

Reviewer's opinion on the doctoral dissertation titled „Development of machine learning, process modeling and optimization algorithms for supporting Industry 4.0” by András Rácz-Szabó

The main focus in the PhD dissertation of András Rácz-Szabó is on the application of positioning data based optimization with the help of different data science techniques including data reconciliation, multi-layer networks, the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm. The author shows profound understanding of the presented problems and the methods described to solve these problems by outlining the theoretical framework and illustrating them with case studies. These real-world manufacturing case studies demonstrate the adaptability and effectiveness of the methodologies in improving resource management and operational efficiency.

The author presents the scientific theses in the dissertation throughout five chapters, totaling almost 100 pages including the introduction, conclusion, and the references. The first chapter serves as an introduction of the topic by giving an outlook to the state-of-the-art concepts and methods related to Industry 4.0 and an overview of the corresponding literature. The second chapter includes the five research questions and outlines the thesis. This chapter also lists the developed algorithms, methods and the related published papers by the author. The consequent chapters until (chapter seven) represent the main chapters of the dissertation: the author presents his independent work. In each of these chapters, a new method or concept is introduced, or previous methodologies are combined to solve complex real-world problems.

The third chapter provides a comprehensive overview of the application and development possibilities of a real-time locating system (RTLS) in manufacturing and investigates how positioning data can be utilized by data science techniques to optimize manufacturing processes. The author presents case studies where RTLS data could improve decision-making or operational efficiency. The use case described in section 3.5 shows the value of applying a RTLS system in production: the identification of the bottleneck in the production steps could directly improve the efficiency of manufacturing. The publicity of the theses of the third chapter is assured by the journal article published in “Sensors”.

The fourth chapter presents a data reconciliation approach aimed at enhancing the accuracy of position data by employing data reconciliation. In the demonstrated use case data is obtained by attaching two tags to forklifts, and by utilizing the physical distance between the tags the layout-induced noises in the production and logistics environment the data precision can be improved. The general method is explained in detail using formal matrix and vector notation and equations. The complexity of the notations makes understanding the ideas quite challenging, but the examples in the use case aid comprehension. In the reviewer's opinion it is worth considering reformulating the notation to make the train of thought more readable, which could also make validation of the equations easier. The publicity of the theses of the fourth chapter is assured by the journal article published in “MethodsX”.

The fifth chapter focuses on evaluating the efficiency of network flows by differentiating value-added and non-value-added activities using a multi-layer network approach. The theoretical background is explained in detail, and the illustrations of the definitions and concepts helps the reader of the dissertation. The author mentions that the problem can be solved by linear programming and provides an overview of the most widespread solvers available. The reviewer suggests considering the solver SCIP. The multi-layer network approach proves to be a valuable tool in the analysis of supply chains and urban transportation systems. This methodology is illustrated by taxi route optimization, where position data is analyzed to optimize resource flows. The publicity of the theses of the fifth chapter is assured by the journal article published in "Complexity 2024".

The sixth chapter combines the methods described in the previous chapter with a modified DBSCAN algorithm. The positional data of the five forklifts collected by an indoor positioning system (IPS) are reconciled to improve data accuracy, the modified DBSCAN algorithm forms clusters of the measurements, then value-added and non-value-added activities are differentiated, and a multi-layer network is constructed. The way the use case is presented is easily understandable for the reader and the figures illustrate the problem and the results well. The publicity of the theses of the sixth chapter is assured by the journal article published in "Emerging Science Journal".

The seventh chapter describes a method for optimizing micro-logistics and material handling processes using positional data. The author introduces the reader to reinforcement learning and Markov Decision Process (MDP) models and presents a case study applying the method on Automated Guided Vehicle (AGV) scheduling. Section 7.2 explains MDPs and presents the theoretical background on an abstract level. The use case described in section 7.3 is quite complex, but the author explains the steps of constructing the model step by step. The publicity of the theses of the seventh chapter is assured by the journal article published in "SAMI", and another related paper is already submitted.

The eighth chapter briefly summarizes the dissertation.

The author formulates five theses, each corresponding to the main chapters of dissertation described above. The independent scientific work of the author is supported by these theses, which obtained sufficient publicity: five scientific journal articles, and one paper already submitted.

The author is familiar with the relevant literature as indicated by a bibliography of 236 references. Many of the cited papers are published recently in international journals, which shows that the topics presented in the dissertation are of current interest.

Overall, the structure of the dissertation helps the reader to understand the complex concepts and methods presented. One point to highlight is the sixth chapter, which builds upon the previous chapters. The case studies in the chapters illustrate the theoretical frameworks.

The dissertation is generally well written. The design of the figures and tables aid easier comprehension. The author formalizes the presented methods mathematically: these formulae are precise enough, although sometimes quite complex.

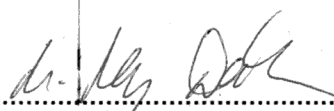
The dissertation shows the candidate's substantial effort. The presented results of the author are regarded suitable for public defense. Therefore, the reviewer recommends the

dissertation for acceptance. Considering the scientifically and practically significant results, recommending the conferment of the PhD title is anticipated.

The reviewer's questions related to the dissertation:

- Was there a case study conducted investigating the utilization of RTLS for layout planning? In the presented case were there hidden production steps in the temporary station?
- In the use case of the proposed data reconciliation approach could the physical dimensions of the forklifts be used beside the computed velocity as constraints? Suppose that the distance between two received positional data record (around the same time instance) belonging to two different forklifts contradicts the physical dimensions of those forklifts.
- What were the minimal and maximal efficiency of the networked process in the taxi case study?
- Was the reconstructed and the optimized cumulated idle time investigated in the forklift case study? What are the benefits of using DBSCAN instead of just classification based on Table 6.1?
- Was there any data of the AGVs available which would enable to construct a state transition matrix of probabilities?

Veszprém, January 2, 2025



.....
Dr. Dávid Csaba Mezőfi
Senior System Engineer
Continental Automotive Hungary