

Grow Circular. (n.d.). 9R Framework. Retrieved October 2, 2024, from https://grow-circular.eu/knowledge-base/9r-framework/

Potocnik, J. (2013). Towards the circular economy-economic and business rationale for an accelerated transition. *Ellen MacArthur Foundation.*

Optimizing Reverse Logistics for Waste Materials: A Multi-Stage Processing and Transportation Model

Clémence Renkin & Sabine Limbourg
HEC Liège Management School of the University of Liege,
Department of Operations Research QuantOM
e-mail: clemence.renkin@uliege.be
e-mail: sabine.limbourg@uliege.be

Companies currently face substantial challenges in adopting recycling technologies and establishing facilities due to their high costs ([2]). Nevertheless, European regulations targeting the plastics strategy, which seek to transform the design, production, usage, and recycling of plastic products within the EU, emphasize the critical need to improve efforts in the plastic recycling sector ([1]). Advancing this sector will require the development of efficient supply chains optimized to manage reverse waste flows, focusing on minimizing both costs and environmental impacts. In fact, transport costs, sorting and recycling logistics can account for a substantial share of the economic and environmental cost of recovered material. The study focuses on polyvinyl chloride (PVC), extensively used in construction and building for products like window frames, pipes, cables, and flooring. However, challenges persist, as conventional mechanical recycling cannot process all PVC waste. Advanced recycling offers a promising solution by purifying PVC and removing harmful additives.

We formulate a bi-objective, multi-stage material processing and transportation model to support the reverse logistics of PVC materials derived from the construction, renovation and demolition sectors. The objectives are twofold: minimizing the costs related to transport, processes, and inventories, as well as the negative environmental impacts by incorporating Life Cycle Assessment (LCA) throughout the whole reverse supply chain. This case-specific model has several unique features. First, the production setup involves two plants, each with distinct operations. The model includes two types of inventory: one before processing and one post-processing at each plant. With regard to distribution, only a single type of transport mode is considered in this research phase. Transportation costs account for travel time, loading, and unloading, based on estimates from the transportation partner. Batch sizes are also customized for this project, allowing containers of $10m^3$, $20m^3$, or $40m^3$ depending on the quantity collected at each source.

A critical challenge in reverse logistics is managing uncertainties in supply site locations, yield capacities, and operational durations. To address this, the current

deterministic model establishes a robust baseline for project-specific conditions. Future enhancements will introduce stochastic elements and expand scalability to an industrial level, broadening applicability to diverse environments and industries.

This initiative is expected to produce several significant results. The project aims to develop bi-objective optimization models for the collection, sorting and recovery of PVC waste specifically from the construction, renovation and demolition sectors. The resulting findings will support industry managers in making strategic and tactical decisions by identifying optimal production lot sizes and determining the quantities of materials to transport between multiple processing sites over a defined planning horizon. In addition, waste traceability will be improved, allowing data collection to monitor the progress and effectiveness of recycling activities. This approach aims to make recycling competitive with the low costs of incineration and disposal techniques.

Acknowledgment

The authors gratefully acknowledge the support provided by the Walloon Region to fund this research through the CIRC-PVC project (grant number 8909), as well as the Walloon clusters Greenwin and Logistics in Wallonia.

References

[1] European Commission. A european strategy for plastics, 2023. Accessed: 2024-12-02.

[2] Eleftherios Iakovou, Efstratios N. Pistikopoulos, Julien Walzberg, Funda Iseri, Halil Iseri, Natasha J. Chrisandina, Shivam Vedant, and Catherine Nkoutche. Next-generation reverse logistics networks of photovoltaic recycling: Perspectives and challenges. Solar Energy, 271:112329, 3 2024.

Conformal Prediction: Calibrated Decision-Making

Akash Singh
HEC - University of Liege
email:Akash.singh@uliege.be

Ashwin Ittoo
HEC - University of Liege
email:Ashwin.ittoo@uliege.be

Pierre Ars
Lead Actuarial Innovation, Ethias
email:Pierre.ars@ethias.be

Elise Vandomme
HEC - University of Liege
email:Elise.vandomme@uliege.be

Keywords: Uncertainty, Conformal prediction, Machine learning

Quantifying uncertainty is crucial for assessing the trustworthiness of Machine learning (ML) predictions. Factors like data noise or transformations can influence uncertainty in ML predictions. Standard ML models like Random Forest(RF), XGBoost (XGB), and Neural Networks(NN) provide accurate point predictions but lack calibrated confidence or statistical guarantees, often producing uncalibrated probabilities without inherent uncertainty representation. While these methods perform well, they are not inherently designed to provide coverage guarantees. Coverage guarantee is theoretical assurance that the prediction sets produced by the conformal method will contain the true label of a given instance with a specified probability $1 - \alpha$, where α is the chosen significance level [1]. Conformal prediction (CP) [1, 2] addresses these limitations by providing explicit coverage guarantees through prediction intervals or calibration sets, correcting the miscalibrated probabilities, and ensuring that the true value lies within the interval with a specified probability (e.g., 90%). A CP set is a set of all possible class labels that a given data point could be classified into. Additionally, CP is a model-agnostic framework that is flexible and versatile and can be applied to any ML model.

In this work, we apply CP to a leading Belgian insurance company's data and binary churn prediction model and to several widely available public churn and fraud prediction datasets from prior research [4, 5, 6]. In a binary churn classification model, each customer's profile x is assigned to one of two discrete classes i.e. *churn (0)* or *not churn(1)*. For a binary classification problem, the "CP set" can take one of the following four forms: $\emptyset, \{0\}, \{1\}, \{0, 1\}$. Conformal Prediction generalizes the concept of hypothesis testing by calculating *p-values* for each potential class. We further define the Significance level (α) that specifies the desired coverage level ($1 - \alpha$). For instance, if α=0.1, the prediction set is designed to contain the true label at least 90% of the time. The threshold in conformal prediction is a value that determines the size of the prediction set. It's calculated based on the distribution of nonconformity scores. Thus if the *p-value* for a class exceeds a threshold (based on the "significance level α"), that label is

included in the prediction set. The threshold is derived from the α-quantile of a set of nonconformity scores.
To illustrate the methodology, we represent a profile in the portfolio with x where **significance level** is $\alpha = 0.1$ with corresponding **threshold** of 0.30. Let $P(C|x)$ denote the churn probability output of a predictive model (e.g., RF or XGB). Let $CC(x)$ represent the "Conformal Prediction set (or class)".
The crucial result is that it can be proven that the final decision regarding the profile x (churn or not churn) belongs to the Conformal Prediction set $CC(x)$ with a probability of at least $1 - \alpha$ [1]. This result applies globally to the entire dataset for both classes.
For example consider three instances of profile x_1, x_2, x_3 where each instance yields a distinct CP set.
1. **Profile** x_1: Assume $P(C|x_1) = 0.15$. Then $P(C|x_1) < 0.30$, we have $1 - P(C|x_1) \geq 0.30$. Therefore, $CC(x_1) = \{0\}$. With significance level $\alpha = 0.1$, the conformal prediction for x_1 not churning is: $P(x_1 \text{ will not churn}) \geq 0.90$.
2. **Profile** x_2: Assume $P(C|x_2) = 0.75$. So $P(C|x_2) \geq 0.30$ and $1 - P(C|x_2) \leq 0.30$. Therefore, $CC(x_2) = \{1\}$. Because $\alpha = 0.1$ the conformal prediction of x_2 will churn is: $P(x_2 \text{ will churn}) \geq 0.90$.
3. **Profile** x_3: Assume $P(C|x_3) = 0.60$. Then $P(C|x_3) \geq 0.30$ and $1-P(C|x_3) \geq 0.30$. Thus $CC(x_3) = \{0, 1\}$. For x_3, there is insufficient evidence to confidently classify it as either churn or not churn with a confidence level of at least 0.90. The Conformal Prediction set $\emptyset$ can occur if, for example, the threshold is 0.60 and $P(C|x) = 0.55$. The CP framework enables decision-makers to reduce uncertainty. For example, a fraud manager can confidently act on cases classified as solely fraudulent or non-fraudulent (CP sets 0 or 1), with confidence, knowing the fraud probability is less than 10% or at least 90%, respectively. Uncertain cases (CP set {0,1}) can be flagged for further analysis. CP offers reliable uncertainty quantification, calibrated probabilities, flexible decision-making, and theoretical guarantees, addressing the limitations of traditional machine learning.

References

[1] Vladimir Vovk, Alexander Gammerman, and Glenn Shafer, *Algorithmic Learning in a Random World, Second Edition*, Volume 29, Springer, 2022.

[2] Eyke Hüllermeier and Willem Waegeman, *Aleatoric and Epistemic Uncertainty in Machine Learning: An Introduction to Concepts and Methods*, Machine Learning, Volume 110, Number 3, Pages 457–506, Springer, 2021.

[3] Ryan Tibshirani, *Conformal Prediction, Advanced Topics in Statistical Learning*, Spring 2023, Berkeley.

[4] Aman Chauhan, *Fraud Detection*, Kaggle.com, Available at: `https://www.kaggle.com/datasets/whenamancodes/fraud-detection/data`, 2015.

[5] BlastChar, *Telco Customer Churn*, Available at: `https://www.kaggle.com/datasets/blastchar/telco-customer-churn`, 2018.

[6] Shubham Kumar, *Churn Modelling Bank*, Available at: `https://www.kaggle.com/datasets/shubh0799/churn-modelling`, 2020.

Optimizing for forecast stability in distribution-free probabilistic forecasting

Jente Van Belle
KU Leuven, Faculty of Economics & Business
e-mail: jente.vanbelle@kuleuven.be

Honglin Wen
Shanghai Jiao Tong University, Department of Electronic Engineering

Wouter Verbeke
KU Leuven, Faculty of Economics & Business

Pierre Pinson
Imperial College London, Dyson School of Design Engineering

Rolling origin forecast instability can be defined as the variability in forecasts for a specific time period induced by updating the forecast for this period when new observations become available, i.e., as time passes [1]. Forecast updates are considered to result in both benefits and costs. The benefits are caused by an increase in forecast quality due to a shorter forecast horizon, while the costs stem from induced forecast instability, which may lead to costly changes to plans formulated based on the forecasts. Additionally, forecast instability can erode trust in the forecasting system, potentially prompting unwarranted judgmental adjustments by users.

Methodologies exist to enhance forecast stability without compromising forecast quality in both point and Gaussian probabilistic forecasting settings [1, 2]. Furthermore, application of these methodologies can lead to improvements in both forecast stability and quality, suggesting that they can also serve as time-series-specific regularization mechanisms.

In this paper, we aim to integrate forecast stability alongside quality into the optimization of distribution-free probabilistic time series forecasts to extend the above findings to the distribution-free setting. To achieve this goal, we propose a method to generate stabilized forecasted conditional quantile functions. These quantile functions are modeled using linear isotonic regression splines, with the parameters learned through training a neural network using a discretized approximation of the continuous ranked probability score as the loss function [3]. Unlike approaches based on parametric probability density functions or those forecasting only a fixed set of quantiles, our proposed method offers the flexibility to produce full density forecasts for various output distributions without explicit specification, while also preventing quantile crossing. Furthermore, the condi-

tional quantile function approach for characterizing density forecasts provides the essential flexibility to optimize forecast stability alongside quality. Specifically, it enables us to control the level of dissimilarity allowed at each forecast update, while also offering the flexibility to place varying importance on different parts of forecast distributions (e.g., central part vs. tails).

We empirically demonstrate the effectiveness of our proposed approach on multiple datasets that exhibit different statistical properties.

References

[1] Van Belle, J., Crevits, R., & Verbeke, W. (2023). Improving forecast stability using deep learning. *International Journal of Forecasting, 39*(3), 1333-1350.

[2] Van Belle, J., Crevits, R., Caljon, D., & Verbeke, W. (2024). Probabilistic forecasting with modified N-BEATS networks. *IEEE Transactions on Neural Networks and Learning Systems, 35*(12), 18872-18885.

[3] Gasthaus, J., Benidis, K., Wang, Y., Rangapuram, S. S., Salinas, D., Flunkert, V., & Januschowski, T. (2019). Probabilistic forecasting with spline quantile function RNNs. *Proceedings of the Twenty-Second International Conference on Artificial Intelligence and Statistics*, in *Proceedings of Machine Learning Research* 89:1901-1910.

Literature Review Time Series in Process Mining

Aurélie Leribaux, Johannes De Smedt, Jochen De Weerdt
KU Leuven, Research Centre for Information Systems Engineering (LIRIS)
e-mail: aurelie.leribaux@kuleuven.be

Time Series in Process Mining Time series data, which are characterized by their sequential and temporal nature, offer a powerful tool for discovering, analyzing, and improving business processes. For example, utilization rates of resources such as loan officers recorded continuously can be merged with discrete actions such as opening a loan application file and submitting a recommendation. Over recent years, many approaches combining these two temporal domains have been introduced, however, most techniques offer ad-hoc combinations which have not been mapped in a comprehensive way. This literature review investigates the role of time series data in process mining, focusing on identifying the specific ways in which such data are utilized in the field.

The reviewed studies highlight an increasing focus on integrating time series data into process mining to uncover deeper insights, identify inefficiencies, and enhance outcomes. This integration, coupled with advanced analytical techniques, is demonstrating significant potential in real-world applications. Industries such as manufacturing (Bertrand et al., 2023), healthcare (Su et al., 2023), and tourism (Kołakowska and Godlewska, 2022), where business processes and continuous monitoring are essential, are particularly benefiting from these advancements.

Despite promising developments, several gaps remain in the integration of time series data with process mining tasks. Specifically, these gaps relate to unexplored or underdeveloped combinations of time series types and the process mining tasks they can address. Bridging these gaps is essential for advancing the process mining field. Future research should prioritize the development of robust analytical frameworks capable of addressing these challenges. By overcoming these limitations, process mining can continue to evolve, unlocking more sophisticated applications and driving progress across diverse industries.

Proposed Approach The study categorizes the findings from recent literature along three primary dimensions: (1) the types of time series data employed in a business process context, (2) the process mining tasks addressed, and (3) the techniques applied to create, transform, or analyze time series data. First, the review distinguishes between native time series data, which are directly recorded time series next to data generated from a business process execution (e.g. prices of tourist prices over time), and derived time series, which are generated through the aggregation or transformation

of event log data. This distinction is important because it highlights the variety of ways time series data can be integrated into process mining workflows and the different challenges each type presents.

Second, the review examines the specific process mining tasks that benefit from the use of time series data. These tasks include process discovery, conformance checking, predictive process monitoring, performance analysis, action-oriented process mining, and comparative process mining (van der Aalst, 2022). Moreover, three task types were added, including event log creation, drift detection, and decision mining in order to have a more broad inclusion of process mining task types applied to or incorporating time series.

Third, the review explores the techniques applied to time series data for a specific process mining task type. Three particular techniques were uncovered: Machine learning (ML), traditional time series analysis techniques and visualization techniques. ML, including supervised and unsupervised learning algorithms, have been particularly successful for several tasks. Traditional time series analysis methods, such as trend and seasonality decomposition and statistical forecasting, are also commonly employed. The intersection of these techniques with process mining tasks has led to a richer understanding of process dynamics and opened up new opportunities for enhancing process performance, process discovery, and many more. Visualization techniques aim at visualizing the intersection of time series data and process mining tasks, enabling more intuitive insights.

In conclusion, this literature review highlights the growing importance of time series data in process mining, while also pointing out key areas for future research. Bridging these gaps will enable the development of more sophisticated analytical frameworks, which could significantly enhance the ability to analyze and optimize business processes in a wide range of real-world scenarios.

References

Bertrand, Y., De Weerdt, J., and Serral, E. (2023). A novel multi-perspective trace clustering technique for iot-enhanced processes: A case study in smart manufacturing. In Di Francescomarino, C., Burattin, A., Janiesch, C., and Sadiq, S., editors, *Business Process Management. BPM 2023. Lecture Notes in Computer Science*, volume 14159 of *Lecture Notes in Computer Science*, pages 423–438. Springer, Cham.

Kołakowska, A. and Godlewska, M. (2022). Analysis of factors influencing the prices of tourist offers. *Applied Sciences*, 12:12938.

Su, Z., Yu, T., Lipovetzky, N., Mohammadi, A., Oetomo, D., Polyvyanyy, A., Sardina, S., Tan, Y., and van Beest, N. (2023). Data-driven goal recognition in transhumeral prostheses using process mining techniques. *arXiv preprint arXiv:2309.08106*. Presented at the 5th International Conference on Process Mining (ICPM 2023).

van der Aalst, W. M. P. (2022). *Process Mining: A 360 Degree Overview*, pages 3–34. Springer International Publishing, Cham.

Distributed e-Fuel Hubs (DEFH): A case study of a Belgian Fischer-Tropsch liquids hub

Samy Mokeddem[1*] Bardhyl Miftari[1]
Victor Dachet[1] Guillaume Derval[1] Damien Ernst[1]
[1] Smart Grids Lab, University of Liège, Belgium
e-mail: *samy.mokeddem@uliege.be

This paper introduces the concept of distributed e-fuel hubs (DEFHs), which consist of power-to-x facilities strategically located near e-fuel demand centers. "Distributed" refers to rather small units scattered across a country close to e-fuel demand centers (EDCs). DEFHs valorize the local industry's waste outputs, to convert them into e-fuel, and the e-fuel synthesis by-products. The EDCs can be airports for plane refueling, ports for ship refueling, or heavy industries that require e-fuels for high-temperature processes.

This work explores the potential of DEFHs through a case study of a DEFH in Belgium synthesizing Fischer-Tropsh (FT) products (i.e. e-kerosene, e-diesel, e-naphtha) with a lime plant as the source of CO_2. This case study is illustrated in Figure 1. CO_2 is captured from the plant's flue gases using post-combustion carbon capture (PCCC). The waste heat generated during e-kerosene synthesis is valorized in a district heating network. The other e-fuel by-product considered is oxygen which is valorized on market price.

Using this case study, we illustrate the appeal of our approach in terms of production cost reduction. This problem is tackled as a linear program where we optimize the asset's capacities and operations. As our approach is new there is no established market and there is a lot of uncertainty regarding the resale price of e-fuels, by-products, and CO_2 taxes. To assess the impact of these uncertainties on the final production cost, we conduct several sensitivity analyses:

- **e-fuel resale prices**, which varies between 0€ and 500€ per MWh.
- **Waste Heat price**, which varies between 0€ and 40€ per MWh.
- **Oxygen price**, which varies between 0€ and 40€ per ton.
- **the regional CO_2 taxes**, which varies between 0€ and 80€ per ton converted into e-fuel.

Based on these sensitivities, we conclude that the combination of by-product valorization can play a significant role in the adoption of the DEFHs.

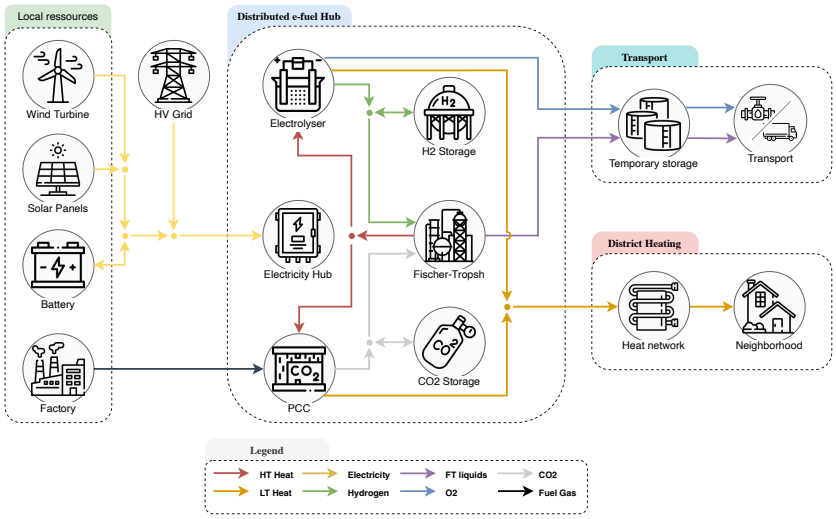


Figure 1: Diagram of the key component of the study case's DEFH. The "Electricity Hub" is an abstract connection point to which each component is considered to have access. For clarity, individual electricity connections are not drawn.

Reconstruction and Compression of sparse network constrained trajectory data

Gilles Dejaegere
Université libre de Bruxelles (ULB), DES-LAB
e-mail: gilles.dejaegere@ulb.be

Christos Doulkeridis
University of Piraeus, University of Piraeus, Department of Digital Systems
e-mail: cdoulk@unipi.gr

Mahmoud Sakr
Université libre de Bruxelles (ULB), DES-LAB
e-mail: mahmoud.sakr@ulb.be

During the last decades, the rapid proliferation of mobile devices equipped with tracking capabilities has led to a surge in the production of spatio-temporal data. The growing availability of this spatio-temporal data has become a valuable resource for various applications such as traffic management, urban planning, and navigation. This significant increase in trajectory data generation however yields high storage and transmission costs. A current avenue of research alleviating these costs consists in compressing these trajectories directly on the edge (online, in real-time). Furthermore, GPS data suffers from inaccuracies and network constrained trajectories require mapmatching techniques to align recorded points with the road network. While traditional mapmatching methods perform well with dense location data, sparse datacaused by low sampling rates or intermittent signal losspresents significant challenges. This work addresses these two challenges by proposing an integrated online methodology for both trajectory reconstruction and compression with sparse GPS data on road networks.

State of the art trajectory compression techniques can be differentiated according to whether or not they rely on historical data to compress currently considered trajectories. Methods relying on historical data, for instance, often exploit compression techniques based on similar reference trajectories or build and exploit dictionaries of frequently used subpaths [1, 2]. These approaches are however less suitable for the contexts with sparse data and when no historical data is available. Instead, methods which are not based on historical data often try to simplify trajectories by exploiting some heuristics reflecting assumptions on most drivers behaviours. This heuristic can for instance be based on the tendency of drivers to favor the shortest paths or to minimize the number of sharp turns [2, 3]. Indeed, if some intermediate points of a trajectory can be deterministically recomputed based on the defined heuristic as well as other points of

this trajectory, then these recomputable points do not need to be transferred nor kept in memory. While compression based on shortest path assumptions provide efficient compression, one of their major drawback is the computational cost of computing the shortest path, rendering it less adequate for real time usage at the edge (especially with sparse input data).

This work aims at tackling this limitation by exploiting a more computational efficient approximate shortest path algorithm. This heuristic is used both for the reconstruction and the compression of the trajectories. First, it is used to fill large gaps in the input GPS point sequence and then it compresses these dense and enriched trajectories.

This approach is illustrated and evaluated using real life trajectories collected by a taxi service in the city of Singapore. It is be compared to state of the art techniques in term of accuracy, computational cost and applicability.

References

[1] Li, T., Chen, L., Jensen, C. S., & Pedersen, T. B. (2021). TRACE: Real-time compression of streaming trajectories in road networks. *Proceedings of the VLDB Endowment*, 14(7), 1175-1187.

[2] Han, Y., Sun, W., & Zheng, B. (2017). COMPRESS: A comprehensive framework of trajectory compression in road networks. *ACM Transactions on Database Systems (TODS)*, 42(2), 1-49.

[3] Chen, C., Ding, Y., Xie, X., Zhang, S., Wang, Z., & Feng, L. (2019). TrajCompressor: An online map-matching-based trajectory compression framework leveraging vehicle heading direction and change. *IEEE Transactions on Intelligent Transportation Systems*, 21(5), 2012-2028.

A rollout strategy for electric vehicle charging stations in urban European environments

Bryan Coulier
KU Leuven, VITO, EnergyVille, Department of Computer Science
e-mail: bryan.coulier@kuleuven.be

Hatice Çalık
KU Leuven, EnergyVille, ELECTA

Thijs Becker
VITO, EnergyVille

Greet Vanden Berghe
KU Leuven, Department of Computer Science

The electrification of transportation is a cornerstone of the global transition to green energy. As Electric Vehicle (EV) adoption continues to increase across Europe, the demand for EV Charging Stations (CSs) continues to rise. However, current CS deployment strategies are largely reactive and fragmented, potentially leading to inefficiencies, higher costs and strain on electricity grids due to a lack of robust planning tools.

This study tackles two key challenges: (i) ensuring that current CS placement decisions avoid long-term grid strain and (ii) quantifying the financial savings achievable through an integrated CS deployment approach. We propose a matheuristic algorithm for optimal CS placement within European low-voltage networks, designed to mitigate electricity congestion and minimize wasted resources. Computational results from large-scale urban scenarios demonstrate the scalability of the algorithm and highlight how an integrated rollout strategy significantly reduces grid congestion while delivering substantial cost savings.

By comparing our approach to the current strategy adopted by city planners, we provide valuable insights for policymakers and city planners concerning on how to prevent foreseeable problems and achieve cost-effective, future-proof CS deployment.

The strategic assembly line feeding problem

Hannah Verplancke
Ghent University, Department of business informatics and operations research
FlandersMake@UGent core lab CVAMO, Ghent, Belgium
e-mail: hannah.verplancke@ugent.be

Veronique Limère
Ghent University, Department of business informatics and operations research
FlandersMake@UGent core lab CVAMO, Ghent, Belgium
e-mail: veronique.limere@ugent.be

El-Houssaine Aghezzaf
Ghent University, Department of Industrial Systems Engineering and Product Design
FlandersMake@UGent core lab ISyE, Ghent, Belgium
e-mail:elhoussaine.aghezzaf@ugent.be

Emmanouil Thanos
Ghent University, Department of business informatics and operations research
FlandersMake@UGent core lab CVAMO, Ghent, Belgium
e-mail: emmanouil.thanos@ugent.be

The rising demand for customized products has driven manufacturing companies to adopt mixed-model assembly line systems, where multiple variants of the same product are manufactured on a single line. This shift has significantly increased the number of distinct parts required at the border of line, posing substantial challenges to intralogistics systems. Managing the transportation and storage of these parts is further complicated by the limited space available at the assembly line border, making it impossible to store all parts in pallets.

To overcome these logistical challenges, many manufacturing companies have introduced preparation cells, where pallets are repacked into smaller load carriers. The manufacturing sector has developed distinct feeding policies, each varying in load carrier size and content. By selecting the optimal feeding policy for each part based on its characteristics, the assembly line feeding problem establishes a smoother flow of parts from the warehouse to the assembly line through preparation cells.

While the traditional assembly line feeding problem focuses on assigning feeding policies to individual parts, the strategic assembly line feeding problem addresses long-term operating structures. This research considers two key decision levels:

1. Facility layout decisions, which determine the number, location, size, and policy-handling capabilities of preparation cells.
2. Vehicle fleet decisions, which define the type and number of vehicles required to transport parts efficiently from the warehouse to the preparation cells and assembly line.

Existing literature typically addresses these dimensions separately: the supermarket location problem focuses on facility layout, while the transport vehicle selection problem addresses fleet composition. However, these two aspects are highly interdependent and significantly influence the overall assembly line feeding decisions.

We propose an integrated approach to the strategic assembly line feeding problem, combining facility layout and vehicle fleet decisions while accounting for part-specific feeding policies. We developed a mathematical model that minimizes the total long-term cost of the intralogistics system. The model considers detailed capacity constraints for both the preparation areas and the border of line, as well as routing restrictions for the employed transportation vehicles. It incorporates both fixed costs (e.g., vehicle acquisition and preparation cell construction) and variable costs (e.g., vehicle operation, part preparation, and part handling), providing manufacturing companies with a comprehensive tool for strategic intralogistics optimization. The performance of the proposed model is assessed using a set of real-world based instances.

Tactical optimization for part feeding in assembly lines

Gohar Porbar
Ghent University, Department of Industrial Systems Engineering and Product Design
FlandersMake@UGent core lab ISyE, Ghent, Belgium
e-mail: gohar.porbar@ugent.be

Emmanouil Thanos
Ghent University, Department of business informatics and operations research
FlandersMake@UGent core lab CVAMO, Ghent, Belgium
e-mail: emmanouil.thanos@ugent.be

El-Houssaine Aghezzaf
Ghent University, Department of Industrial Systems Engineering and Product Design
FlandersMake@UGent core lab ISyE, Ghent, Belgium
e-mail: elhoussaine.aghezzaf@ugent.be

Veronique Limère
Ghent University, Department of business informatics and operations research
FlandersMake@UGent core lab CVAMO, Ghent, Belgium
e-mail: veronique.limere@ugent.be

In today's competitive market, customers increasingly demand products tailored to their specific needs. This requires manufacturers to produce a wide range of product variants, and operate in a high-mix, low-volume environment. To meet this challenge, mixed-model assembly lines are employed, enabling these product variants to be assembled on the same line. However, the space available at the Border of Line (BoL) is limited, making it impossible to store the required parts in large quantities at the BoL.

To cope with this BoL space limitation manufacturing companies use different part-feeding policies. For instance, instead of storing all parts in complete pallets at the BoL, only parts needed for a specific variant can be pre-picked and placed in kits, which are then delivered to the line. This reduces the space required at the BoL, as only the required parts are delivered when needed. Of course, this policy comes with additional costs for preparing and transporting the kits. A trade-off between the utilized space at the BoL and the logistics costs must be made. Optimization approaches must thus provide efficient part feeding solutions while coping with the BoL space limitation.

Research on the tactical level of the part feeding problem very rarely consider

real-world limitations, such as fixed layout, preparation cell sizes and limited vehicle fleets. In our research we consider five different part-feeding policies:

1. Line Stocking: Supplying parts in full pallets directly to the BoL.
2. Boxed Supply: Repackaging parts into smaller boxes before delivering them to the BoL.
3. Sequencing: Presenting parts in the exact order they will be needed on the line.
4. Stationary Kitting: Grouping parts required for one specific product and one assembly station in a kit.
5. Traveling Kitting: Preparing kits which travel along the line through multiple stations.

These policies vary in terms of their space requirements at the BoL, the mix of parts in the containers and the related logistics costs. Kitting, sequencing and boxed supply require additional handling, which are done in preparation areas known as supermarkets. In those areas parts are repacked into smaller containers or grouped into kits which contain all parts required for a specific product.

We propose an exact solution approach to efficiently assign feeding policies to parts such that logistics costs are minimized. The developed MIP formulation considers real-world routing restrictions, in addition to the layout, preparation cell sizes, vehicle-fleet and BoL-space limitations. The models performance is assessed using a set of diverse real-world instances. By incorporating real-world constraints, the model serves as a decision-support tool, providing practical insights to help manufacturers efficiently manage limited BoL space and reduce logistics costs.

A local-search-based heuristic method for efficient assembly line part delivery

Gül Gündüz Mengübaş, Kenneth Sörensen, and Christof Defryn
Department of Engineering Management, Antwerp Operations Research Group-ANT/OR, University of Antwerp, Antwerpen, Belgium
e-mail: gul.gunduzmengubas@uantwerpen.be

This study addresses the critical challenge of optimizing assembly line feeding systems in the automotive industry, particularly under the just-in-time (JIT) concept. Efficiently delivering parts to workstations using tow trains is essential for maintaining seamless production flow and minimizing delays. The research aims to optimize the ALFP by addressing three key objectives: minimizing travel distance, reducing the number of tow trains, and managing penalty costs associated with split deliveries. Unlike exact optimization methods, which often struggle with scalability and efficiency in complex operational settings, this study introduces a heuristic approach tailored to balance these objectives.

This study investigates split delivery to reduce the number of required tow trains and enhance system flexibility. Split delivery allows multiple vehicles to visit the same workstation, enabling incremental part deliveries. This approach increases tow train capacity utilization and improves adaptability; however, it may add complexity and increase travel distance, requiring careful consideration of the trade-offs.

A central focus of this work is the analysis of the relationship between cycle time, delivery volume, and split delivery requirements, considering their impacts on operational planning, routing, and capacity utilization. The findings provide insights into how these factors interact in a multi-objective framework, offering practical strategies for optimizing ALFP systems. The proposed methodology is validated through computational experiments, demonstrating its ability to generate near-optimal solutions efficiently across a range of scenarios.

This research contributes to understanding how operational decisions in ALFP affect cycle time and resource allocation, providing a foundation for further studies in assembly line logistics and optimization.

Native design bias in Large Language Models

Manon Reusens
KU Leuven, LIRIS
e-mail: manon.reusens@kuleuven.be

Philipp Borchert
KU Leuven
LIRIS

Jochen De Weerdt
KU Leuven
LIRIS

Bart Baesens
KU Leuven, LIRIS
University of Southampton,
School of Management

Keywords: Large Language Models, Native Design Bias

Large Language Models (LLMs) are trained on massive text corpora dominated by English, the global lingua franca. However, these corpora primarily represent privileged English dialects spoken in urban, Western contexts, often overlooking the diversity of English usage worldwide. Prior research has demonstrated that such biases can influence model performance across English dialects spoken by native speakers from different regions. Yet, English extends far beyond native-speaking populations, serving as an additional language for millions globally.

In this study, we investigate whether LLM performance differs when prompted by native English speakers versus non-native English speakers. Our findings reveal notable disparities: we find performance differences when models are prompted by native speakers compared to non-native ones. Furthermore, a strong anchoring effect emerges when the model identifies or is made aware of a user's linguistic background, further degrading response quality for non-native users. These findings are based on a novel dataset of over 12,000 annotations from 124 participants, enriched with detailed metadata on their native language and English proficiency. This study underscores the urgent need to address these biases, emphasizing the importance of equitable LLM design and deployment to ensure fair and consistent performance across diverse user groups worldwide.

Making Sense of BERTopic: A Deep Dive into Topic Reduction Techniques

Wannes Janssens
Ghent University, Department of Marketing, Innovation and Organisation
FlandersMake@UGent-corelab CVAMO, Ghent, Belgium
e-mail: wanjanss.janssens@ugent.be

Matthias Bogaert
Ghent University, Department of Marketing, Innovation and Organisation
FlandersMake@UGent-corelab CVAMO, Ghent, Belgium

Dirk Van Den Poel
Ghent University, Department of Marketing, Innovation and Organisation
FlandersMake@UGent-corelab CVAMO, Ghent, Belgium

Keywords: BERTopic, Topic Reduction, Topic Modelling, Large Language Models, NLP

Introduction: BERTopic [1] is a topic modelling framework, widely adopted by practitioners and researchers, thanks to its user-friendly Python package[1] and proven ability to generate high-quality topics. Its default algorithm involves four modules: (i) document embedding using SBERT, (ii) dimensionality reduction via UMAP, (iii) clustering with HDBSCAN, and (iv) a representation module combining CountVectorizer and c-TF-IDF. Since each module relies on its own underlying algorithms, each governed by a distinct set of parameters, BERTopic becomes a complex framework with a large parameter space for tuning. In addition, the clustering module often generates an excessive number of topics which need to be reduced either for downstream tasks or for comparison to other topic modelling techniques.

BERTopic includes built-in functionalities for both direct topic reduction, performed during the fitting process of the BERTopic model, and indirect topic reduction, applied after the model has been fitted. These methods can in turn be applied on different topic embeddings. Users, however, have limited insight into how the framework arrives at a specific set of topics and how different reduction techniques affect the quality of the reduced topics. Furthermore, while the ability of large language models to aid in the topic modelling process has been researched for certain aspects, such as for creating more interpretable topic representations, or enabling end-to-end topic modelling, no study has yet explored their potential to assist in the topic reduction process.

[1] https://maartengr.github.io/BERTopic/index.html.

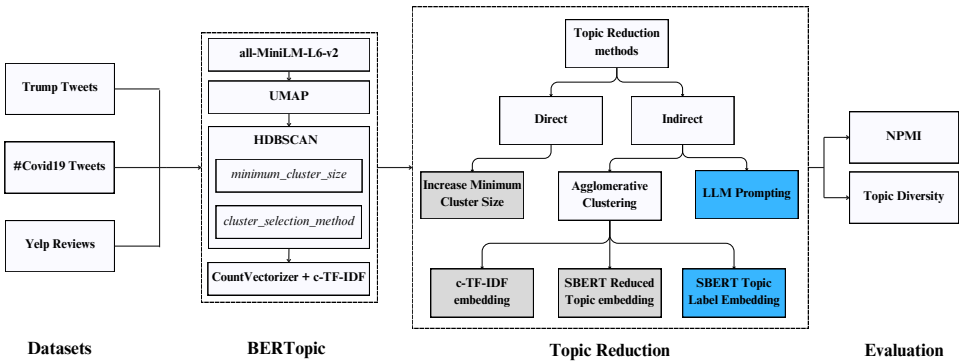


Figure 1: Research Methodology

Methodology: Our study provides a detailed interpretation of the HDBSCAN algorithm and its underlying parameters, as well as a comprehensive description of existing topic reduction techniques for BERTopic. In addition to these, we introduce a number of new methods to the analysis, some of which leverage LLMs, either by prompting an LLM to directly merge topics or by using the SBERT embeddings of the LLM-generated topic labels. We then evaluate the impact of these methods on the coherence and diversity of the reduced topics across two social media datasets from Twitter/X and one review dataset from Yelp as well as across a range of initial HDBSCAN configurations. A full overview of the research methodology is given in figure 1. Through this analysis, we aim to provide insights into some of the black-box aspects of BERTopic and help users make more informed choices in topic reduction.

Results: The preliminary findings reveal significant differences in the quality of the reduced topics across different reduction methods. Indirect methods tend to produce more diverse topics, with a slight decrease in coherence compared to direct methods. Additionally, variations in coherence and diversity are observed depending on the initial BERTopic configuration and the type of topic embeddings used for clustering. Finally, LLMs can effectively identify suitable topics for merging, leading to a reduced set of topics that is more diverse, though not always more coherent. Additionally, larger models appear to outperform smaller (open-source) models.

References

[1] Maarten Grootendorst. Bertopic: Neural topic modeling with a class-based tf-idf procedure. *arXiv preprint arXiv:2203.05794*, 2022.

Ordinal Regression for Question Difficulty Estimation with Transformer-Based Neural Networks

Arthur Thuy
Ghent University, Department of Marketing, Innovation and Organisation
FlandersMake@UGent, core lab CVAMO
e-mail: arthur.thuy@ugent.be

Ekaterina Loginova
Dedalus Healthcare
e-mail: ekaterina.d.loginova@gmail.com

Dries F. Benoit
Ghent University, Department of Marketing, Innovation and Organisation
FlandersMake@UGent, core lab CVAMO
e-mail: dries.benoit@ugent.be

Question Difficulty Estimation (QDE), also referred to as question calibration, involves predicting the difficulty of a question based solely on its text and answers. This task is critical in personalized learning systems, such as computerized adaptive testing, which tailor question selection to match a student's skill level. Providing questions that are too easy or too challenging can demotivate learners, slowing their overall learning progress.

Traditionally, QDE has relied on manual calibration and pretesting, processes that are both time-consuming and expensive. To address these challenges, recent research has turned to natural language processing (NLP) techniques. These approaches involve training machine learning models to predict question difficulty based on its text. Once trained, these models can efficiently calibrate new questions, significantly reducing the dependence on manual calibration and pretesting.

Question difficulty can be represented as either a continuous or discrete value. For continuous difficulty levels, existing research exclusively employs regression models. In contrast, QDE with discrete difficulty levels is essentially an ordinal regression problem, given their inherent ordinal nature and the typically unknown distances between levels. This allows for approaches using regression, classification, or ordinal regression methods. Current studies choose between regression and classification models without comparing their performance, potentially leading to suboptimal outcomes due to inappropriate assumptions. Additionally, specialized ordinal regression techniques remain unexplored. As such, there is no consensus on the optimal method for handling discrete difficulty levels.

This study addresses the gaps in the literature by comparing the performance of standard regression and classification models with specialized ordinal regression

architectures, which have not yet been applied to QDE. Specifically, it evaluates ordinal regression methods based on the extended binary classification framework, namely OR-NN [3], CORAL [2], and CORN [4]. The study focuses on end-to-end Transformer-based neural networks, which have demonstrated superior performance over traditional NLP methods that separate feature engineering from modeling [1]. To this end, two pre-trained Transformers, BERT and DistilBERT, are fine-tuned on both imbalanced and balanced versions of the RACE++ and ARC datasets.

Experimental results reveal that the specialized ordinal method OR-NN consistently ranks among the top-performing approaches, frequently surpassing the more advanced methods CORAL and CORN. In contrast, the regression and classification methods commonly used in existing literature show inconsistent performance, ranging from comparable to OR-NN to significantly worse. Leveraging the full ordinal information inherent in the difficulty levels opens new avenues for advancing personalized learning systems, ultimately benefiting students through a more effective and adaptive educational environment.

References

[1] L. Benedetto. A quantitative study of nlp approaches to question difficulty estimation. In *International Conference on Artificial Intelligence in Education*, pages 428–434. Springer, 2023. doi: 10.1007/978-3-031-36336-8_67.

[2] W. Cao, V. Mirjalili, and S. Raschka. Rank consistent ordinal regression for neural networks with application to age estimation. *Pattern Recognition Letters*, 140:325–331, 2020. doi: 10.1016/j.patrec.2020.11.008.

[3] Z. Niu, M. Zhou, L. Wang, X. Gao, and G. Hua. Ordinal regression with multiple output cnn for age estimation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4920–4928, 2016.

[4] X. Shi, W. Cao, and S. Raschka. Deep neural networks for rank-consistent ordinal regression based on conditional probabilities. *Pattern Analysis and Applications*, 26(3):941–955, 2023. doi: 10.1007/s10044-023-01181-9.

Learning Analytics Dashboard for Conceptual Modeling Education

Elena Tiukhova
KU Leuven, Research Center for Information Systems Engineering
e-mail: elena.tiukhova@kuleuven.be

Charlotte Verbruggen
KU Leuven, Research Center for Information Systems Engineering
e-mail: charlotte.verbruggen@kuleuven.be

Tinne De Laet
KU Leuven, Tutorial Services
e-mail: tinne.delaet@kuleuven.be

Bart Baesens
KU Leuven, Research Center for Information Systems Engineering
e-mail: bart.baesens@kuleuven.be

Monique Snoeck
KU Leuven, Research Center for Information Systems Engineering
e-mail: monique.snoeck@kuleuven.be

Teaching conceptual modeling requires a well-structured instructional strategy, as the tasks involved require the development of complex cognitive skills. To assist learners, educators frequently incorporate technology, such as modeling and prototyping tools, as well as online materials that allow for extensive collection of learning data. Student-generated data can be leveraged to produce learning analytics and use them to provide visual feedback on student learning activity, academic performance, and overall study progress through a unified interface known as the Learning Analytics Dashboard (LAD). To aid students in interpreting the visuals given in the LAD, a reference frame is provided, often in the form of peer comparison. Although peer comparison in LADs has been widely studied, its effectiveness remains debated. We examine the impact of offering both individual and peer-referenced versions of the LAD in a conceptual modeling education context. By analyzing student engagement, performance, and LAD usage, we explore how these factors interrelate across students with varying academic backgrounds. The findings indicate that these relationships are significant only when peer-enhanced LADs are used and are particularly pronounced among students with limited modeling experience, highlighting the importance of peer comparison in supporting novice modelers.

Predicting restaurant hygiene scores: an interpretable machine learning approach

Xiao Wang
Ghent University, Department of Marketing, Innovation, and Organization
e-mail: Xiao.Wang@UGent.be

Bram Janssens
Ghent University, Department of Marketing, Innovation, and Organization
FlandersMake@UGentcorelab CVAMO, Belgium
Research Foundation Flanders, Belgium
e-mail: Bram.Janssens@UGent.be

Matthias Bogaert
Ghent University, Department of Marketing, Innovation, and Organization
FlandersMake@UGentcorelab CVAMO, Belgium
e-mail: Matthias.Bogaert@UGent.be

Health departments have adopted inspections to regulate restaurant hygiene, ensuring public food safety and mitigating the risk of foodborne illnesses. To address the challenges associated with inspection practices, two types of raw data (i.e., inspection records and review data) are utilized to develop predictive models aimed at improving the effectiveness of restaurant hygiene monitoring [1]. Whereas research has made great improvements in predicting hygiene scores [1, 2], we identify several gaps in current literature. Firstly, current studies often overlook certain important variables (e.g., food type and number of violations), resulting in the inability to effectively assess value-added features. Secondly, current studies are primarily rooted in traditional text analytics approaches, with few attempts made to explore more advanced methodologies such as embeddings and sentiment analysis tools. There is a need to develop a unified framework that not only integrates value-added features but also enhances the representations of raw features. Finally, current studies lack interpretability, despite the fact that decision-making processes consistently place emphasis on model explainability.

In this study, we introduce an interpretable approach to assist in the decision-making process for restaurant hygiene inspections with two groups of raw data (i.e., inspection records from health departments and review data from Yelp). To mitigate shallow representations of raw data, we propose an approach that involves extracting additional features from inspection records, deriving sentiment scores, and leveraging textual representations to enhance model performance. By integrating all the features, we propose and evaluate various predictive models [1, 3]. Furthermore, we improve the interpretability of our approach by analyzing and highlighting the importance of individual features. Our experimental results

demonstrate that this comprehensive approach can achieve high accuracy and provide enhanced interpretability.

We find that (1) prediction performance derived from inspection records is robust and reliable, (2) key features such as inspection type, number of violations, sentiment scores, and food types play a pivotal role in predictive accuracy, (3) current ensemble learning methods may yield few improvements when paired with more advanced feature engineering, and (4) the utility of review data seems to depend on the frequency of reviews. Therefore, our method holds value for all practitioners within the restaurant industry, providing insights for enhancing food safety practices and regulatory compliance.

References

[1] Siering, M. (2021). Leveraging online review platforms to support public policy: Predicting restaurant health violations based on online reviews. Decision Support Systems, 143, 113474.

[2] Jeon, J., Kim, E., Wang, X., & Tang, L. (2023). Predicting on restaurant's hygiene rating: Does customer review emotion and content matter?. British Food Journal, 125(11), 3871-3887.

[3] Liu, Z., De Bock, K. W., & Zhang, L. (2025). Explainable profit-driven hotel booking cancellation prediction based on heterogeneous stacking-based ensemble classification. European Journal of Operational Research, 321(1), 284-301.

Predicting core returns in B2B machine operations using machine learning

Yaël De Rocker, Wannes Janssens, Brammert Termont
Matthias Bogaert, Dirk Van den Poel
Ghent University, Department of Marketing, Innovation and Organisation
FlandersMake@UGent - corelab CVAMO, Belgium
{yael.derocker, wanjanss.janssens, brammert.termont}@UGent.be
{matthias.bogaert, dirk.vandenpoel}@UGent.be

Remanufacturing involves collecting and processing returned product cores, to repair, recycle, refurbish or remanufacture them. Managing this reverse manufacturing supply chain requires decision-making at both the strategic and operational levels. On the operational level, accurately predicting core returns is critical for achieving sustainable and operational efficiency. In an industrial B2B setting, predicting when and why products are returned is challenging due to the high value, extended usage periods, and complex repair needs of the products involved. As a result, this area of research has received limited attention.

Existing research on product returns has primarily focused on B2C settings such as electronics, fashion and e-commerce [Mishra and Dutta, 2024], where products are typically returned by consumers shortly after purchase. In contrast, studies focusing on B2B settings, where products are returned after longer periods of time, primarily leverage operational data to analyze return trends and optimize inventory management on an aggregated level [Yang et al., 2024]. Moreover, these studies often employ traditional statistical optimization methods rather than machine learning approaches. However, they pay little attention to product-specific return predictions, which consider factors such as product defects, warranty conditions, customer behavior, and supply chain dynamics.

This study addresses these gaps by developing a machine learning-based framework for predicting product returns while also identifying key return drivers. We use a real-world dataset from a European B2B company specializing in the production, sale, and repair of cores for large industrial machines. To analyze these data, we leverage a broad set of operational and product usage variables such as supplier defect rates, product usage patterns, and warranty dynamics. Additionally, instead of forecasting the total number of product returns [Cui et al., 2020], this study employs an instance-specific prediction approach which determines the likelihood of individual product returns within a certain time window. These predictions are then refined using explainable AI methods, which offer insights into the causal factors behind returns and generate actionable recommendations for operational decision-making.

References

[Cui et al., 2020] Cui, H., Rajagopalan, S., and Ward, A. R. (2020). Predicting product return volume using machine learning methods. *European Journal of Operational Research*, 281(3):612–627.

[Mishra and Dutta, 2024] Mishra, A. and Dutta, P. (2024). Return management in e-commerce firms: A machine learning approach to predict product returns and examine variables influencing returns. *Journal of Cleaner Production*, 477:143802.

[Yang et al., 2024] Yang, C.-H., Su, X.-L., Ma, X., and Talluri, S. (2024). A data-driven distributionally robust optimization approach for the core acquisition problem. *European Journal of Operational Research*.

On strong integrality properties of the perfect matching polytope

Roland Grappe
LAMSADE, Université Paris-Dauphine, France
e-mail: roland.grappe@lamsade.dauphine.fr

Mathieu Lacroix
LIPN, Université Sorbonne Paris Nord, France
e-mail: lacroix@lipn.fr

Francesco Pisanu
CORE, Université Catholique de Louvain, Belgium
e-mail: francesco.pisanu@uclouvain.be

1 Introduction

The *perfect matching polytope* is the convex hull of the characteristic vectors of all perfect matchings of a graph. A graph is *matching covered* if each of its edges belongs to a perfect matching. These graphs have been introduced by Edmonds, Pulleyblank, and Lovász in [5] while standing the affine hull of the perfect matching polytope. In this work, we study a class of graphs that — unlikely to the general case [6] — allows a compact description of the perfect matching polytope together with two properties stronger than integrality.

A *tight cut* of a matching covered graph is a cut that shares exactly one edge with each perfect matching. A *brick* is a nonbipartite matching covered graph having no nontrivial tight cuts. A matching covered graph G is *solid* if it has no separating cuts that differ from the tight one, where a cut $\delta(X)$ is *separating* if the two graphs G/X and $G/(V \setminus X)$ — obtained by contracting X and $V \setminus X$ respectively to a single vertex — are also matching covered.

A rational linear system is totally dual integral (TDI) if for every integer linear function for which the optimum is finite the associated dual problem has an integer optimal solution. A TDI system is box-TDI if adding any rational bounds on the variables preserves its TDIness. Box-TDI systems are systems that yield strong min-max relations such as the one involved in the Max Flow-Min Cut Theorem of Ford and Fulkerson. A polyhedron is box-TDI if it can be described by a box-TDI system.

Box-totally dual integral systems and polyhedra received a lot of attention from the combinatorial optimization community around the 80's. A renewed interest appeared in the last decade and since then many deep results appeared involving such systems (see for exemple [2]).

Ding et al. [4] characterized the box-TDIness of the matching polytope. Though the perfect matching polytope is a face of the latter, their result does not characterize the box-TDIness of the perfect matching polytope.

2 Contributions

In this work, we investigate the connection between solidity and box-TDIness by proving a strict inclusion, that is every graph whose perfect matching polytope is box-TDI is solid, and some links with edge-coloring.

First, extending the work of de Carvalho et al. [3], we characterize the perfect matching polytope of solid matching covered graphs and independently prove that also box-TDI perfect matching polytope allows compact descriptions.

Theorem 1 *The perfect matching polytope of a solid matching covered graph is the intersection of its affine hull and the nonnegative orthant.*

Theorem 2 *The perfect matching polytope of a matching covered graph G is box-TDI if and only if there exists a totally unimodular matrix describing it. Moreover, G is solid.*

Then, combining a characterization of box-TDI polyhedra of Chervet et al. [2], we prove the following.

Corollary 3 *Recognizing whether the perfect matching polytope of a given matching covered graph is box-TDI can be done in polynomial time.*

A last consequence built upon the results of Baum and Trotter [1] on the *integer decomposition property* is the following.

Corollary 4 *If every brick obtained from the tight cut decomposition of a d-regular matching covered graph G does not have two vertex-disjoint odd cycles, then G is d-edge-colorable.*

References

[1] S. Baum and Jr. L. E. Trotter. Integer rounding and polyhedral decomposition of totally unimodular systems. In R. Henn, B. Korte, and W. Oettli, editors, *Optimization and Operations Research*, volume 157 of Lecture Notes in Economics and Mathematical Systems, pages 15-23. Springer, Berlin, 1977. Proc. Bonn 1977.

[2] P.Chervet, R.Grappe, and L.H.Robert. Box-total dual integrality, box-integrality, and equimodular matrices. *Mathematical Programming*, 188(1):319-349, 2021.

[3] M.H.de Carvalho, C.L.Lucchesi, and U.S.R. Murty. The perfect matching polytope and solid bricks. *Journal of Combinatorial Theory, Series B*, 92(2):319-324, 2004.

[4] G.Ding, L.Tan, and W.Zang. When is the matching polytope box-totally dual integral? *Mathematics of Operations Research*, 43(1):64-99, 2018.

[5] J.Edmonds, W.R.Pulleyblank, and L.Lovász. Brick decompositions and the matching rank of graphs. *Combinatorica*, 2:247-274, 1982.

[6] T.Rothvoss. The matching polytope has exponential extension complexity. *Proceedings of the forty-sixth annual ACM symposium on Theory of computing*, 2013.

Maximal Matchings

Anh Trieu
Maastricht University, Department Quantitative Economics
e-mail: anh.trieu@maastrichtuniversity.nl

Iwan Bos
Maastricht University, Department of Organisation, Strategy, and Entrepreneurship
e-mail: i.bos@maastrichtuniversity.nl

Marc Schröder
Maastricht University, Department Quantitative Economics
e-mail: m.schroder@maastrichtuniversity.nl

Dries Vermeulen
Maastricht University, Department Quantitative Economics
e-mail: d.vermeulen@maastrichtuniversity.nl

There are many situations in which policymakers are primarily concerned with the availability and accessibility of goods or services. Examples include electricity, food, housing, medical supplies, *et cetera*. In such cases, the social goal may be to maximize the number of transactions, which we refer to as a maximal matching, instead of total utility. This paper presents a mechanism that implements this objective in a unit-demand and unit-supply setting, where each consumer needs one unit of a goods and each producer can produce one unit of such goods. The mechanism satisfies the incentive and participation constraints, but requires external funding. We also attempts to answer the question whether such maximal matching outcome can be obtained in a more complicated setting such as unit-demand but multi-supply, or even multi-demand and multi-supply.

The expected size of maximum matchings in bipartite graphs

Rob Zuidwijk
Erasmus University, Rotterdam School of Management
e-mail: rzuidwijk@rsm.nl

The expected size of maximum matchings in a random Erdős Rényi Gilbert bipartite graph $B(n, m, p)$ is a polynomial in the edge probability p of order nm. This paper provides an explicit expression for such expected size of maximum matchings by means of a recursive formula. This explicit expression is obtained by an extension of the Dulmage and Mendelsohn decomposition. In addition, simple lower and upper bound polynomials are established for expected size of maximum matchings. The study of expected size of maximum matchings of these random graphs is motivated by matching problems in transportation.

The study of maximum matchings in random bipartite graphs can be used as an instrument to analyse the synergies of freight transportation matching. A truck hauler receives import orders and export orders with attributes that follow probability distributions f_I and f_E, respectively. Order attributes, such as geographical location and delivery and pick-up time windows, are in attribute space X. An import order with attributes $x_I \in X$ and an export order with attributes $x_E \in X$ may match in the sense that their combined execution creates synergies. The combined execution of the import and export orders in a single truck haul may result in a trip with smaller truck distance traversed without cargo and associated cost savings. We may put $g(x_I, x_E) = 1$ when a match creates sufficient synergies, and $g(x_I, x_E) = 0$ when there are no or insufficient synergies. The probability of a match between an arbitrary import order and export order is then equal to

$$p = \int_X \int_X g(x_I, x_E) f_I(x_I) f_E(x_E)\, dx_I\, dx_E. \tag{1}$$

We are interested in the expected maximum number of matches that can be made when a number of import orders and a number of export orders are available. This expected number of matches coincides with the expected maximum size of a matching in a random bipartite graph with vertex sets that represent the import and export orders.

A realisation of a bipartite graph $B(n, m, E)$ has edge set $E \subseteq \mathbb{N}_n \times \mathbb{N}_m$. Here $\mathbb{N}_n = \{1, \ldots, n\}$. If E consist of $k(E)$ edges, then the probability of its occurrence equals $P(E) = p^{k(E)}(1-p)^{nm-k(E)}$. The expected maximum number of matches of a random bipartite graph $B(n, m, p)$ is given by the polynomial

$$R_{n,m}(p) = \sum_{E \subseteq \mathbb{N}_n \times \mathbb{N}_m} p^{k(E)}(1-p)^{nm-k(E)} \mu(E). \tag{2}$$

The first objective of this paper is to come up with a recursive formula that helps to arrive at an explicit expression for $R_{n,m}(p)$. Although the expression for $R_{n,m}(p)$ is quite involved in general, the formula for $R_{n,1}(p)$ is simple and can be derived easily as any instance of $B(n, 1, p)$ has maximum matching equal to one if and only if it has edges, so

$$R_{n,1}(p) = 1 - (1-p)^n \text{ for n} \geq 1. \tag{3}$$

The second objective of this paper is to provide explicit upper and lower bounds for $R_{n,m}(p)$ involving these simple expressions. For $n \geq m$ and $n \geq 2$, the following inequalities will be discussed:

$$R_{n-1,1}(p) \leq \frac{1}{m+1} R_{n,m+1}(p) \leq \frac{1}{m} R_{n,m}(p) \leq R_{n,1}(p). \tag{4}$$

Achieving the two objectives of this research requires careful analysis of maximum matchings in bipartite graphs. The analysis involves an extension of the Dulmage and Mendelsohn decomposition. It involves a decomposition of the subset $N_1(E) \subseteq \mathbb{N}_n$ of vertices that are missed by some maximum matching in E. The recursive formula mentioned earlier is based on the study of extensions $E \cup J$ of $E \subseteq \mathbb{N}_n \times \mathbb{N}_m$ with edge sets $J \subseteq \mathbb{N}_n \times \{m+1\}$. In particular, expressions are derived of how the decomposition of $N_1(E)$ informs the decomposition of $N_1(E \cup J)$.

Constraint evaluation techniques for the nurse rostering problem

Robin Tourlamain, Greet Vanden Berghe, Pieter Smet
KU Leuven, NUMA
e-mail: {robin.tourlamain, greet.vanden.berghe, pieter.smet}@kuleuven.be

The nurse rostering problem has received much attention in the operations research community as an optimization problem with important real-world roots. The complex reality of scheduling nurses with different contracts and preferences translates itself into equally difficult constraints. Vanhoucke et al. [1] argues that for an instance to be considered nurse rostering it must have at least coverage constraints, and be limited to one shift per day, per nurse. Other well-known constraint types include but are not limited to:

1. Shift succession constraints
2. Nurse skill levels
3. Nurse preferences
4. Nurse workload constraints
5. Constraints on number of consecutive shifts
6. Constraints on number of consecutive off days
7. Holidays/ vacations

Over the years, multiple public datasets have become available for benchmarking and competition purposes [1, 2, 3], all of which adopt their own subset of constraints. Given some roster, calculating the penalty for some of these constraint types requires a non-negligible amount of computation time. As time limits are a widely used stop criterion, efficient constraint evaluation is an important part of any time-sensitive algorithm. Unfortunately, many publications omit a description of the evaluation procedure. A notable exception is [4], where a flexible evaluation method is well-described. We propose an efficient evaluation method for frequently used constraint types. Our method is designed with a local search context in mind and can significantly reduce the evaluation time of the most complex nurse rostering constraints.

References

[1] M. Vanhoucke and B. Maenhout, "Nsplib–a nurse scheduling problem library: A tool to evaluate (meta-) heuristic procedures," in *Operational research for health policy: making better decisions, proceedings of the 31st annual meeting of the working group on operations research applied to health services*, pp. 151–165, 2007.

[2] S. Haspeslagh, P. De Causmaecker, A. Schaerf, and M. Stølevik, "The first international nurse rostering competition 2010," *Annals of Operations Research*, vol. 218, pp. 221–236, 2014.

[3] S. Ceschia, N. Dang, P. De Causmaecker, S. Haspeslagh, and A. Schaerf, "The second international nurse rostering competition," *Annals of Operations Research*, vol. 274, no. 1, pp. 171–186, 2019.

[4] E. K. Burke, P. De Causmaecker, S. Petrovic, and G. V. Berghe, "Fitness evaluation for nurse scheduling problems," in *Proceedings of the 2001 Congress on Evolutionary Computation (IEEE Cat. No. 01TH8546)*, vol. 2, pp. 1139–1146, IEEE, 2001.

Practical challenges in workforce scheduling

Lotte Berghman & Leendert Kok
ORTEC, Houtsingel 5, 2719 EA, Zoetermeer, The Netherlands
e-mail: lotte.berghman@ortec.com; leendert.kok@ortec.com

Abstract

There is a lot of academic research on workforce scheduling (and nurse rostering in particular). However, these results can not always be applied to practical cases. The trade-off between model accuracy and algorithmic efficiency is challenging and is made differently by academic researchers than in practice. We highlight some of the main differences between academic and practical research concerning workforce scheduling. In practice, many challenges appear in workforce scheduling. Availability of employees, skill levels, and labor rules put a lot of complexity in the modeling of hard constraints. In addition, fuzzy and even contradicting objectives make it challenging to determine when a schedule is good or better than another one. Finally, changes between publication and execution of the roster ask for a tradeoff between robustness and quality of a roster.

1 Benchmark versus practical instances

When we analyze benchmark instances that are usually used in scientific research ([1, 2]), we see quite some differences with the instances we see at our clients. As a result, the methods developed for those scientific instances are not necessarily good methods for real life instances.

A first set of differences concerns the employees that are available. In real life, employees differ from one another in multiple ways. For example, some industries have a majority of employees that do not work fulltime. They have contracts for a limited number of hours, which might be predefined (for example only during the weekends) or not. Another difference concerns the skills employees might have. In real life, there are a lot of different skills, and for each of these skills even different levels. These skill levels may even be department dependent, e.g., when you work in one department on level three, this may become level two if you work temporarily in another department.

Another set of differences concerns the labor rules. If we take the Dutch labor rules (Arbeids Tijden Wet ([3]) as an example, the number of constraints is not small. Moreover, there are often several options and exceptions, which makes the modeling part challenging. To give a small example, some rules put a bound on the number of "consecutive shifts". However, the definition of consecutive shifts (two shifts with at most 32 hours in between) does not necessarily imply shifts on consecutive days. In fact, there are examples of two shifts starting on two consecutive days that are not consecutive shifts. And there are also examples of two shifts starting on day d and on day $d + 2$ that are consecutive shifts.

2 Different objectives

Defining the objective function in rostering problems is a big challenge in practice. Also, it is industry- or even client-specific. There are different stakeholders who all have their own objectives, and they are often contradictory. Certain employees might, for example, prefer to work many consecutive night shifts because of the extra payment. But management sees that those people are more often sick, so they prefer to spread out the night shifts over all employees and to limit the number of consecutive night shifts.
Moreover, some preferences are only present in the planners head because of their sensitive nature. Certain employees, for example, prefer not to work with particular colleagues. Another example is the fact that the size of the problem of not having the entire workload covered depends on the particular employees that are already assigned to that shift. This is because some have less problems working a bit harder or are less sensitive to the extra pressure that comes with under-staffing.
Another challenging question is how to combine all those objectives. A Pareto front does not give very practical information to a planner because he will have a hard time picking the right solution. A weighted sum might seem logical, but is difficult to interpret. How should one sum over different metrics and what does the objective value mean in that case? Moreover, a company typically has no idea about the weights of each individual objective.

3 Robustness

A roster is typically made a few months in advance, because it is important (usually even obligated) to inform employees early about their working hours. But between the moment that the roster is communicated and the moment that the roster is executed, a lot of events might change the input data for that roster. This asks for a trade off between roster quality (in terms of the objective) and robustness of the roster.
After publication, many events occur. Typical events are for example employees calling in sick, (special) leave requests, and changes in the estimated workload. When handling such events, one wants to minimize the number of changes to the published roster. Namely, for every employee who is included in the changes, his/her acceptance is needed. And they might accept a change only under certain conditions which give rise to new events.

References

[1] Haspeslagh, S., De Causmaecker, P., Stølevik, M., & Schaerf, A. (2010). First international nurse rostering competition.

[2] Curtois, T., & Qu, R. (2014). Computational results on new staff scheduling benchmark instances. *Technical report ASAP Research Group, School Comput. Sci., Univ. Nottingham, Nottingham, UK.*

[3] Arbeidstijdenwet. (2022). `https://wetten.overheid.nl/BWBR0007671/2022-08-02`

How to optimize the number of nurses in a hospital

Fabien Nießen
KU Leuven, Department of Computer Science
e-mail: fabien.niessen@kuleuven.be

Paul Paschmanns
Universität Bonn, Forschungsinstitut für Diskrete Mathematik
e-mail: paschmanns@dm.uni-bonn.de

Personnel scheduling aims to generate a timetable for a group of workers while adhering to specific constraints. These constraints may include meeting daily staffing requirements, limiting the number of consecutive night shifts, and potentially optimizing an objective function. Due to its extensive real-world applications, this problem has been studied mostly from a practical point of view. To address the complex constraints involved, approaches such as integer linear programming (ILP) and LP-based heuristics are commonly used [Bur+04].

Although certain forms of the personnel rostering problem are recognized as NP-hard, there is a lack of theoretical analysis for many of its specific variants. Most studies merely state the NP-hardness of the general problem without examining whether the particular cases they address inherit this computational complexity.

We focus on the decision variant of the Days On Days Off Scheduling Problem (DODOSP), a personnel scheduling problem introduced by Brunner, Bard, and Köhler [BBK13], which involves a homogeneous workforce. The DODOSP considers a sequence of days, each with minimum staffing requirements, and incorporates various constraints: global limits on the total number of work days and days off per worker (global bounds), as well as local limits specifying upper and lower bounds for consecutive work days and days off (local bounds).

These constraints reflect practical considerations. For instance, labor regulations, such as Germany's "Arbeitszeitgesetz", prohibit working more than 19 consecutive days [Arb24]. Additionally, the total number of work days per employee over the scheduling period may be capped, while vacation days are typically limited. Lower bounds on consecutive workdays or days off can also be practical, reducing unnecessary commuting and proving especially relevant in industries requiring on-site customer visits.

We have demonstrated that the DODOSP is NP-complete if at least one global and one local lower bound are nontrivial. Furthermore, we show that all other cases are represented by two special cases where the problem becomes solvable in polynomial time, as summarized in Table 1. The algorithms we developed for these cases are both efficient and easy to implement.

Our findings offer a clear approach to identifying which subcases of the prob-

bounds		complexity
U_x, l_y with $x, y \in \{w, o\}$	or any superset	
U_w, U_o, u_w, u_o	or any subset	P
u_w, u_o, l_w, l_o	or any subset	P

Table 1: Computational complexity of the DODOSP. U_w and U_o denote the global bounds on work days and off days. u_w and u_o denote the local upper bounds and similar l_w and l_o the local lower bounds.

lem are computationally simple and which are more challenging. This differentiation is especially useful in practical applications, as it can inform decisions about which constraints to adjust or eliminate to simplify the computational complexity of the problem. Additionally, our results can be readily extended to determine the minimum number of workers required to meet the specified staffing demands.

References

[Arb24] Landesinstitut für Arbeitsschutz und Arbeitsgestaltung Nordrhein-Westfalen - KOMNET. *Wieviele Tage darf man nach dem Arbeitszeitgesetz hintereinander arbeiten?* Dialog 4493. Sept. 25, 2024. URL: `https://www.komnet.nrw.de/_sitetools/dialog/4493`.

[BBK13] Jens O. Brunner, Jonathan F. Bard, and Jan M. Köhler. "Bounded flexibility in days-on and days-off scheduling". In: *Nav. Res. Logist.* 60.8 (2013), pp. 678–701.

[Bur+04] Edmund K. Burke et al. "The State of the Art of Nurse Rostering". en. In: *Journal of Scheduling* 7.6 (2004), pp. 441–499. ISSN: 1099-1425.

Index

E

F

G

H

I

J

K

L

W

Y

Z