Response to Reviewer's Opinion - #Doctoral Dissertation

Reviewer: Dr. Dávid Földes

Title: Development of machine learning, process modeling and optimization algorithms for supporting Industry 4.0

Author: András Rácz-Szabó

Supervisor: Dr. János Abonyi, Dr. Tamás Ruppert

Dear Dr. Dávid Földes,

I would like to express my sincere gratitude for your thorough review and valuable feedback on my doctoral dissertation titled *"Development of machine learning, process modeling and optimization algorithms for supporting Industry 4.0 / Ipar 4.0 támogató gépi tanulási, folyamatmodellezési és optimalizációs algoritmusok fejlesztése."* I greatly appreciate the time and effort you have invested in providing such detailed and constructive feedback. Your comments and suggestions are extremely valuable and will contribute significantly to improving the quality of my research. I am particularly grateful for your insightful observations, which have helped me identify key areas for further refinement and development. I provide detailed responses to the questions in this letter. I hope my answers will be satisfactory.

Yours sincerely,

András Rácz-Szabó

Chapter 3 – Thesis 2: Elaboration of a method for improving the accuracy of indoor positioning data

1. *During the interpolation process, did the Candidate consider assumptions other than linear interpolation? Why was the linear interpolation chosen?*
2. *What potential development opportunities were identified through tracking the operational area of the vehicles?*

Thank you for your questions. Several options were indeed considered during the interpolation process. Since 4-5 position data points were recorded per second during active states, one possibility was to average the data points within each second. However, given the high density of recorded data, linear interpolation does not introduce inaccuracies and allows the use of actual measured data (after filtering out extreme outliers).

By tracking the operational areas of the vehicles, several development opportunities can be identified. These include optimizing route planning to reduce idle time and improve efficiency, identifying high-demand zones for more effective resource allocation, and uncovering vehicle usage patterns that could lead to a reduction in the total number of taxis required. Additionally, the data can be used to analyze and address bottlenecks or inefficiencies in specific areas and to develop dynamic scheduling systems that adapt to real-time demand and traffic conditions. These insights support more efficient logistics and resource allocation in similar operational environments.

Chapter 4 – Thesis 3: Design of an algorithm to reduce idle runs

1. *What other methods, besides linear programming, could be employed to address the transportation problem? Did the Candidate explore heuristic methods, such as genetic algorithms or ant colony optimization, or consider the application of classical approaches like Dijkstra's algorithm?*
2. *What was the justification for using Kernel-density estimation in the analysis? Could the Candidate have assumed a specific distribution (e.g., multimodal distribution corresponding to morning and afternoon peaks) for taxi usage patterns instead?*
3. *How can the resilience of the system be defined in this context? What resilience indicators were considered, and what were their results in the case study?*

Thank you for your questions. The thesis focuses on the multi-layer network-based approach and does not deal with the use of different optimization algorithms or compare their results. However, this is an excellent suggestion for possible future research directions.

Kernel density estimation was used as an approximation in the analysis. Exploring other possibilities, such as assuming a specific distribution, such as the multimodal distribution corresponding to morning and afternoon peaks, is an excellent suggestion for future research. The thesis also points out that variations in traffic intensity over different time periods were not considered in the analysis. Extending the research to account for these temporal variations could further improve the accuracy of the optimization. In addition, the optimization in the thesis relied on average zone-to-zone travel times, but incorporating more accurate location data and traffic dynamics would allow for further refinement of the process.

The case study focuses exclusively on reducing idle time; however, in a transportation example, multiple factors must be considered. The focus is on developing a methodology that can be applied to indoor position data, where the resilience issue is much simpler. In the case study, the resilience indicators were not evaluated quantitatively, but where the resilience issue is much simpler. In the case study, the resilience indicators were not evaluated quantitatively, but rather addressed through the robustness of the proposed

scheduling and optimization methods. Future research could include defining specific resilience metrics and evaluating them quantitatively in similar contexts.

Chapter 5 – Thesis 4: Development of a modified DBSCAN method for identifying activity types; and representing activity types on a novel multi level network graph

1. *How are different activities (value-added and non-value-added) currently estimated in the profession? Is the process manual or does it involve any automated methods? Are any observation-based measurements or standard times used in the estimation process?*
2. *How could the method be adapted for use with material handling equipment other than forklifts? What modifications would be necessary to apply the method to other material-handling equipment?*

Thank you for your questions. Typically, there is no fully developed or standardized process for identifying value-added and non-value-added activities, making this line of research particularly novel and valuable. Observation-based measures and standard times are often used in the estimation process, sometimes cross-referenced with data from Enterprise Resource Planning (ERP) systems. However, this is usually done at a relatively basic level. Challenges remain in the standardization of digitization, data visualization, and data collection processes, highlighting the need for further advances in this area.

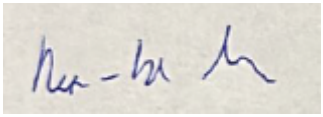
The DBSCAN algorithm can be easily modified and adapted for use with any type of material handling equipment. However, a framework must be established to define what constitutes value-added and non-value-added activities for the specific equipment. This framework could use data from enterprise resource planning (ERP) systems (e.g., if a piece of equipment is moving to retrieve an item, it can be classified as idle time, while storing an item can be considered value-added). Alternatively, data can come from position tracking sensors or other types of sensors that can provide objective information to define equipment states. These inputs ensure that the method is flexible and can be adapted to different material handling scenarios, enabling accurate classification and analysis.

Chapter 6 – Thesis 5: Development of an MDP-based model to enhance the efficiency of production systems

1. *What are the cost implications of implementing these developments, including hardware, software, and integration costs? What return on investment metrics can be associated with these advancements? Under what conditions would it be worthwhile to undertake such investments?*

Thank you for your questions. Implementing such developments involves significant costs, particularly in terms of integration into the production system. However, without these improvements, automated material handling without production downtime would not be feasible. In the case study, the identified issues with logistics, inventory strategies, and production scheduling were effectively addressed, eliminating the need for additional task scheduling optimization. ROI metrics could include savings from reduced manual labor, improved efficiency, and minimized errors. Such investments would be worthwhile in cases of high labor costs, large-scale operations, or where automation aligns with long-term strategic goals, such as implementing Industry 4.0. Ultimately, the decision should weigh upfront costs against potential long-term benefits, such as increased efficiency and reduced errors.

Veszprém, 2025.01.06.



András Rácz-Szabó