Prof. Dr. Florian Jaehn
Management Science and Operations Research
Helmut Schmidt University Hamburg, Germany
florian.jaehn@hsu-hh.de

January 8, 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 dissertation considers some port logistics problems and focuses on modeling and solving such problems from a discrete optimization perspective. More specifically, after an introduction and an intensive literature review, four selected port logistics problems are considered, which can be summarized as follows. First, the Deterministic Quay Partitioning Problem (DQPP) is analyzed. This is a tactical problem of partitioning the quay of a port into berths, each of which can accommodate at most one or two ships at a time (two problem variants). Second, the Berth Allocation Problem (BAP) is considered. This is an operational problem in which arriving ships are assigned to berths, given that berths can only accommodate ships of a certain size and only one (or two) at a time. If no berth is available for a ship, the ship must wait, which should be avoided as much as possible (minimizing the sum of weighted completion times). Third, historical vessel data is statistically analyzed to generate realistic test instances that can be used to evaluate the performance of the algorithms on port problems. Fourth, the quay partitioning problem is extended by stochastic input data, i.e., uncertainty is considered. This problem combines the previous problems, as the DQPP is the basis, the historical data from the second problem is used, and the BAP is used to evaluate the stochastic QPP.

The problems considered are certainly scientific and the research question (page 10) is well developed, although is very general and not as precise as it would be required in the social sciences, for example. Certainly, the topic is of high practical relevance. Decisions in port logistics often have a huge economic impact and therefore need to be taken carefully. In particular, the problems considered here are very relevant because lay days in ports are very costly and are often limited by scheduled arrival and departure times that are enforced by the tides. Thus, the research question studied in this thesis considers theoretical challenges that are relevant to practice.

2. Contribution

The most relevant contribution of this thesis can be found in Chapters 3 - 6. However, it should be mentioned that the literature review presented in Chapter 2 seems to be exhaustive and provides a good introduction to the topic. Let us consider each chapter separately.

Chapter 2 formalizes the DQPP. Since BAP is a subproblem of DQPP, and BAP is known from the literature to be strongly NP-hard, DQPP is also strongly NP-hard. The author shows that the partitioning subproblem of DQPP, i.e., omitting BAP, remains NP-hard in the ordinary sense. A

further theoretical result is that in both problem variants (1in1, i.e., at most one vessel can be assigned to a berth at a time, and 2in1, i.e., at most two vessels can be assigned to a berth at a time) an equidistant partition of the quay can lead to an arbitrarily bad solution compared to an optimal solution. For both problem variants, a mixed-integer linear optimization problem is presented that defines the original problem. Both models are then tested in an extensive numerical study. The study is based on self-generated test instances and includes several aspects to transfer the theoretical results to practice.

Chapter 3 examines the BAP in detail and presents many algorithms for solving it. There are 12 sorting rules, which are then used by 5 greedy algorithms that need a sorting rule. So there are already 60 algorithms plus the super greedy algorithm, which is the best of the above. Their solutions are the basis for several metaheuristic improvement methods, all based on local search. All these algorithms are used in an algorithm selection problem (ASP). I.e., the author not only presents these algorithms for solving the BAP, but also presents two methods (Section 4.6) for selecting an algorithm portfolio for a given instance that is likely to provide good solutions for that instance.

In Chapter 4, historical data are analyzed using statistics to simulate port logistics. First, ship sizes are clustered. Second, ship processing times are considered. Since a linear regression did not give promising results, several distributions were tested to see how they describe the processing time for the ship size clusters (maximum likelihood estimation was used to determine the parameters). Finally, ship arrival times are estimated. All this knowledge is then used to generate test instances that can be tailored to specific ports.

The results of the previous three chapters are summarized in Chapter 5. Here, a stochastic quay partitioning problem is considered. This means that the deterministic version of the problem from Chapter 2 is generalized. The evaluation requires the solution of the BAP, which is performed using the algorithm selection from Chapter 3. The numerical study is performed using the results of Chapter 4. In addition, algorithms for the stochastic partitioning problem are presented, again mostly based on greedy strategies and local search methods. These algorithms are compared in a numerical study before the full simulation study is performed. The latter is done using the port of Le Havre as an example.

In summary, the contribution is twofold. First, there are theoretical novelties, namely the complexity results, properties and algorithms (especially the algorithm selection algorithm). Second, the transfer of the theoretical results through the numerical tests is a valuable contribution to practice.

3. Correctness

In general, the thesis seems to be on a good level and the results are correct and understandable. It is only when going into details that some questions remain. For example, the proposed objective function of weighted completion time comes ad hoc and no justification or discussion is given. Certainly, there are good arguments for using such an objective function (e.g., because it is often used in the scientific literature), but there are also many cases where it contradicts practice. Often, ship departure times are determined at a higher decision level and used as input for all operational port logistics problems. This is due, among other things, to tidal restrictions when entering and leaving the port under consideration or subsequent ports. Similarly, the reasoning is somewhat incomplete in the following case. The author argues that an NP-hard problem can only be solved for small instances (page 23). This is

generally true for NP-hard problems in the strong sense, which is present here. However, Theorem 1 (page 29) only shows NP-hardness in the ordinary sense, which certainly has less practical impact.

In Chapter 3, the approach seems to be correct as well. A reader might wonder why the ASP, applied to a specific problem, only considers meta-strategies, but not a customized algorithm developed for the BAP. Surely, the BAP has a special structure that allows for such customized algorithms. Moreover, it is questionable whether the generated data represent real life situations and allow a reasonable ASP solution. Especially the fact that a uniform distribution is chosen seems questionable. For example, there may be algorithms that perform well or poorly in cases where there are very few large vessels but many small ones (a real scenario in ports with feeders from inland and open sea vessels).

Chapter 4 uses statistics and thus a very different methodology than the other parts of the dissertation. It has to be mentioned that this area is not the core expertise of the reviewer. The author relies on standard methods, which can be seen by the fact that the publicly available packages of the programming language R are mostly used. This is a viable approach and gives confidence in the results obtained. However, to criticize on a high level, some conclusions could have benefited from more statistical rigor. For example, on page 102, the author concludes that the longer the ships, the longer the return times. This is based on a positive correlation. Since all data are given, hypothesis testing could have easily verified this statement. Later, on page 137, hypothesis testing is indeed performed, but by misusing the p-value. To test whether pairs of algorithms are equal, the candidate first calculates p-values and then sets the significance level (which the candidate incorrectly calls the "confidence level"). However, the p-value must be set before the p-values are calculated. In addition, the candidate performs multiple hypothesis tests, but does not count for the family-wise error rate. A correction, e.g., using the Holm-Bonferroni method, would have been necessary.

4. Knowledge of the candidate

It can be confirmed that the candidate has a good general knowledge and understanding of the discipline of Information and Communication Technology. He shows a broad knowledge of algorithms, theoretical background, data handling and implementation. In general, the chapters are at a very good level, with some exceptions. Especially when it comes to neighboring fields, in this case statistics, the high quality of the other parts of the thesis cannot be maintained. Regarding formal aspects, the thesis is on a good, but not a very good level. There are still some typos in the text and errors in the list of references (e.g., in reference [60] only the first names of the authors are listed).

5. Other remarks

The thesis is extensive and the effort of the candidate is obvious. Congratulations to the author for a fine piece of work.

6. Conclusion

Taking into account what I have presented above and the requirements imposed by Article 187 of the A*ct 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:

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

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 Combinatorial Optimization and Operations Research?

[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*

Florian Jaehn

Signature