Proceedings of the 2021 Winter Simulation Conference
S. Kim, B. Feng, K. Smith, S. Masoud, Z. Zheng, C. Szabo and M. Loper, eds.

AUTOMATION IN THE PROCESS OF KNOWLEDGE DISCOVERY IN SIMULATION DATA

Jonas Genath

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

ABSTRACT

In contrast to classical simulation studies, the method of knowledge discovery in simulation data uses a simulation model as a data generator (data farming). Subsequently using data mining methods, hidden, previously unknown and potentially useful cause-effect relationships can be uncovered. So far, however, there is a lack of support and automatization tools for non-experts or novices in knowledge discovery in simulation data, which leads to a more difficult use in industrial applications and prevents a broader utilization. In this work, we propose a concept which provides an approach for automating and supporting knowledge discovery in simulation data.

1 INTRODUCTION

Discrete event simulation is an established tool for planning, evaluating, and investigating the dynamic behavior of manufacturing and logistics systems. In order to go beyond questions of classical simulation studies, it is possible to use the simulation model for data farming (Sanchez and Sanchez 2017). This involves an extensive experiment design that covers the system behavior as completely as possible. Feldkamp et al. (2020) developed and validated the process of knowledge discovery in simulation data, which extends data farming with data mining methods for data analysis and various tools for interactive data exploration. Attempts to transfer the process into operational practice have so far failed due to a lack of experience and support for users. They often do not know which parameters and parameter values they should use or how the results should be visualized. In this work, we propose a concept which provides an approach for automating and supporting the aforementioned process, so that even inexperienced users can take advantage of the resulting benefits in practice.

2 TOWARDS AUTOMATION IN THE PROCESS OF KNOWLEDGE DISCOVERY IN SIMULATION DATA

Figure 1 shows the proposed concept. The left side of the figure shows the standard data farming process, which includes the design of experiments, the simulation and the resulting simulation data. Normally, this data is then evaluated by the user according to the iterative knowledge discovery in simulation data process (blue area) using data mining and visual analytics to generate knowledge about the simulation model. To make this process more accessible to novice users, we propose an automation system that uses various methods to provide automation and support for the user. These methods include meta learning, automated data mining, and enhanced or rule-based selection of visualizations. The input data of these methods is based on meta data and information, which is either automatically extracted from the simulation data itself using a meta data generator or queried by the user via an understandable interface. In a recent paper, we proposed a concept that identifies potential statistical anomalies in result parameters for subsequent data

mining methods using meta learning (Genath et al. 2021). Since the characterization of the result data often poses great difficulties for inexperienced users in practice, this can provide a much easier entry point for the following analyses. With automated data mining, we aim to reduce the time and iterations required for the algorithms used by means of hyperparameter tuning. In addition, the required knowledge about the parameters used in the data mining methods can be reduced. Furthermore, we try to deduce rules for improved visualizations to reduce errors in interpretation and to facilitate the recognition of information. Concluding, the user is asked if specific results of the simulation study should be stored as new meta data to improve the support provided by the automation system.

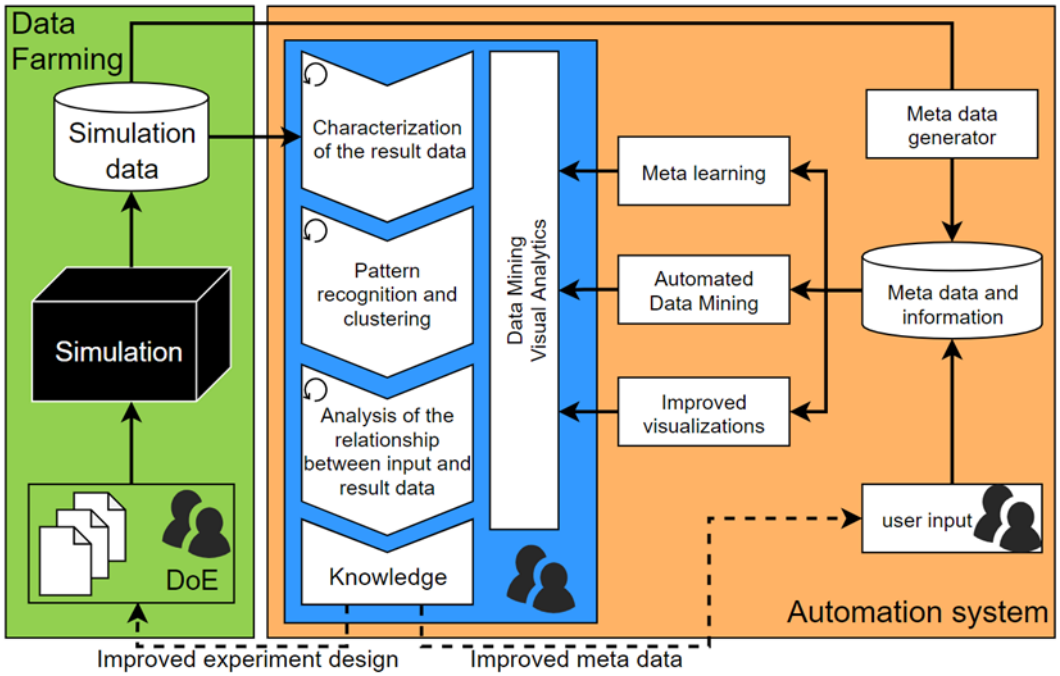

Figure 1: Conceptualization for automating and supporting the process of knowledge discovery in simulation data..

3 SUMMARY

In summary, this work proposes an approach to automate and support the process of knowledge discovery in simulation data using various methods. These methods are mainly from the field of machine learning but also include rule-based decision support for visualizations. A deeper insight into the proposed concept and the particular methods will be provided at WSC 2021.

REFERENCES

Feldkamp, N., S. Bergmann, and S. Strassburger. 2020. "Knowledge Discovery in Simulation Data". *ACM Transactions on Modeling and Computer Simulation* 30(4):1–25.

Genath, J., S. Bergmann, N. Feldkamp, and S. Straßburger. 2021. "Automatisierung im Prozess der Wissensentdeckung in Simulationsdaten – Charakterisierung der Ergebnisdaten". In *Simulation in Produktion und Logistik 2021*, edited by J. Franke and P. Schuderer, 177–186. Göttingen: Cuvillier Verlag.

Sanchez, S., and P. J. Sanchez. 2017. "Better Big Data via Data Farming Experiments". In *Advances in Modeling and Simulation*: *Seminal Research from 50 Years of Winter Simulation Conferences*, edited by A. Tolk, J. Fowler, G. Shao, and E. Yücesan, 159–179. 1st ed. Springer International Publishing.