dr hab. inż. Krzysztof Rządca
Institute of Informatics
University of Warsaw
and Google LLC
krzadca@mimuw.edu.pl

28th March 2025

Reviewer's opinion on Ph.D. dissertation authored by Jakub Wawrzyniak entitled: Combinatorial optimization problems in port logistics

1. Problem and its impact

The thesis studies problems of optimizing resource allocation in maritime container ports, in particular the quay partitioning problem (QPP) and the berth allocation problem (BAP). Container logistics is a vital part of the modern supply chain, modern manufacturing and, thus, one of the key elements of the modern world. While this industry stayed somewhat under the radar for years, its importance became apparent during the COVID pandemic, when container port delays and the resulting increases in shipping fees made headlines around the world and significantly contributed to the widespread inflation. Thus, in my opinion, the problem area of the thesis is timely, practically important and additionally relatively less explored than resource allocation problems in other domains (such as computing).

The quay partitioning problem (QPP), introduced in this thesis (Chapter 3), consists of partitioning a long, continuous quay into berths, i.e., slots, or parts of the quay, where ships moor, with an assumption that a ship moors at a single berth; and one, or at most two, ships moor at a single berth at the same time. The thesis motivates the problem using real-world data on ship locations (AIS) in major container ports (Fig 1.1). Starting with this problem, the thesis then studies its natural extensions. Optimizing algorithm portfolio for the berth allocation problem (BAP, Chapter 4) is a key to evaluate the QPP at scale. The ship traffic model (Chapter 5) models the ship location data (AIS) probabilistically, which allows the research to go beyond just a few instances taken directly from the real world data, and instead to generate many realistic instances. Finally, the stochastic QPP (Chapter 6) studies the quay partitioning under many instances, therefore addressing directly the problem of optimization of the container port design.

To conclude, in my opinion, the thesis shows a complete application of operational research methodology: starting from a real world problem, the thesis models it in a sufficient way to capture the key elements in a formal, mathematical model; then discusses realistic algorithms and generation of instances; finally solving these instances and discussing the results.

2. Contribution

The thesis has a number of original contributions to the resource allocation literature. In my opinion, the most important one is the formulation of a new resource allocation problem, the Quay Partitioning Problem (QPP). The thesis motivates the introduction of this problem very well, in particular by analysing data about ship placement in major container ports; and by linking QPP to the known berth allocation problem (BAP). As the solution of the QPP is the input of the BAP, optimizing QPP should lead to better overall resource allocation. I find this contribution especially important, as the introduction of a new, reasonable and real-world-motivated optimization problem is, in my opinion, significantly more difficult than just solving an existing problem with a new method.

In the context of this new problem, the thesis uses the classic operational research methodology: a formal combinatorial hardness proof (Chapter 3.2), an ILP formulation (Chapter 3.4-3.4), and a stochastic formulation (Chapter 6), with an extended analysis of the solutions produced (Chapters 3.6, 4.8, 6.4) and the quality of algorithms (Chapter 6.5).

As solving the QPP includes solving a BAP, but on a much larger scale than previously considered, the thesis also considers two tangential problems: choosing a portfolio of heuristic algorithms to solve the BAP (Chapter 4); and generating a large number of test instances based on a realistic container ship traffic model (Chapter 5). In Chapter 4, the algorithms used are, in general, natural heuristics or meta-heuristics - in contrast to, say, elaborate combinatorial optimization algorithms with worst-case performance guarantees.

To summarize, in my opinion, the contribution of the thesis lies principally in the problem definition, which is then solved using a standard operational research methodology, rather than in novel solution methods. The contribution is indisputable and on a sufficient level for a PhD thesis.

Partial results of the thesis were published in quality journals: Chapter 4 in the European Journal of Operational Research (one of the leading journals in operational research); Chapter 5 in the International Transactions in Operational Research; and Chapter 3 in the International Journal of Applied Mathematics and Computer Science.

3. Correctness

Overall, the principal arguments of the thesis are correct to the degree I could verify. While I list a number of questions or suggested edits below, none of them are substantial enough to undermine the principal conclusions.

The proposed ship traffic model (Chapter 5) is especially important for the field. As the AIS (real world) data is expensive and thus difficult to get and then to share, such probabilistic generators should allow the field to progress.

I appreciate that the thesis shares detailed information on the instances used (e.g., Appendix D and reference [60]). While the research relied on the AIS data, this data cannot be publicly shared; thus creating a probabilistic instance generator for realistic instances greatly improves the reproducibility of the results. I could not, however, find the source code for the generator, nor for the remaining software.

I also appreciate extensive evaluation and then analysis of experimental results, e.g., in Chapter 3.5, 4.8 or 6.4. The thesis not only says which heuristic is better, but also studies difficulty of instances or properties of solutions (e.g., what is the structure of the proposed berth partitions) - which are then summarized as high-level conclusions.

I have the following questions and comments about the results presented in the thesis.

Chapter 2 ("Related Work") starts with "without claims to be exhaustive". While I appreciate that the literature is vast, I believe a PhD thesis should be the place to summarize it. I am not expecting an exhaustive list of every published paper, but rather citing literature reviews for sub-fields deemed further from the main topic of the thesis, and the detailed descriptions of results directly related to the problem studied. Similarly, Table 2.1 should have, in my opinion, additional structure (perhaps with columns categorizing the papers, rather than just listing them). To aid understanding, the chapter would ideally feature a simple schematic drawing, enabling a landlubber to grasp the port organization and its connection to the combinatorial problems defined later.

The definitions of optimization problems and MIPs are generally clear and use standard scheduling notation. Yet, I'd suggest K to describe lengths "counts", rather than "frequencies"; and w to be the value of unit delay of servicing a ship, as the current definition might suggest a problem with rejections. Similarly, in Chapter 3.3, (k,l) is a pair rather than a "couple"; and it is initially difficult to understand what is a "copy" in the definition of t_j^{kl} .

The NP hardness proof (Theorem 1) is difficult to follow. Formatting the mathematics equations and breaking the text more often would make it more accessible. The intuition of the proof (unit processing times, no release dates) are given after the proof.

In Chapters 3.5 and 4.4, experiments are performed on random instances with parameters drawn from uniform distributions. Only after reading the description of the results did I understand why non-realistic instances do not cause methodological problems here. Thus, I think the approach should be justified better, especially given the elaborate model introduced in Chapter 5.

Until Chapter 6, the thesis ignores parallel and distributed programming, a de-facto standard for solving large, difficult combinatorial optimization instances. I wonder how the results of, e.g., Chapter 3 or the discussions in Chapter 4.1, would change if we use a modern machine with hundreds of CPU cores or tens of thousands of GPU cores (note that perhaps some parallelism was used internally by CPLEX in its "standard settings", Chapter 3). Similarly, the ILP for the algorithm portfolio selection (p. 70) counts "cost" as the total execution time of the portfolio - which corresponds to sequential execution or to the notion of work in parallel and distributed programming.

On p.71, when discussing the ILP formulation of the algorithm portfolio, the text gets too informal when saying that "we maximize the chance of covering a new (unseen) instance with the best possible solution". In my opinion, the text should discuss here the issue of the representativeness of the training set, as well as avoid referring to "chance", which might be interpreted probabilistically.

While reading the model of Chapter 5, I missed a frank discussion of the time scale (or time scales) of the problem. I am not sure why the considered instances model arrivals over, approximately, a year - while service times are concentrated around 5 to 30 hours (Figure 5.3). These time scales remind me of scheduling in HPC systems, where it is customary to use (multiple) weeks-long traces, rather than a single year-long trace, as in longer traces some phenomena may average out. In QPP, why is a 1-year long instance better than, say, many instances, each covering a month?

The distance function used to cluster ships' lengths (Eq. 5.1) is unusual (wedge-like). What is the motivation for such a function?

The numerical range in Fig. 5.2 (a) is just 5 - I think it is too small to distinguish between the proposed theoretical distributions. Similarly, in Fig 5.2 (b), it seems that there are in fact two clusters - one with lengths [346,351] and another with lengths [363, 368], with no lengths in between these two ranges.

The processing time model (Chapter 5.3.4) uses the Anderson-Darling statistical test as a goodness-of-fit test. However, the text does not explain why this particular test (and not, say, the Kolmogorov-Smirov test) is used. Also, the A-D test is numerical; yet Table 5.6 uses its results as if it were just ordinal (as the table compares just the number of wins).

The ship arrival model in Chapter 5 does not specify how to set ships' weights. While I understand that it would be very hard to get the ground truth - monetary value of servicing a ship - I would like to see some discussion here, rather than just proposing a simple linear model in Chapter 6.2, which seems somewhat arbitrary.

The introduction to Chapter 6 lacks a clear motivation to study the Stochastic Quay Partitioning Problem (SQPP). I think the goal here is to produce a partitioning that is robust to future (unknown) ship traffic. On a similar note, I lacked a broader discussion of the classic measures of robustness in combinatorial optimization or operational research.

In SQPP, the goal function F(K) (Eq. 6.1) is somewhat unusual as it averages over BAP algorithms (I believe the specific algorithms used are listed in Chapter 6.2, SPT-PRIO, SPTGI-Prio, ..., RND-LA5). The thesis motivates this choice by saying "[This] exposes the quay design process to a greater variety of possible scheduling methods, and thus, builds a broader view on potential dispersion of the partition scores under the future, yet currently unknown, BAP algorithm". One could argue that the broader the portofolio - and the poorer the algorithms - the smaller the impact of quay partitioning on the final result. And, in my opinion, the port designer can suggest not to use a particular class of BAP algorithms on a particular port design.

Analysis of SQPP results contain an interesting study on the similarity of solutions (Chapter 6.4.3). However, the chosen measure of similarity is a Student t-test, which is proposed without significant discussion of alternatives. In particular, the t-test might be overly sensitive to solutions that are statistically different, yet practically leading to very similar numerical outcomes. Statistical methods that estimate the mean distance between solutions are usually more robust in these scenarios. Similar methods might be also used in Chapter 6.5, where an algorithm would be significantly better than Rnd10 if the estimated range of improvement does not contain zero.

In the discussion of the distributed implementation of the SQPP solver (Chapter 6.6), I missed the information about raw runtimes. Also, the text hypotheses that the queuing time depends on "the number of active nodes in the computing cluster": does it mean that the nodes were dynamically switched on and off?

In the discussion of reliability (Chapter 6.3.3), I was surprised by the high number of failed executions (26, compared with 189 successes), resulting in a high estimate of failure probability (0.341). This might suggest supercomputer or scheduler misconfiguration.

4. Knowledge of the candidate

A few chapters of the dissertation specifically show that the candidate is well read in the discipline of information and communication technology.

Combinatorial optimization is covered in chapter 2.4 (Algorithm Selection Problem); and then in precise formulations of the considered problems (e.g., Chapters 3.2, 4.3, 6.1); as well as the usage of classic methods to solve them: ILP formulations (e.g., Chapter 3.3, 3.4, 6.3.2), modern heuristics (e.g., Chapter 6.3.3, or 4.3, particularly well written), or meta-heuristics (e.g., Chapter 6.3.5).
Parallel and distributed programming is covered in Chapter 6.6.
Elements of **machine learning** and **statistics** are covered in Chapter 5.3.
The presentation of experimental results shows that the candidate has adequate knowledge of **descriptive statistics.**

The thesis has 106 references, which I find sufficient, especially since the problem studied in the thesis is new and the wider problem area does not belong to the core of algorithmics / information technology.

The thesis relies on many numerical experiments that required the candidate to actually implement the proposed algorithms and run them in an HPC center (Chapter 6), which proves that the candidate has considerable hands-on programming experience.

5. Other remarks

The following remarks concern the presentation of the thesis, which, in general, is easy to read and to follow, but, nevertheless, has minor issues.

- The thesis is written in correct technical English. Occasional grammatical errors (p.8: “Berths they play important managerial role”, p. 69: “exemplary instance”) do not compromise understanding of the text. As a general remark, I would suggest using passive voice less often, especially in the description of experiments: it is the author who performed these.
- Introduction is, at times, imprecise. Container shipping is used not only to move consumer goods, but also (perhaps principally) to move semi-finished goods in the global supply chain. Demand for fast fashion products is not, unfortunately, well predictable, leading to even more waste.
- As a landsman, I missed a high-level, intuitive description of what a quay and a berth is.
- It is also unclear what is the AIS position of the ship in, e.g., Fig 1.1 (a) and (b), especially for large ships. Do all ships have their AIS in the same place (e.g., on a bridge)? If not, this could result in additional noise in the data (with one ship reporting the position of its bow, while another one of its stern).
- Besides histograms in Fig 1.1, I’d appreciate having some quantitative measure of goodness of fit between reported ship positions and the official berth placements.
- Chapter 1.4 has Chapters 4 and 5 swapped.

- The problem (and the notation) of Chapter 4 should be explicitly compared with the problem in Chapter 3.
- In Figure 4.5, I'd suggest to use two Y axes, rather than clump the number of algorithms and the cost on the single Y axis; and to convert X axis ticks to meaningful units (10s, 100s, etc.).
- Ships' lengths in Figure 5.1 are not to scale with Table 5.2.
- Table 5.5 is difficult to decrypt; I'd suggest adding three sub-columns.
- Figures with sub-figures (e.g., 6.3) would be more readable with sub-captions (e.g.: "a) Le Havre, intensity 50").
- Ref 89, containing authors' descriptions of the datasets, has no URL.

6. Conclusion

Taking into account what I have presented above and the requirements imposed by Article 187 of the *Act of 20 July 2018 - The Law on Higher Education and Science* (with amendments)[1], my evaluation of the dissertation according to the three basic criteria is the following:

A. Does the dissertation present an original solution to a scientific problem? (the selected option is marked with **X**)

X				
Definitely YES	*Rather yes*	*Hard to say*	*Rather no*	*Definitely NO*

B. After reading the dissertation, would you agree that the candidate has general theoretical knowledge and understanding of the discipline of **Information and Communication Technology**, and particularly the area of **....**?

X				
Definitely YES	*Rather yes*	*Hard to say*	*Rather no*	*Definitely NO*

C. Does the dissertation support the claim that the candidate is able to conduct scientific work?

X				
Definitely YES	*Rather yes*	*Hard to say*	*Rather no*	*Definitely NO*

Signature

[1] http://isap.sejm.gov.pl/isap.nsf/DocDetails.xsp?id=WDU20190000276