Proceedings of the 2017 Winter Simulation Conference
W. K. V. Chan, A. D'Ambrogio, G. Zacharewicz, N. Mustafee, G. Wainer, and E. Page, eds.

KNOWLEDGE DISCOVERY IN SIMULATION DATA – A CASE STUDY FOR A BACKHOE ASSEMBLY LINE

Niclas Feldkamp
Soeren Bergmann
Steffen Strassburger

Department for Industrial Information Systems
Ilmenau University of Technology
P.O. Box 100 565
98684 Ilmenau, GERMANY

Thomas Schulze

School of Computer Science
Otto-von-Guericke-University Magdeburg
Universitätsplatz 2
39106 Magdeburg, GERMANY

Praneeth Akondi

John Deere Technology Center
Tower-14, Magarpatta City
Pune, 411013, INDIA

Marco Lemessi

John Deere GmbH & Co. KG
Mannheim Regional Center
John-Deere-Straße 70
Mannheim, 68163, GERMANY

ABSTRACT

Discrete event simulation is an established and popular technology for investigating the dynamic behavior of complex manufacturing and logistics systems. Besides conventional simulation studies that focus on single model aspects answering project specific analysis questions, new methods of broad scale experiment design and system analysis emerge alongside new developments of computational power and data processing. This enables to investigate the bandwidth of possible system behavior in a more in-depth way. In this work we applied our previously developed methodology of knowledge discovery in simulation data onto an industrial case study for a backhoe loader manufacturing facility.

1 INTRODUCTION

Discrete event simulation is a broadly accepted tool for planning, operating, and evaluating manufacturing and logistics systems. Traditional simulation studies usually focus on a specifically outlined project scope that leads to definite and quantifiable analysis questions. Simulation studies are frequently conducted in a way that experimentation is done by varying distinct and preselected parameters that are influential to the project goal as supposed by the simulation expert (Law 2014). Besides single aspect analysis and scenario evaluation, hidden relations and effects outside of a defined project scope may exist in a complex system that manual experimentation might never discover.

We therefore developed a method called knowledge discovery in simulation data that aims to uncover those relations and effects by means of broad scale experimentation, data mining, and visually aided analysis (Feldkamp et al. 2015a). This approach goes beyond what could be learnt from analyzing real data collected from the physical system, as our experiment design can cover system variants and parameter combinations that the real system might never experience.

Recent advancements in computing power enable to conduct large and broad scale simulation ensembles (Matkovic et al. 2015; Theodoropoulos 2015). The concept of knowledge discovery in simulation data consists of three components, namely *data farming*, *data mining,* and *visual analytics*. Data Farming describes a methodology of using a simulation model as a data generator to maximize data yield and therefore information gain by covering a large bandwidth of possible input parameter value combinations or system behavior, respectively. This is achieved by applying intelligent and efficient experimental design alongside performant and distributed computing of simulation experiments (Horne and Meyer 2010; Sanchez 2014).

Our approach is driven by applying data mining algorithms on large amounts of simulation output data combined with an interactive, visual aided analysis approach (Feldkamp et al. 2015b). Combining data mining and visualizations is based on the research area called visual analytics that aims to combine the strengths of machines, e.g., for processing huge amounts of data, with those of humans, e.g., for pattern recognition and drawing conclusions (Keim et al. 2008).

2 CASE STUDY MODEL

The investigations in this case study were carried out on a simulation model of a manufacturing facility for backhoe loaders. The simulation model was created in the design phase of this manufacturing line. Our investigations assumed this model to be valid within the parameter range of input parameters discussed below.

The facility has a main line separated in five areas with each area having multiple stations and also at least one sub-line for component assembly. Input parameters of the model are buffer capacity, the mixture of product types, worker efficiency (accepted task speed), and worker skill flexibility that describes the ability of workers of how many different tasks on various workstations they can perform, whereby each parameter can be individually set for each zone.

Special challenges within this case study consisted in modelling product mixes and implementing different worker skill flexibility levels in the model.

The main application of the knowledge discovery in simulation data concept is to convey interesting patterns and relations between all modeled input and output parameters. In this model, the relation of input parameters to throughput and worker utilization was of superior interest. To guideline the analysis process, we derived multiple analysis questions from that:

- Which input factors are most influential to worker utilization? What is the main source of variation for worker utilization?
- Which area is most affected by changes in the input parameter values? What are possible bottlenecks for throughput limitation?
- How robust is the line against variations in the product mix?

From these three coarse investigation questions, more specific analysis questions can be derived if analysis reveals anything else that yields potential for in-depth analysis coverage.

3 RESULTS

Experimental setup contained 24 different input parameters in a *nearly orthogonal latin hypercube design* containing 512 design points crossed with 100 different designs points for the product mix resulting in 51.200 total experiments. Crossed designs are important when robustness of configurations against variation in the product mix needs to be evaluated. Results of the analysis showed reasonable and helpful insights about the system. The main finding is that, although connected to each other, performance values are very diverse among different areas.

As mentioned before, further aspects of the investigation have been the relations between different performance values, the relations and effects of corresponding input factors to output performance as well as robustness analysis of system configurations. Processing of the simulation data has been carried out by using data mining, for example clustering algorithms. Processing those findings in an interactive, visually aided analysis revealed interesting insights about the system. For example, not only did we discover that the effect of worker efficiency on performance was very dissimilar on different areas of the assembly, we also recognized a strong interaction between worker efficiency and skill flexibility in some of the areas.

For a certain area, we were able to show that an increase of the skill level of workers had a superior effect on both the overall system performance and on the worker utilization in the entire system. For a different area, we were able to show that increasing worker efficiency (in contrast to worker flexibility) had a positive effect on system performance.

In summary, this case study shows how data farming and knowledge discovery in simulation data can be applied on real world applications and simulation models in order to create and enhance decision support for planning and managing manufacturing systems.

Concerning future work, the investigations with this model could be enhanced towards a robustness analysis of the system (Feldkamp et al. 2017).

REFERENCES

Feldkamp, N., S. Bergmann, and S. Strassburger. 2015a. "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, NY, USA: ACM.

Feldkamp, N., S. Bergmann, and S. Strassburger. 2015b. "Visual Analytics of Manufacturing Simulation Data." In *Proceedings of the 2015 Winter Simulation Conference,* edited by L. Yilmaz, W. K. V. Chan, I. Moon, T. M. K. Roeder, C. Macal and M. D. Rossetti, 779–790. Piscataway, N.J.: IEEE Inc.

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 W. K. V. Chan, A. D'Ambrogio, G. Zacharewicz, N. Mustafee, G. Wainer, and E. Page (Accepted for Publication).

Horne, G. E., and T. Meyer. 2010. "Data farming and defense applications." In *MODSIM World Conference and Expo,* edited by R. Armstrong, J. McNamara and T. E. Pinelli, 74–82. Hampton, VA: Langley Research Center.

Keim, D. A., F. Mansmann, J. Schneidewind, J. Thomas, and H. Ziegler. 2008. "Visual Analytics: Scope and Challenges." In *Visual Data Mining: Theory, Techniques and Tools for Visual Analytics,* edited by S. Simoff, M. H. Boehlen and A. Mazeika. Berlin, Heidelberg: Springer.

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

Matkovic, K., D. Gracanin, M. Jelović, and H. Hauser. 2015. "Interactive Visual Analysis of Large Simulation Ensembles." In *Proceedings of the 2015 Winter Simulation Conference,* edited by L. Yilmaz, W. K. V. Chan, I. Moon, T. M. K. Roeder, C. Macal and M. D. Rossetti, 517–528. Piscataway, N.J.: IEEE Inc.

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, N.J.: IEEE 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, NY, USA: ACM.