Proceedings of the 2019 Winter Simulation Conference
N. Mustafee, K.-H.G. Bae, S. Lazarova-Molnar, M. Rabe, C. Szabo, P. Haas, and Y-J. Son, eds.

KNOWLEDGE DISCOVERY IN SIMULATION DATA: AN INDUSTRIAL CASE STUDY ON INTRALOGISTICS IN AN AUTOMOTIVE PRODUCTION

Niclas Feldkamp
Soeren Bergmann
Steffen Strassburger

Technische Universität Ilmenau
Information Technology in Production and Logistics
P.O. Box 100 565
Ilmenau, 98684, GERMANY

Kristina Sokoll
Matthias Clausing

AUDI AG
I/PI-532
SimOffice
Ingolstadt, 85045, GERMANY

ABSTRACT

Discrete-event simulation is an established and popular technology for investigating the dynamic behavior of complex manufacturing and logistics systems. In addition to conventional simulation studies that focus on single model aspects answering project specific analysis questions, new developments in computational power, data processing, and availability of big data infrastructures enable the execution of broad scale simulation experimentation. In this work, we applied our developed methodology of knowledge discovery in simulation data onto an industrial case study of a production process from the AUDI AG. The main focus of interest in this case study was the routing for intralogistics material supply trains. We used broad scale experiment design in combination with data mining methods to investigate a very large number of possible routes. The goal was to find patterns and combinations for routes and associated factors, and to investigate how they affect the system performance.

1 INTRODUCTION

Simulation is a well-established tool for planning, operating and monitoring complex systems. Usually, simulation projects are conducted in order to account for specific project goals that are formulated beforehand (Law 2014). Besides single aspect analysis and scenario evaluation, hidden relations and effects outside of a defined project scope may exist that the system expert did not think about before. In order to detect hidden relations and get in-depth insight into the system, we need a broad coverage of possible model behavior. Recent advancements in computing power enable to conduct large and broad scale simulation ensembles (Theodoropoulos 2015). The data farming methodology promotes the usage of intelligent experiment design in combination with high performance computing to achieve this (Horne et al. 2005; Sanchez 2014). In order to analyze these large amounts of farmed simulation data efficiently, we developed a concept called knowledge discovery in simulation data on top of the data farming methodology, that combines data mining and interactive visual analysis (Feldkamp et al. 2015).

2 CASE STUDY

This case study was carried out on a simulation model of a manufacturing system at Audi in Ingolstadt. Primary focus of this model were the intralogistics processes regarding the material supply for the manufacturing lines. The main application of our concept was to convey interesting patterns and relations between all modeled input and output parameters. In this model, especially the routes for the milk runs that ensure the delivery of material should be investigated. The routing of those trains includes the number of routes, the number of trains, the frequency of tours and the allocation of stations to each route. The overall objective was the application of the knowledge discovery method to uncover previously unknown interactions, relations, problems and solutions for decision support for future logistics planning projects. To guideline the data analysis process, some key questions were derived that are shown below:

- Which routes are particularly performant? Do these routes exhibit specific similarities or patterns?
- What distinguishes good routes from bad routes?
- How are these patterns related to the number of trains, and how does the number of trains affect the system performance? Which influence has the number of trailers per train?
- Which stations are particularly critical? Do critical groups or areas of stations exist? Are there significant differences in the material flow between the top and the main floor?

In order to accomplish the study, we used a combination of different tools including Java for the generation of the experiment designs, Apache Hadoop and Apache Spark for persisting and processing large amounts of data, and Matlab and R for data mining and statistical analyses.

3 RESULTS

The set of milk run routes for each experiment is a complex factor. We calculated 10^{52} different possible combinations for the routes, so that a full factorial experiment design was not feasible. We therefore developed a method for generating a smart experimental design for this factor based on sampling and graph optimization. This method reduced the number of experiments to 100,000 while still maintaining desirable features of a good experiment design that enables a valid and bias-free coverage of the system behavior. Experiments were conducted on a computing cluster with a runtime of almost seven days. We collected more than 20 key parameters for each route on each experiment during the simulation runs which resulted in more than 16 GBytes of raw data. For the analysis of data we used interactive visualizations which were supported by the output of data mining algorithms. The data mining operations included clustering, classification trees, and frequent pattern mining. We successfully generate knowledge and conclusions on how the composition of routes affects certain parts of the system. We also were able to give recommendations which patterns in the composition of routes are advantageous and which should be avoided. For potential future work, the investigation could be extended towards a robustness analysis (Feldkamp et al. 2017) regarding the behavior of the system against variance of external noise factors.

REFERENCES

Feldkamp, N., S. Bergmann, and S. Strassburger. 2015. “Knowledge Discovery in Manufacturing Simulations”. In *Proceedings of the 3rd ACM SIGSIM Conference on Principles of Advanced Discrete Simulation,* edited by S. J. E. Taylor, N. Mustafee and Y.-J. Son, 3–12. New York, New York: Association for Computing Machinery.

Feldkamp, N., S. Bergmann, S. Strassburger, and T. Schulze. 2017. “Knowledge Discovery and Robustness Analysis in Manufacturing Simulations”. In *Proceedings of the 2017 Winter Simulation Conference,* edited by V. Chan, A. D’Ambrogio, G. Zacharewicz and N. Mustafee, 3952–3963. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Horne, G. E., and T. E. Meyer. 2005. “Data Farming: Discovering Surprise”. In *Proceedings of the 2005 Winter Simulation Conference,* edited by M. E. Kuhl, N. M. Steiger, F. B. Armstrong and J. A. Joines, 1082–1087. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Law, A. M. 2014. *Simulation Modeling and Analysis,* 5th edn. New York, New York: McGraw Hill Book Co.

Sanchez, S. M. 2014. “Simulation Experiments: Better Data, Not Just Big Data”. In *Proceedings of the 2014 Winter Simulation Conference,* edited by A. Tolk, S. D. Diallo, I. O. Ryzhov, L. Yilmaz, S. Buckley and J. A. Miller, 805–816. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Theodoropoulos, G. 2015. “Simulation in the Era of Big Data: Trends and Challenges”. In *Proceedings of the 3rd ACM SIGSIM Conference on Principles of Advanced Discrete Simulation,* edited by S. J. E. Taylor, N. Mustafee and Y.-J. Son, 1. New York, New York: Association for Computing Machinery.