DOI: https://doi.org/10.46354/i3m.2019.mas.008

INTEGRATION OF PROCESS MINING TECHNIQUES IN SIMULATION RESULTS ANALYSIS

Irīna Šitova[a], Jeļena Pečerska[b]

[(a), (b)] Riga Technical University, Riga, Latvia

[(a)]irina.sitova@rtu.lv, [(b)]jelena.pecerska@rtu.lv

ABSTRACT

The research is carried out in the area of analysis of simulation results. The aim of this research is to explore the applicability of process mining techniques, and to introduce the process mining techniques integration into results analysis of discrete-event system simulations. As soon as the dynamic discrete-event system simulation (DESS) is based on events list or calendar, most of simulators provide the events lists. These events lists are interpreted as event logs in this research, and are used for process mining. The information from the events list is analysed to extract process-related information and perform in-depth process analysis. Event log analysis verified applicability of the proposed approach. Based on the results of this research, it can be concluded that process mining techniques in simulation results analysis provide a possibility to reveal new knowledge about the performance of the system, and to find the parameter values providing the advisable performance.

Keywords: discrete-event system simulation, process mining, queueing system, simulation results analysis

1. INTRODUCTION

Simulation is one of operations research approaches that is realized in form of software tools. A simulation is the imitation of a complex system that gives a possibility to display and describe the behaviour of a system in a detailed way (Pidd 2004). In the present paper, the discussion is limited to the DESS approach, further referenced as “simulation”. Simulations are most effectively used for the analysis of dynamics of complex artificial material systems. This choice is a reasonable one, as soon as this approach aims to investigate business, manufacturing, service and other processes. Researchers simulate processes to gain insight into the operation of systems. In the process of simulation, random factors influencing the system are taken into account, as well as their changes over time. Researchers obtain information about the system under consideration after experiments with a verified, calibrated and validated model (Law and Kelton 2000).

Technically the approach is implemented with the software tools; most of them simulate artificial event logs. The event log in form of events lists or events calendars is a mean of model entity management. Simulation is an important tool to explain how process performance indicators react in the face of controllable factors and environmental factors (Banks et al. 2010). Any changes of the system state during the process execution are recorded in the simulation events lists. Simulation results after running models of this type make it possible to obtain estimates of various performance measures: productivity and throughput measures, resource utility measures, and service level measures (Merkurjevs et al. 2008). The obtained simulation results are used for understanding the behaviour of the system, to formulate forecasts, to compare alternatives or to solve the optimization tasks of system parameters. Events lists usually are not analysed as a simulation outcome.

Process mining is described by its authors as a tool to extract non-trivial and useful information from process execution logs. Process mining uses event data to extract the information. This discipline is built on process model-driven approaches and data mining, and is used to support process improvements (van der Aalst 2011).

Thus the correspondence with the process mining tool in the area of the research, as well as in using natural or simulated event logs, and implementation goals – process improvement – is detected. The event logs in process mining are used as the source of information about the process and may be aimed on simulation models discovery. In DESS the models are based on the information about the process, models are realised through simulation of the events, organized as events lists, and produce simulation results.

The paper proposes to use a combined approach to the analysis of the outputs, both simulation results and events lists, of the simulation model, as shown in Figure 1.

The findings in the area, authors’ previous work and the basics of the proposed approach are provided in the following sections.

2. RELATED WORK

There are a considerable number of articles and projects concerning application of process mining techniques for analysis of various process types.

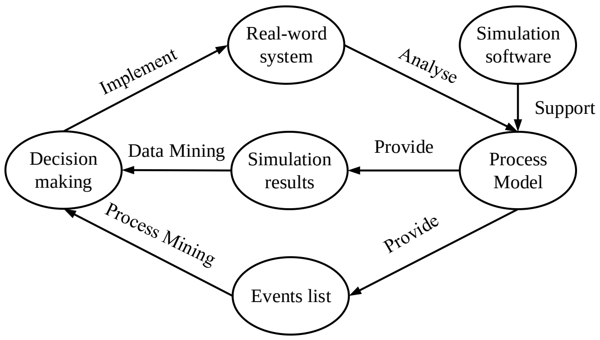

Figure 1: Process Mining for Simulation Outcomes Analysis

Some of them deal with efficient resource utilization in manufacturing processes, business process organization in warehousing, insights in human behaviour from realization of smart house processes and business rules discovery from manufacturing processes. Some of the most characteristic publications are considered in this section.

2.1. Analysis of Resource Behaviour

Authors in (Nakatumba and van der Aalst 2009) developed an approach that applies process mining techniques to provide detailed insights analysing resource behaviour. They investigate the relationship between workload and services time using linear regression analysis on historic data from Process-Aware Information Systems (PAISs). Researchers note a strong connection between workload and performance of workers which is confirmation of the "Yerkes-Dodson Law of Arousal". The authors used real-life logs to validate their approach using implemented new plug-in in ProM. The authors showed how better resource description helps to create more realistic and precise simulation models and provides better work allocation.

2.2. Analysis of Warehouse

Haav and Kalja in (Haav and Kalja 2014) presented the application of automatic process mining for solving a real warehousing problem. The goal of the research was to map a specific business problem to the relevant process mining problem definition, perform the analysis and translate the technical results into business insights for the evaluation. The goal was reached on the basis of the case study results. The case study was performed on data mined from the Enterprise Resource Planning (ERP) system of a logistics company, rendering warehousing services. The authors applied the developed methodology for extracting, transforming and loading the data using ProM software. Two process mining algorithms – Alpha and HeuristicMiner were used for data analysis. Thanks to the experiment the company, which provided data, received insights that helped to adjust the strategy and operations of logistic enterprise.

2.3. Analysis of Daily Life

The method described in (Tax et al. 2018) provided extracting insights in (un)healthy living habits from smart home environment data. To deal with overgeneralizing process models abstraction of the events to a higher-level interpretation was made in order to reach better understanding for analytics. For abstraction from sensor-level to human activity purpose framework using supervised learning methods that is based on the XES IEEE standard for event logs were developed. This enables direct analysis of process models showing real human behaviour. In confirmation of the correctness of the methodology the authors demonstrate the added value on the example of the case study.

2.4. Analysis of Semiconductor Manufacturing

Another research in the area (Khemiri et al. 2018) showed an example of use of data mining algorithm for discovery of business rules in semiconductor manufacturing process inspired by process mining ideas. Authors explained the importance of the topic with the lack of knowledge of the operating rules in industrial companies, which leads to not the best possible production performance. The authors, using the case study results, proved that the developed method helps to gain knowledge about the existing rules, to detect problems in the system by using process modelling and serve as a base to knowledge capitalization.

2.5. Current Findings and Future Development

Comparing the approaches from the researches described above, one can conclude that the process mining is aimed at:

- Data-based facilitation of process models creation or improvement (Nakatumba and van der Aalst 2009);
- Using ProM tool for data filtering, analysis and visualization (Haav and Kalja 2014);
- Extracting event log from real life process (Tax et al. 2018);
- Revealing business rules from the output decision trees (Khemiri et al. 2018).

The goal of process mining is to use event data to extract process-related information (van der Aalst 2011). The DESS approach deals with the simulated processes that are producing event data. The objective of creation and use of simulation models is in-depth analysis of the simulated processes and obtaining knowledge about these processes. The generic goal of the simulation studies is transformation of obtained knowledge into actions to improve the performance indicators or structure of the investigated processes. There is a clear similarity of the goals of process mining and DESS, as well as the common subject of the research – processes – and model-oriented nature of process mining. Thus a logical motivation arises for adaptation of process mining techniques for simulation results analysis and simulation-based process analysis for process mining purposes.

Taking these similarities into account and expecting potential findings, the authors of the present article in the fourth section propose the approach to simulation results analysis that is based on process mining techniques.

3. THE AUTHORS' PREVIOUS WORK

The present research is a logical continuation of authors' previous research (Šitova and Pečerska 2018). Regarding this it is shortly described below.

3.1. The Developed Approach

The authors (Šitova and Pečerska 2018) provided a review of approaches in the area of analysis of simulation results by using data mining techniques. The goal of the research was to explore the applicability of data mining techniques in the area of simulation result analysis and to introduce an application scheme of data mining techniques for the analysis of simulation results. After analysing the reviewed scientific publications (Brady and Bowden 2001, Brady and Yelling 2005, Painter et al. 2006, Feldkamp et al. 2015, Kibira et al. 2015, Feldkamp et al. 2016), as well as several books (Dunham 2003, Witten and Frank 2005, Han and Kamber 2006, Cassandras and Lafortune 2008) devoted to simulation, data mining and integration of these technologies, the authors had developed their approach. The approach includes the analysis of simulation results using several data mining techniques in the proposed order. As a result of the theoretical study, a two-stage approach was formulated, combining the fundamental principles of data farming and knowledge discovery. In the developed approach, each of the used technologies has its own purpose. Data farming, the purpose of which is to get results from a simulation model that corresponds to some real process. Knowledge discovery, the purpose of which is to find patterns in a simulation model state variable behaviour, in order to uncover knowledge which would otherwise be hidden.

A variety of data mining techniques, including correlation analysis, clustering and several visualization mechanisms of results, were used for the knowledge discovery. The developed approach was applied to the analysis of experimental data of a simple DESS model.

The authors hypothesized, tested and confirmed that data mining techniques may provide a better interpretation of simulation output as well as visualization of outputs. Also, the authors showed that data mining reveals not only trivial knowledge from simulation output. As a result of data mining techniques application in simulation results analysis, the knowledge and decision rules were obtained from simulation results coupled with the relevant visualization, which increases the simulation efficiency. It provides a more versatile output analysis and deals with a potentially huge amount of simulation output data.

The overall scheme of the developed approach application, which reflects data farming and knowledge discovery phases, is shown in Figure 2.

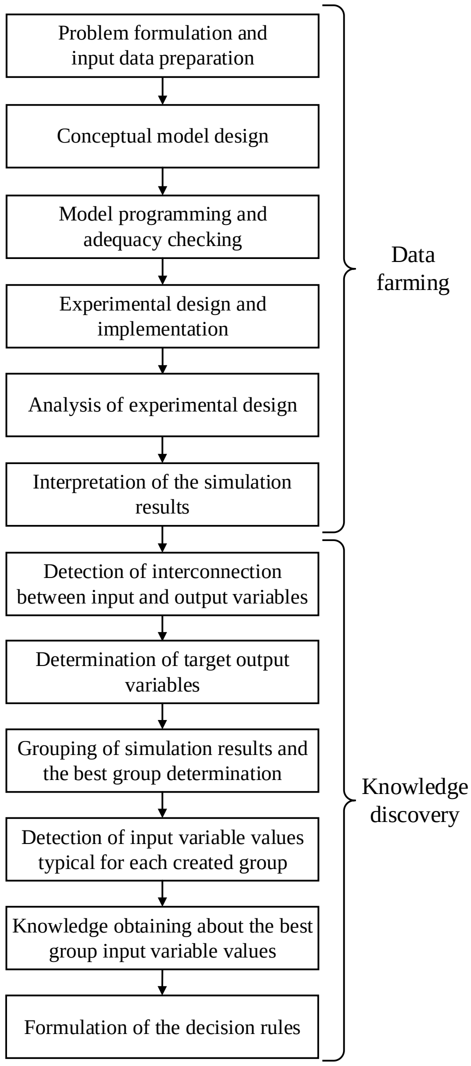

Figure 2: The Developed Approach Application Scheme

3.2. Case Study: Queueing System Simulation and Results Analysis

To achieve the goal of the research (Šitova and Pečerska 2018), the process was consistently implemented as follows. First comes data farming realization: the simulation model of the system under consideration was constructed, simulation experiments were designed, replicated and results were obtained. Second – knowledge discovery realization: the simulation results were analysed using data mining techniques. The domain of acceptable values of experimental factors was revealed and visualised.

To obtain data for mining, 28 model experiment scenarios had been developed with the realistic experimental factor values. Each of the scenarios had been replicated 25 times, providing 700 experiment results sets.

The general scheme of knowledge discovery consists of three steps:

- Correlation analysis of a relationship between input and output variables;
- Clustering algorithm application against target output variables;
- Results visualization with box charts, histograms, radar charts.

Based on the experimental results, the decision principles for the case study problem have been formulated for detecting the input variable values for experimental factors that are specific for each cluster and in particular for the best cluster. The values for the input variables that are experimental factors of the best cluster were defined. Thus the main goal of the case study – the best performance of a queueing system – has been achieved.
In spite of satisfactory results and findings in (Šitova and Pečerska 2018) authors noticed that a whole layer of "additional" results – the events lists data – is not taken into account during the process of knowledge discovery. Therefore, in the present study, the authors decided to conduct knowledge discovery based on data from events lists while data farming remains unchanged. Due to this, the concept of the developed approach has changed as shown in Figure 3. The relevant approach is described in fourth section, as well as the case study is shown in the fifth section.

4. PROCESS MINING IN THE SIMULATION RESULTS ANALYSIS

The present research is a continuation of the previous research described in third section (Šitova and Pečerska 2018). The approach proposed in this research still consists of two stages – data farming and knowledge discovery. However, after analysing the scientific papers (Nakatumba and van der Aalst 2009, Haav and Kalja 2014, Tax et al. 2018, Khemiri et al. 2018, van der Aalst 2012) and books (van der Aalst 2011, van der Aalst 2016) dedicated to process mining and application of process mining techniques, the authors made changes in their approach. These changes affect knowledge discovery stage.
The proposed approach is based on the process mining discipline formulated by van der Aalst in (van der Aalst 2011). The authors consider that not using an additional source of information from DESS models, such as events list, is inconsistent. In this paper, the authors deal with the conventional DESS approach. In most cases, the simulation technique involves the creation and management of events lists to organize and track entities in the simulation model. The term "entity" is used here to designate a unit of traffic. Entities instigate and respond to events (Schriber et al. 2014). Technically, there are several types of events lists involved into the running the simulations. Further, the authors do not delve into the type of events list used; it is assumed that it is possible to use information as necessary. The event logs may be interpreted in a way that is conventional in DESS – a list of particular event notices, ordered by time of occurrence and aggregated if necessary. The typical structure of the DESS model events list is provided as an example in Table 1.

Table 1. The Fragment of the Events List Created by DESS Software

Simulation time	Entity's ID	Event type
…		
481.56	46	Arrival
481.56	46	End of Job on Gate
481.56	46	End of Job at Decision_0
481.56	46	Moved to Server_3 for service
482.45	47	Arrival
482.45	47	End of Job on Gate
482.45	47	End of Job at Decision_0
482.45	47	Moved to Server_4 for service
483.23	45	End of Job at Server_2
…		

While running the simulations, the events list is managed: events are created, sorted, executed and deleted according to the model structure and logic. For the purposes of the current research, the events lists provide the source for the event log data. Some trivial data saving, sorting and processing actions are not introduced in this paper. These actions are applied to transform the conventional events list into event log.
The concept of the events list-based approach is available in Figure 3. The stages of approach implementation include both data farming and knowledge discovery phases.

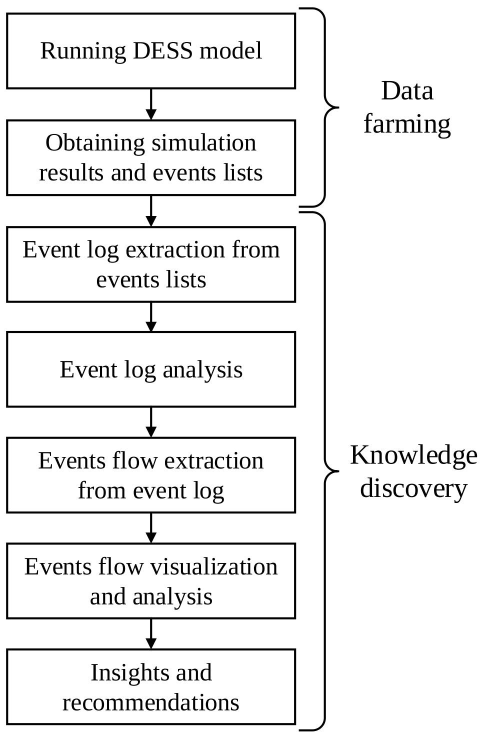

Figure 3: The Stages of the Events List-Based Approach

During data farming stages, the queueing system DESS model provides both simulation results and process history as the record of events list. The events list provides the necessary base for event log extraction, analysis and further processing. The data export from events list provides the possibility to save events flows by event type. For example, in a simple queueing system G/G/*n* (Kendall's notation), there are arrival events, service finish events, as well as other event types, but a particular events list containing only arrival events notices can be used, as well as any particular events list can be constructed. Thus the conventional events list of a simulation model is transformed into particular event logs and further into events flows.

The events flows are introduced as the chronological sequence of the events of the same type, each event is associated with a particular model entity and simulation time. The example of events flow is a sequence of entity arrivals in a service system.

As soon as events are categorized and events flows are described, the knowledge discovery may be implemented not only through visualization, but also through in-depth analysis of interaction of events flows.

The standard procedure for analysis of events flows cannot been proposed, as it depends on the discoveries in the previous stages. The relevant example of events flow analysis is provided in the case study at the following section.

In the proposed approach events list entries provide additional options for simulation result analysis, as well as extending the approach applicability from analysis of the processes in simulation models to analysis of the processes in real systems.

5. CASE STUDY

The following case study tends to conduce to revealing knowledge from simulation results by applying the process mining techniques.

5.1. Problem Statement

One of the fundamental phases of the simulation procedure is an appropriate analysis of simulation results. However, analysing results commonly events list is not taken into account. The events list keeps all the data concerning the changes of the system state. The event log may be easily extracted from events list and transformed into events flow protocol. Further in the text it is assumed as performed action and the existing of event log is fixed. In order to analyse data from the event log process mining techniques could be used. A simple queuing system with a queue of infinite capacity, exponential distribution of interarrival times, variable average value of this exponential distribution at different daytime, random service time, queue-length based probability of arriving customers not served is an object of the case study. Interarrival times and service times are independent random variables. Service discipline is first in first out. Service time is a function of the demand for product. The variable demand for product is used for evaluation of the revenue and described as empirical distribution. Dropout rule is introduced as a function of the queue length. Kendall's notation G/G/*n* queue is relevant for this system. Thus the system performance is not stable, and indicators – service quality and revenue – are conflicting. As a result, it is necessary to formulate a trade-off approach to achieve the best possible performance.

The detailed analytical study of system performance measures is complicated. For the purposes of performance analysis, a discrete-event simulation model of a system is created. Process mining techniques are applied to determine the parameter values for obtaining such performance.

5.2. The Summary of the Research Approach

In the proposed approach data from the events list is used for enriching opportunities in the area of simulation result analysis for getting extra insights from events list. The research approach combines the fundamental techniques of data farming and knowledge discovery. In the data farming phase, the simulation model is constructed, simulation experiments are designed, replicated and results obtained. Knowledge discovery phase is based on process mining ideas and implemented as follows. The conventional events list of a simulation model is transformed into particular event logs and further into events flows. From that insights and recommendations concerning acceptable values of experimental factors are revealed and visualized.

5.3. Comparison of Case Studies

In the previous research (Šitova and Pečerska 2018) the set of system parameters values was obtained by applying data mining techniques to simulation results from several replications. The simulation model in this case study used for data farming was the same as in the current research. Further this case study is referred as Case Study 1 while the current case study as Case Study 2.

Taking into account the fact that the same model was analysed during the knowledge discovery phase, it is possible to make a relevant comparison of both case studies. Comparative information is provided in Table 2. A detailed description of Case Study 2 realization is provided in the following sections.

5.4. Simulation Model of a Queuing System

For the purposes of performance analysis of the system, a discrete-event simulation model is used, providing the estimates of the relevant performance measures.

Outputs of the simulation model are statistics of server utilization, customer time spent in the system, a volume of satisfied demand, number of services and dropouts, and some other typical performance measures that are described in (Šitova and Pečerska 2017). There are two types of dropouts – working-schedule based and dropout rule-based. There are two experimental factors used for experimentation – number of servers and working schedule.

Table 2. Comparison of Case Studies

Features	Case Study 1	Case Study 2
Type of analysed simulation outputs	Time-persistent and observational simulation results (e.g. number of dropouts; queue length statistics)	Events flow (e.g. flow of served customers; flow of dropouts because of gate being closed)
Techniques applicable for analysis	Data mining techniques: correlation analysis, clustering and visualization	Process mining techniques: events list, event log and events flow extraction and analysis, visualization
Target performance indicator	Composite objective function	Integral value of dropouts flow

The conceptual model of the queueing system (QS) under consideration is shown in Figure 4. The model is created in accordance with the recommendations from (Robinson 2015).

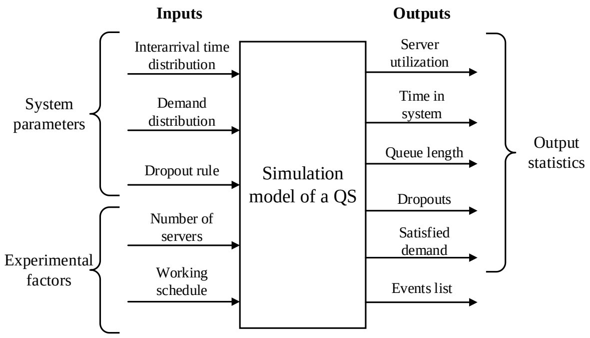

Figure 4: The Conceptual Model for the Case Study

Since this study is a continuation of Case Study 1 the same simulation model is used. Simulation model is created step by step as suggested in (Martin 2018).
The goal of the previous study was achieved, i.e. the combination of input variable values and adjustable working schedule, providing the best values of the composite objective function, was found. Therefore, the simulation model was configured as proposed in the work (Šitova and Pečerska 2018). However, these suggestions concerning experimental factors were not strictly defined. Case Study 1 resulted in proposed number of servers ranges from 2 to 3, the selected number – 3, number of service working hours ranges from 13 to 19, and the selected working schedule includes working times 07:00–15:00 and 16:00–24:00.
Initial simulation model layout as proposed in Case Study 1 is shown in Figure 5. The initial model parameters are selected as proposed in Case Study 1. After running the simulation model, results were

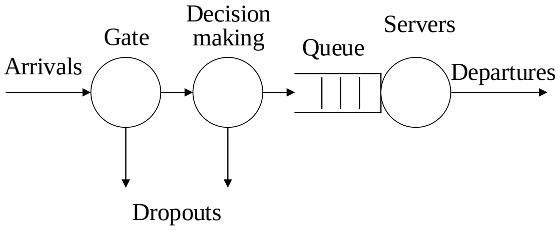

Figure 5: Initial Simulation Model

obtained in the form of events list, which were used for the further analysis.

5.5. Log data transformation: Events list, Event log, Events flow

For the analysis of simulation results, the MS Excel is used as a powerful tool with available visualization. Resulting data from the simulation model in form of events list were uploaded to MS Excel. All further operations were held in it.
Number of possible occurring events in created simulation model is finite. In this case there are 7 different event types. Each event has its ID and description, this information is provided in Table 3.

Table 3. Possible Event Types in Simulation Model

Event's ID	Event description
1	Arrival
2	Crossing gate
3	Start of service
4	End of service
5	Facing closed gate
6	Facing unsatisfying queue size
7	Departure

Events list contains information about simulation time, entity's ID and event type. The full set of events in the list includes records about all cases, associated with them events with certain timestamps i.e. each entity in the system and each change of its transitions in time. Here case, event and timestamp are interpreted as process mining terms in (van der Aalst 2011). Decomposition of the QS process into cases and events is shown in Figure 6.

Case ID	Event ID	Time-stamp
I	1	10:01
	2	10:01
	3	10:01
	4	10:04
	7	10:04
…	…	…
555	1	14:22
	2	14:22
	6	14:22
	7	14:22
…	…	…
999	1	23:07
	5	23:07
	7	23:07
…	…	…

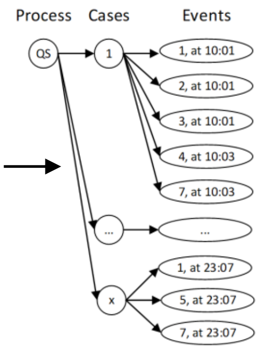

Figure 6: Decomposition of the QS Process

Results in the form of events list displayed as in Table 1 are transformed into event logs according to defined event types played in Table 3. As a result, logs of all event types occurring at a specific time are obtained. After events categorization events flows are described. As soon as the events flows are the chronological sequence of the events of the same type, seven particular events flows are obtained.
Events flow example, namely flow of arriving customers in the system, is shown in Figure 7.

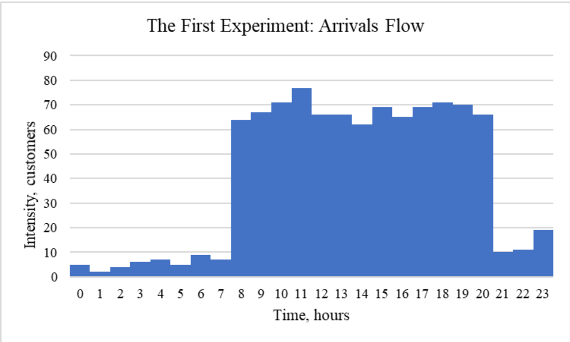

Figure 7: The Flow of Arriving Customers in the System

Analysis of created events flows, and experiments associated with them are discussed in the next subsection.

5.6. Analysis of Events Flows

Analysing arrival flow in Figure 7, it is clear that an inefficient working schedule was chosen for the initial simulation model (hereinafter is referenced as the first experiment). In the first experiment the gate is being closed at 00:00–07:00 and 15:00–16:00, i.e. queueing system is not working. While a decision not to work at night-time is a reasonable one, an hour in the middle of the day is inappropriate. During this hour a large number of potential customers arrives, which are forced to leave the system immediately. Given this fact, the second experiment is conducted. For this experiment the working hours of the system are changed. Based on Figure 7 the working hours are changed to 08:00–21:00. Figure 8 shows changes in dropouts flow after applying the new working schedule.
Figure 8 shows an approximately 25 percent decrease in the number of dropouts because of the gate being closed in the second experiment. Moreover, the total amount of working hours decreased from 16 to 13. This led to the conclusion that the efficiency of the QS after changing working schedule increased.
To detect the best number of servers, the third experiment is conducted. This time the simulation model contained 2 servers. To analyse this experimental factor impact, two events flows were analysed simultaneously. Changes in dropouts flow and served customer flow after changing number of servers from 3 to 2 respectively are displayed in Figure 9 and Figure 10.

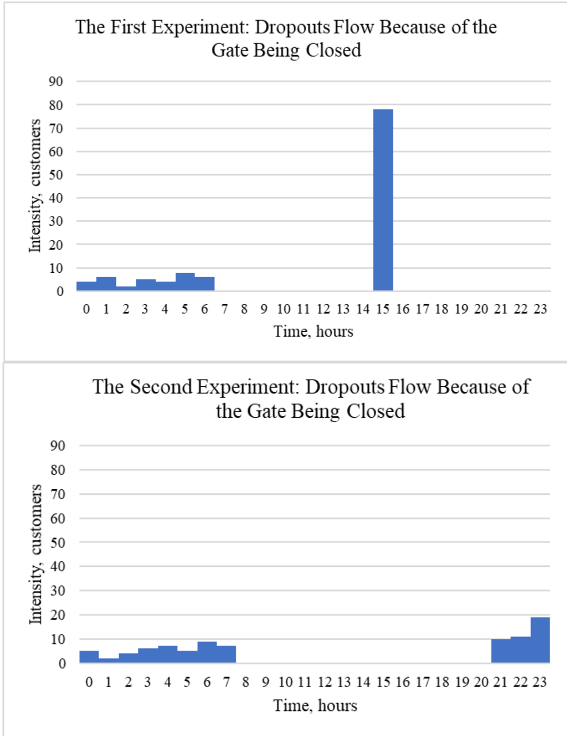

Figure 8: The First and the Second Experiments: Changes in Dropouts Flows

Figure 9 and Figure 10 show an approximately 20 percent drop in the number of served customers.
A swift decrease in the number of served customers is a critical factor for the service system.

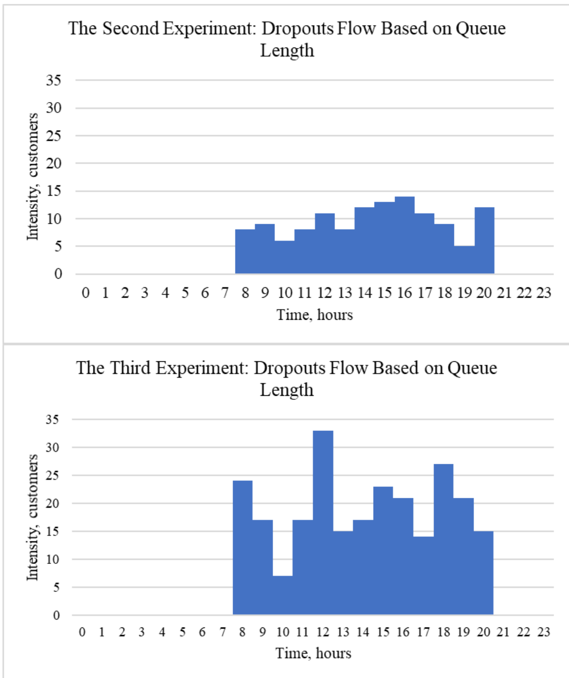

Figure 9: The Second and the Third Experiments: Changes in Dropouts Flows

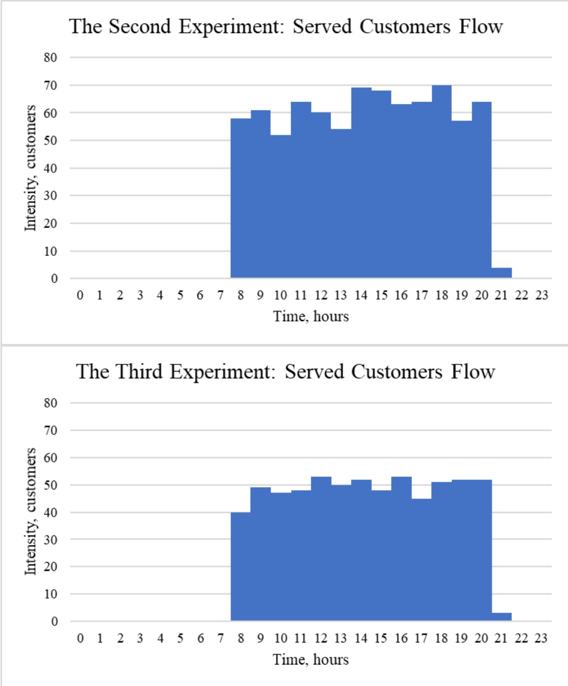

Figure 10: The Second and the Third Experiments: Changes in Served Customers Flows

Based on this, the authors decide that 3 is a required number of servers for efficient performance of this system.

5.7. Analysis of Experiments Results and Recommendations

Information summary about experimental factors in all the experiments is shown in Table 4.

Table 4. Comparison of Experimental Factors

Experi-ment Nr.	Working hours	Working schedule	Nr. of servers
1	16 hours	07:00–15:00; 16:00–24:00	3
2	13 hours	08:00–21:00	3
3	13 hours	08:00–21:00	2

According to the analysis of simulation results based on events flows, it may be concluded that the experimental factors in the second experiment provide the improved results: 25 percent decrease in the number of dropouts and 20 percent increase in the number of served customers.

On the basis of the results of the case study the following recommendations are made for efficient operation of the queueing system: number of servers is 3, there are 13 working hours and the selected working schedule includes working time from 08:00 to 21:00.

Thus the goal of the case study – to find experimental factors values leading to the best performance of a queueing system with a help of process mining techniques – has been achieved.

6. CONCLUSIONS

The experiments performed in the research confirming applicability of the proposed approach helped authors to answer the main question of the research "Is it possible to reveal insights and formulate recommendations from the simulation results by applying the process mining techniques?" positively.

Data transformation from events list of a simulation model into particular event logs and then into events flows helps to reveal the new knowledge about the performance of the system and to find the parameter values providing the advisable performance. The number of dropouts and the number of served customers are key factors in decision-making about the best performance of a queueing system.

To conclude, the developed events list-based approach is applicable to extracting process-related information and performing in-depth process analysis of queuing systems as well as other simulation-based projects.

In spite of satisfactory results of the research there is one limitation. The research did not include application of process mining algorithms. Thus the authors plan to continue research in the field of analysis of simulation results. In the future works, the authors are going to use process mining algorithms to verify their applicability in simulation results analysis.

ACKNOWLEDGMENTS

This work has been supported by Riga Technical University doctoral research grant program for academic year 2018/2019.

REFERENCES

Banks J., Carson J.S., Nelson B.L., Nicol D., 2010. Discrete-event System Simulation, 5th ed. Upper Saddle River, NJ. USA: Prentice Hall.

Brady T. F., Bowden R. A., 2001. The effectiveness of generic optimization routines in computer simulation languages. Proceedings of the 2001 Industrial Engineering Research Conference. Dallas, Texas.

Brady T. F., Yelling E., 2005. Simulation data mining: A new form of computer simulation output. Proceedings of the 2005 Winter Simulation Conference, pp. 285–289. https://doi.org/10.1109/WSC.2005.1574262.

Dunham M., 2003. Data Mining: Introductory and Advanced Topics. Pearson Education, Inc. ISBN-13: 9780130888921.

Cassandras C. G., Lafortune S., 2008. Introduction to Discrete Event Systems, Second edition., Boston, MA, USA: Springer. https://doi.org/10.1007/978-0-387-68612-7.

Feldkamp N., Bergmann S., Strassburger S., 2015. Knowledge Discovery in Manufacturing Simulations, Proceedings of the 3rd ACM SIGSIM Conference on Principles of Advanced Discrete Simulation. https://doi.org/10.1145/2769458.2769468.

Feldkamp N., Bergmann S., Strassburger S., Schulze T., 2016. Knowledge discovery in simulation data: A case study of a gold mining facility. Proceedings of the 2016 Winter Simulation Conference. https://doi.org/10.1109/WSC.2016.7822210.

Haav H.-M., Kalja A., 2014. Business Process Mining in Warehouses: a Case Study. Proc. of the 11th International Baltic Conference, Baltic DB&IS 204. Tallinn: Tallinn University of Technology Press, 387−394.

Han J., Kamber M., 2006. Data Mining: Concepts and Techniques, Second Edition, San Francisco, CA, USA: Elsevier Inc.

Khemiri, A., Hamri, M. A., Frydman, C., Pinaton, J. Improving business process in semiconductor manufacturing by discovering business rules, Proceedings of the 2018 Winter Simulation Conference, pp. 3441-3448. https://doi.org/10.1109/WSC.2018.8632509.

Kibira D., Hatim Q., Kumara S., et al., 2015. Integrating data analytics and simulation methods to support manufacturing decision making. Proceedings of the 2015 Winter Simulation Conference. https://doi.org/10.1109/WSC.2015.7408324.

Law A.M. and Kelton W., 2000. Simulation modeling and analysis. 3nd ed. Mc Graw Hill Higher Education.

Martin N., Depaire B., Caris A., 2018. A synthesized method for conducting a business process simulation study. Proceedings of the 2018 Winter Simulation Conference. 276-290. 10.1109/WSC.2018.8632298.

Merkurjevs J., Pečerska J., Tolujevs J., 2008. Simulation-Based Analysis of Logistic Systems, Humanities and Social Sciences. Latvia, vol. 4, no. 57, pp. 27–48.

Nakatumba J. and van der Aalst W. M. P., 2009. Analyzing Resource Behavior Using Process Mining. In Business Process Management Workshops, volume 43 of Lecture Notes in Business Information Processing, pages 69–80. Springer. ISBN 978-3-642-12185-2.

Painter M. K., Erraguntla M., Beachkofski B., et al., 2006. Using simulation, data mining, and knowledge discovery techniques for optimized aircraft engine fleet management. Proceedings of the 2006 Winter Simulation Conference. https://doi.org/10.1109/WSC.2006.323221.

Pidd M., 2004. Systems Modelling: Theory and Practice. Chichester: John Wiley & Sons Ltd.

Robinson S., 2015. A tutorial on conceptual modeling for simulation, Proceedings of the 2015 Winter Simulation Conference. https://doi.org/10.1109/WSC.2015.7408298.

Šitova I., Pečerska J., 2018, Approach to Integration of Data Mining Techniques in Simulation Results Analysis. Information Technology and Management Science. vol. 21, pp. 86–92. https://doi.org/10.7250/itms-2018-0014.

Šitova I., Pečerska J., 2017. A concept of simulation-based SC performance analysis using SCOR metrics, Information Technology and Management Science. vol. 20, pp. 85–90. https://doi.org/10.1515/itms-2017-0015.

Schriber T. J., Brunner D. T., Smith J. S. Inside discrete-event simulation software: how it works and why it matters. Proceedings of the 2014 Winter Simulation Conference. https://doi.org/10.1109/WSC.2014.7019884.

Tax N., Sidorova N., Haakma R., van der Aalst W. M. P., 2018. Mining Process Model Descriptions of Daily Life through Event Abstraction In: Bi Y., Kapoor S., Bhatia R. (eds) Intelligent Systems and Applications. IntelliSys 2016. Studies in Computational Intelligence, vol 751. Springer, Cham. https://doi.org/10.1007/978-3-319-69266-1_5.

van der Aalst W. M. P., 2011. Process Mining - Discovery, Conformance and Enhancement of Business Processes. Springer. https://doi.org/10.1007/978-3-642-19345-3.

van der Aalst W. M. P., 2012. Process Mining: Overview and Opportunities. ACM Trans. Management Inf. Syst., 3(2):7. https://doi.org/10.1145/2229156.2229157.

van der Aalst W. M. P, 2016. Process Mining: Data Science in Action. Springer. https://doi.org/10.1007/978-3-662-49851-4.

Witten I. H., Frank E., 2005. Data Mining: Practical Machine Learning Tools and Techniques, Second Edition, San Francisco, CA, USA: Elsevier Inc.

AUTHORS BIOGRAPHY

Irīna Šitova is a Doctoral student at the Information Technology Institute, Riga Technical University (RTU). She received her M. sc. ing. degree in Information Technology from RTU in 2018. Currently, she is a Scientific Assistant at the Department of Modelling and Simulation of RTU. Her interests include discrete-event simulation, decision making, data mining, machine learning and process mining.

Jeļena Pečerska is a Doctor of Information Technology in the field of system analysis, modelling and development. She has been working for Riga Technical University since 1979. At the moment, she is an Associate Professor at the Department of Modelling and Simulation of Riga Technical University. Professional interests include methodology of discrete-event simulation, combined simulation, supply chain modelling, practical applications of discrete-event simulation and discrete-event simulation in education. She is a member of the Latvian Simulation Society.