Non-Container Port Services Bottlenecks Identification Using Process Mining and Simulation Analysis

Alvin Syarifudin Shahab[1,*] Mohammad Isa Irawan[2]

[1] *Technology Management Department, Institut Teknologi Sepuluh Nopember: alvinsyarifudin@gmail.com*
[2] *Department of Mathemathics, Institut Teknologi Sepuluh Nopember: mii@its.ac.id*
[*]*Corresponding author. Email: alvinsyarifudin@gmail.com*

ABSTRACT

The number of loading and unloading processes is one of the port's main business processes. In carrying out loading and unloading services, there are fundamental indicators used as performance evaluation of a port in Indonesia, namely T/G/H (Ton/Gang/Hours) for the non-container performance of loading and unloading production. At one of Indonesia's ports, the T/G/H indicator from 2018 to 2020 several times did not reach the specified minimum standard. Therefore, it is necessary to evaluate the loading and unloading time of activities with bottlenecks to minimize and improve performance. In this research, the identification of business process flows is based on applicable SOP using YAWL to obtain the event logs. The process mining tools then carry out the event log results from the system to identify the most significant bottlenecks in the business process. The activities are then analyzed using a queuing system simulation analysis based on ARENA tools. This study's results obtained three processes with the biggest bottlenecks in the business process: the loading and unloading process, the timesheet, and the realization entries with the highest waiting time are 67 hours. These results can be a reference for the company to improve its performance in the future.

Keywords: Business process management, event log, non-container loading and unloading, process mining, queuing system simulation analysis.

1. INTRODUCTION

The era of digitalization reaches every aspect of the industry, including the port industry. Every process is monitored into the database as in the non-container loading and unloading activity starts from the ship ties the rope at the dock until the ship releases the mooring rope. These data are beneficial for the company and could be analyzed to gain several aspects of the process. The loading and unloading process requires at least a few days interval is called the Berth Time. Each non - container loading and unloading service at port consists of several activities which vary in every port. At one of the ports in Indonesia, the KPI (Key Performance Indicator) of bulk cargo T/G/H (Ton/Gang/hours) from 2018 to 2020 slightly reached, but there is also not reached the minimum standard as shown in Fig. 1. Therefore it is necessary to identify the business process to find the bottlenecks.

Every business process has a standard operating system (SOP) to reference how the process should work. The typical approach to simulate using workflow tools such as YAWL is one of the newest business process modeling techniques. It is integrated with existing mining process tools to generate the event logs based on the actual process [1]. Event logs store all actions from a

running task to completing the assignment or event in various formats for easy integration with other tools [2].
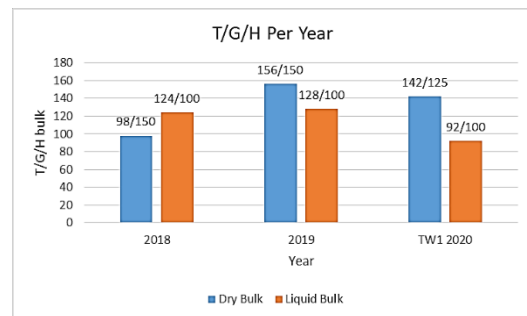

Figure 1 T/G/H KPI value from 2018 to the quarter of 2020

Process mining is an objective way to gain knowledge from the event log generated from BPM (Business Process Model) or database to analyze the business processes from event log analysis results [3]. The process mining also can validate the process flow and finds the bottleneck that runs by using ProM as one of the tools [4]. The mining process is generally used in three activities [5]: Find a process model, conformance checking, and optimization.

The event log data could be generated either from workflow tools or ERP software like SAP or company databases. It must have minimal properties such as id, timestamps, and activity [6]. Process mining can only sometimes not capture the other elements of the existing system that cannot express in analytical solutions to gain the system's dynamic side. Then a computer simulation is needed [7]. Research in [8,9,10] used the queuing system simulation based on the ARENA tools to identify the waiting time, bottlenecks, and service time with detailed identification of the problems. That research applies the simulation model and queue theory to all business process systems directly instead of analyzing the most significant delay or bottlenecks' activity.

According to [11], the determination of waiting line management is based on the process's arrival rate and service rate. The queuing system analysis could be conducted using the ARENA software, which operates as the validation tool to simulate the process, waiting time automatically, and any other properties.

The integration of process mining with simulation methods is needed in carrying out business process reengineering. Simulation and mining processes integration, comparative analysis can be easily carried out using workflow-based simulation tools [12]. This study aims to identify the bottleneck in loading and unloading non-container by using event log data, then analyzed using process mining and simulation analysis. This research simulation is based on the queuing theory and developed the ARENA software model, which we will discuss in this paper.

2. BUSINESS PROCESS AND METHOD

2.1. Business Process

The business process in loading and unloading at the port can be grouped into two primary functions: the loading-unloading process and realization, which consist of 11 activities from the gate until approval, as shown in Fig. 2.
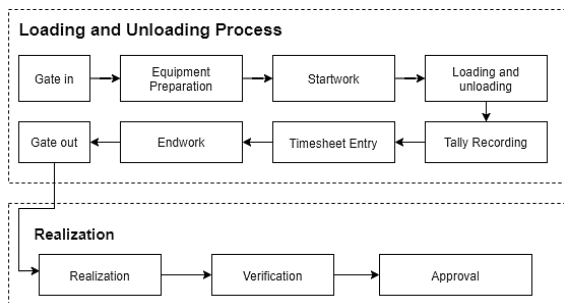

Figure 2 Port business process

This research was conducted in Indonesia at the port of Tanjung Perak, which has three terminals: Jamrud, Nilam, and Mirah terminals with seven jetties and nine

moorings loading-unloading processes. The service system applied to ships' arrival is FIFO (First In, First Out), where the vessel will be served according to its arrival order.

The FIFO system does not take into account the differences in the cost of ships of different sizes. North Jamrud, west Jamrud, east Nilam, and Mirah jetty can handle dry and liquid bulk, then south Jamrud and Mirah-Bogasari can only handle dry bulk. In contrast, the nilam kade intan jetty can only take liquid bulk. The T/G/H KPI is calculated using (1).

$$TGH = \frac{total\ tonnage}{total\ gang * total\ effective\ time} \qquad (1)$$

This research method combined process mining to find the activities that have the biggest bottleneck in the business process and simulation analysis to simulate the existing condition. The bottleneck identification process flows of this research is shown in Fig.3
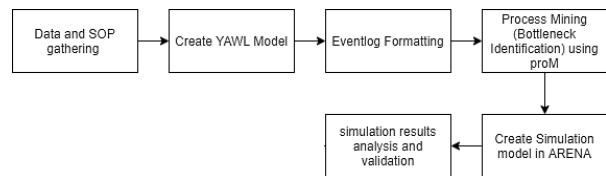

Figure 3 Bottleneck identification process

The data obtained from the company software from 2018 to 2020 is 988 data. The amount of data is data that has completeness starting from the gate process to realization, as shown in Fig. 4.
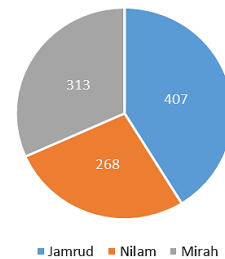

Figure 4 Total data each terminal

The data has several properties in the form of id (ship id number), start date (activity start date), and end date (activity end date) and activity, as shown in Table 1.

Table 1. Data properties Samples

Id	Start Date	End Date	Activity
1912021094	12/24/2019 22:20	12/24/2019 22:20	Gate In
1912021094	12/24/2019 22:25	12/24/2019 22:35	Equipment Prep
1912021094	12/24/2019 22:40	12/24/2019 22:40	Start Work
1912021094	12/24/2019 22:41	12/26/2019 16:01	Loading & unloading
1912021094	12/26/2019 16:02	12/26/2019 16:02	Tally recording

Id	Start Date	End Date	Activity
1912021094	12/26/2019 23:20	12/26/2019 23:20	Time sheet
1912021094	12/26/2019 23:20	12/26/2019 23:20	End Work
1912021094	12/26/2019 23:22	12/26/2019 23:23	Gate Out
1912021094	12/26/2019 23:23	12/27/2019 2:54	Realization

Based on the loading and unloading business process, the YAWL model consists of 2 primary operations, as shown in Fig 5.
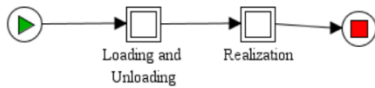

Figure 5 Loading - unloading process YAWL model level 1

Each process block has a sub-process that consists of several activities. The loading and unloading process has eight activities from the gate-in until gate-out, as shown in Fig 6. The loading-unloading primary process is parallel by tally recording because the SOP states that it works in parallel using join and spit blocks.
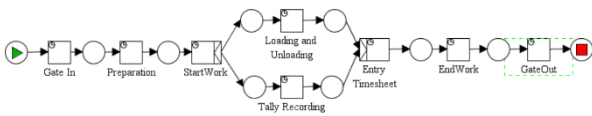

Figure 6 Loading – unloading process YAWL model level 1

The realization process consists of 3 activities, from realization entry to approval, as shown in Fig. 7. Each activity in the model time is arranged based on the system log data in the process block by utilizing the YAWL timer feature. After the timing process is carried out, the model is uploaded to the YAWL engine and run on the Tomcat server that appears by opening a browser on a local server to generate the event log format data.
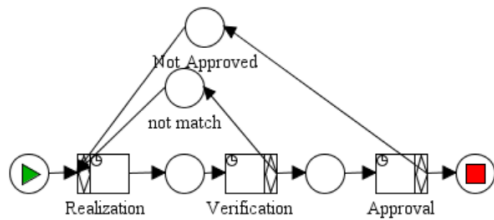

Figure 7 Realization process YAWL model level 1

3. IMPLEMENTATION

The event log generated from YAWL has the (.xes) format in which it consists of all cases that have been executed. The event log data must be filtered and converted into (.mxml) format to perform process mining. The filtering was served with a heuristic miner using ProM 5.2 tools.

To analyze the results of modeling in YAWL with ProM, the alpha ++ algorithm can detect implicit dependencies on a Petri net, as shown in Fig. 8.
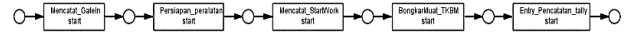

Figure 8 Process Mining Result

Based on the alpha++ algorithm results, it can be seen that the sequence of the paths formed is by the SOP sequence starting from the gate in until the realization. However, there are differences in the YAWL model in the loading and unloading section of the tally, which does not form parallel because the entry process based on log data is complete after loading and unloading. The event log data also shows the throughput of each case where the most extended process is 306 hours and the fastest is 10 hours, as shown in Fig. 9.
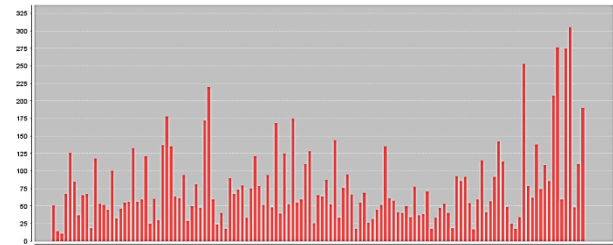

Figure 9 Event log throughput result

After obtaining the Petri net results from the alpha++ algorithm, the data is analyzed using the Petri net performance analysis by setting the time to find the activities that have the most significant bottlenecks. The results obtained that several activities have bottlenecks, as shown in Fig. 10.
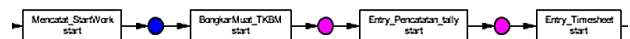

Figure 10 Loading and unloading process bottleneck

This analysis result comes with three color codes showing the waiting time of the transition. The purple color shows a high waiting time. The yellow color shows a normal waiting time, and the blue color indicates a low waiting time. The loading and unloading activity has two transition activities colored by purple, loading-unloading TKBM and timesheet entry, with an average waiting time of 51.3 hours and 4.23 hours on average. Another activity that has a bottleneck is realization entry, as shown in Fig. 11.
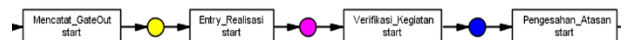

Figure 11 Realization process bottleneck

There is a bottleneck in the entry realization process with an average processing time of 12.86 hours. Based on

the process mining result, the three activities with a high bottleneck will be modeled in ARENA to simulate the process more detail.

3.1. Simulation Model

The first simulation is the loading and unloading process which its primary resource is a crane or pump. An entity served in the system is a ship carrying cargo, whether bulk or liquid cargo. Simulation arrival and service rate for each jetty are collected and analyzed using an ARENA input analyzer to generate the distribution, as shown in Table 2.

Table 2. Arrival and service rate distribution loading-unloading process

Input Data	Distribution	Expression
Time between arrival	Exponential	-0.001 + EXPO(4.03)
North Jamrud jetty dry cargo	Beta	35 + 136 * BETA(1.19, 1.99)
North Jamrud jetty liquid cargo	Gamma	23+GAMM(26.4, 0.936)
South Jamrud jetty dry cargo	Weibull	4+WEIB(60.8, 1.19)
West Jamrud jetty dry cargo	Exponential	24 + EXPO(60.9)
West Jamrud jetty liquid cargo	Triangular	TRIA(16, 38, 60)
East Nilam jetty dry cargo	Erlang	10 + ERLA(9.79, 4)
East Nilam jetty liquid cargo	Normal	NORM(49.9, 18.4)
Nilam Kade Intan jetty liquid cargo	Exponential	23 + EXPO(30.4)
Mirah dry cargo	Weibull	5+WEIB(33.3, 0.909)
Mirah liquid cargo	Normal	NORM(49.8, 20.9)
Mirah bogasari dry cargo	Triangular	TRIA(34, 46.2, 117)

Based on the distribution of arrival rate, service rate, and the tonnage of every jetty collected, it is then assigned in every process, create and assign block in the model.
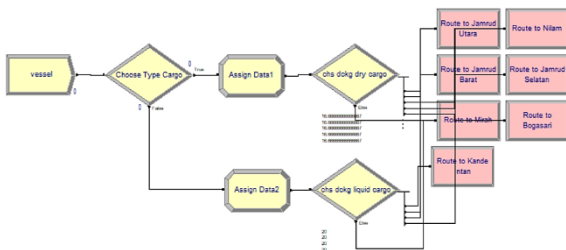

Figure 12 Arrival ship decision model

The first model is the entity generator and the cargo arrangement, as shown in Fig. 12, which uses 52% probability for dry bulk and 48% for liquid bulk based on the data. Meanwhile, the ship mooring decisions process is using equal probability.

The loading-unloading simulation process was carried out with three replications as experiments on the ARENA to get the half-width value and determine the simulation stop time. The total replication is related to the confidence interval value, which states the percentage of error. By using three times replication, the half-width value is 31.94 hours for the total time entity. According to [13], to reduce the weight of the half-width is using (2).

$$n \cong n_0 \frac{h_0^2}{h_1^2} \qquad (2)$$

Where n is the number of replications meanwhile h0 is the initial half-width, and h1 is the new half-width. The value of half-width is set to 5, and the number of repetitions is 122. The duration of each replication process is set to 30 days and 24 hours. Fig. 13 shows the model loading and unloading at the nilam jetty model consists of one mooring which can handle liquid and dry bulk (two process blocks in one mooring).
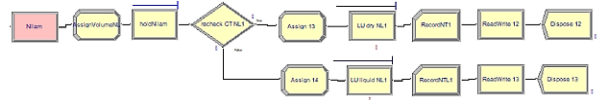

Figure 13 Loading – unloading nilam model

The holding block in the model is used as a queue, a place for the ship to wait until the mooring becoming available because, in the actual process, the loading and unloading of dry bulk and liquid bulk in one mooring have to take turns.

The model in another jetty has a similar model, which differentiates the number of mooring that will affect the number of process blocks in each jetty. The second model is the entry timesheet process as the second bottleneck. The primary resource in this process is human resources who work in each terminal. Each terminal has a different number of these officers.

Figure 14 Entry Timesheet Mirah model
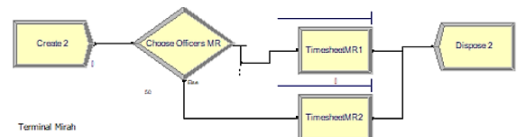
Jamrud terminal has four officers meanwhile, Mirah terminal only has two officers. Nilam terminal has three officers, as shown in Fig. 14 symbolized by process block Time sheetNL1 to 3.

The distribution of arrival and service rate are collected in each terminal based on the data using an input analyzer, as shown in Table 3.

Table 3. Arrival and service rate distribution timesheet process

Input Data	Distribution	Expression
The time between arrival Jamrud Terminal	Exponential	-0.001+EXPO(56.1)
Jamrud terminal service rate	Beta	-0.001+14*BETA (0.835, 2.31)
The time between arrival NilamTerminal	Beta	-0.001 + 140 * BETA(0.739, 4.04)
Nilam terminal service rate	Beta	12*BETA(0.7,1.23)
The time between arrival Mirah Terminal	Exponential	-0.001+EXPO(69.1)
Mirah terminal service rate	Beta	-0.001+12*BETA (0.771,1.83)

Based on the simulation results using three replications, the value of half-width for the total time is 15.06, with an average of 11.65 hours. To reduce the half-width, then using (9) with a new half-width of 3, the number of replications is 76, and the length of one replication is 30 days.

The third model is the realization process, where primary resources are human resources like the timesheet process. The officers who do this work are the same person as the timesheet, so the model is similar, and the only difference is the service and arrival rate. Fig.15 shows the realization model in the Mirah terminal, which has two officers as described before.
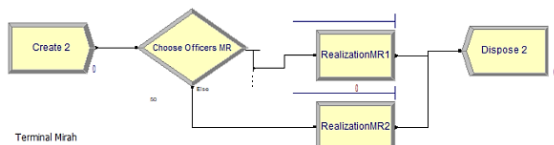

Figure 15 Entry Realization Mirah model

Based on the simulation results using three replications, the value of half-width for the total time is 13.50, with an average of 7.42 hours. The number of repetitions for realization is 61, with a new half-width of 3. The distribution of arrival and service rate are described in table 4.

Table 4. Arrival and service rate distribution Realization Process

Input Data	Distribution	Expression
The time between arrival Jamrud Terminal	Exponential	EXPO(64.2)
Jamrud terminal service rate	Lognormal	LOGN(22.3, 32)
The time between arrival at Nilam Terminal	Weibull	WEIB(35.1,0.54)
Nilam terminal service rate	Lognormal	LOGN(31.5,70.5)
The time between arrival Mirah Terminal	Exponential	EXPO(65.3)
Mirah terminal service rate	Lognormal	LOGN(29.8, 52)

4. RESULT AND DISCUSSION

4.1. Loading and Unloading Simulation Result

The first simulation result is the loading and unloading process. The simulation results are used to compare the system performance with each terminal performance in the actual procedure. Table 5 shows the simulation time results for each terminal served along with the actual loading and unloading process time, KPI, and the time differences.

Table 5. Loading and Unloading Simulation Result

Terminal	Simulation time Hours	Actual Hours	KPI T/G/H liquid bulk	KPI T/G/H dry bulk	Time Diff. Hours
Jamrud	67.354	62.96	94.071	99.57	4.393
Nilam	50.612	55.45	72.117	72.97	4.837
Mirah	51.720	63.54	73.934	241.3	11.82

Based on the simulation results, Jamrud, Nilam, and Mirah terminal have average dry bulk tonnage, respectively 5790, 3002, 12929 Ton and liquid bulk respectively 3519, 2959, 2846 kiloliter. All terminals have KPI values below 100 which have not yet reached the minimum target. The average process time is more than 48 hours in each terminal, showing a primary resource problem.

The simulation result is validated using the Chi-Square test for Association based on the research in [14] shows that Chi-Square test is the best way to validate the simulation model with its existing system using the following hypothesis:

- H0: x2=y2 the model can represent the actual system

- H1: x2≠y2, the model, cannot represent the actual system.

Where x represents the simulation results of the model and y represents the actual system. The p-value result is 0.574763, where the significance level is 0.05 so

that the H0 hypothesis is not rejected can represent the actual system.

In the loading and unloading process, it can be inferred there are several factors that cause the length of the process. The volume, the primary equipment for the loading and unloading process (grabs for dry bulk and pumps for liquid bulk), and the number of gangs per shift for operating the crane or pump.

Based on these three factors, the factors that can be controlled are the equipment and the number of the gang/team, while the volume can vary according to the incoming ship's size. However, the size of the gang used will affect the T/G/H KPI value. The more number of gangs/teams will make the process faster but made the KPI value lower. Therefore, data collection was carried out on the main loading and unloading equipment for liquid and dry bulk at each terminal to validate the problems.

Table 6. Equipment resource in each terminal

Equipment Type	Maximal Capacity	Total	Terminal
Grab	15 Ton	8	Jamrud
Harbour Mobile Crane	150 Ton	10	Jamrud
Shore Crane	40 Ton	1	Nilam
Shore Crane	40 Ton	1	Mirah
Grab	10 Ton	2	Mirah

Based on table 6, the equipment is not equally distributed in terms of the number and characteristics. Jamrud terminal equipment has complete equipment with a maximum load more significant than the other terminals, causing the loading and unloading speed to differ. The three terminals do not have pumping equipment, so the loading and unloading of liquid bulk rely on a ship pump or pipe.

4.2. Timesheet Simulation Result

The second simulation is the timesheet process. Based on the simulation results shown in table 7, the waiting time for officers at jamrud and nilam terminals has a considerable waiting time, respectively 3.42 hours and 2.21 hours. The probability of the service being busy at the two terminals also has a greater value than the Mirah terminal, where the jamrud terminal has a percentage of 11.91%. In contrast, the nilam terminal reaches 21.82%.

Table 7. Timesheet simulation result

Terminal	Sim. Time (Hours)	Real process (Hours)	Time Diff. (Hours)	Busy Probability
Jamrud	3.42	3.93	0.51	11.91%
Nilam	2.21	4.35	2.14	21.82%
Mirah	2.13	3.70	1.57	8.6%

The simulation result is validated using the Chi-Square test for Association, the same as the previous test. The p-value result is 0.91744 with a significance level of 0.05 so that hypothesis 1 is accepted and the model can represent the actual model. The process problem that can be inferred from the simulation result is not enough officers used in each terminal to accomplish all requests because the busy probability is also relatively high.

4.3. Realization Simulation Result

The third simulation is the realization process. Based on the simulation results in table 8, each officer at each terminal has a high waiting time of 53.43 hours at the nilam terminal. At the jamrud terminal, it reaches 21.86 hours, and the Mirah terminal runs 30.67 hours. The value of the service probability to be busy is also high and long waiting time. Furthermore, the average waiting time for each terminal is compared to data in the existing system.

Table 8. Realization Simulation Result

Terminal	Simulation time (Hours)	Real process (Hours)	Time Diff. (Hours)	Busy Probability
Jamrud	21.86	20.78	1.08	36.75%
Nilam	53.43	27.14	26.29	48%
Mirah	30.67	27.41	3.26	42%

Simulation results are validated with the same test as before. Based on the chi-square test results, the p-value is 0.17002, where the significance level is 0.05, so that the null hypothesis is not rejected and the model is validated and can represent the actual system. The problem in this process is the realization process takes a significant amount of time due to the admin officer's number of tasks. Apart from recording the realization, they have to handle the timesheet recording process and the loading and unloading application process.

5. CONCLUSION

There are a total of 3 activities that experience bottlenecks in the non-container loading and unloading process from the mining process. It is the main loading and unloading process with a waiting time of 48.68 hours, then the timesheet recording process with a waiting time of 3.68 hours, and the recording process of realization 12.86 hours. Of the three processes, the enormous waiting time is in the main loading and unloading process influenced by the specifications of the equipment used. The timesheet and realization process that shares the same resources could not handle the number of tasks. These activities must be optimized to increase the T/G/H KPI in the port services.

REFERENCES

[1] K. Vergidis, A. Tiwari, and B. Maieed, "Business Process Analysis and Optimization: Beyond Reengineering," IEEE Trans. Syst. Man Cybern. Part C Appl. Rev., (2008).

[2] M. Adams, A. V. Hense, and A. H. M. ter Hofstede, "YAWL: An open source Business Process Management System from science for science," SoftwareX, vol. 12, p. 100576, Jul. (2020).

[3] S. Aguirre, C. Parra, and J. Alvarado, "Combination of Process Mining and Simulation Techniques for Business Process Redesign: A Methodological Approach," in Lecture Notes in Business Information Processing, (2013).

[4] O. Tamburis and C. Esposito, "Process mining as support to simulation modeling: A hospital-based case study," Simul. Model. Pract. Theory, (2020).

[5] H. Tang, Y. Chen, and J. Lu, "Architecture of process mining based business process optimization," in IET Conference Publications, (2006).

[6] N. R. T. P. Van Beest and L. Mӑruşter, "A process mining approach to redesign business processes - A case study in gas industry," in Proceedings - 9th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing, SYNASC (2007).

[7] S. Ďutková, K. Achimský, and D. Hoštáková, "Simulation of queuing system of post office," in Transportation Research Procedia, (2019).

[8] R. Jain, H. Bedekar, K. Jayakrishna, K. E. K. Vimal, and M. Vijaya Kumar, "Analysis and Optimization of Queueing Systems in Airports—Discrete Event Simulation," in Lecture Notes in Mechanical Engineering, (2020), pp. 1189–1194.

[9] P. Songchitruksa and E. N. Hard, "Queuing simulation of roadside survey station: Blocked traffic lane," Transp. Res. Part A Policy Pract., (2008).

[10] R. Shone, K. Glazebrook, and K. G. Zografos, "Resource allocation in congested queueing systems with time-varying demand: An application to airport operations," Eur. J. Oper. Res., (2019).

[11] T. Bernard W, Introduction to Management Science , 11th ed. Pearson, (2013).

[12] B. Keith Norambuena, "Integration of process mining and simulation: A survey of applications and current research," in Advances in Intelligent Systems and Computing, (2018).

[13] Kelton, W. D., Sadowski, R. P., & Sadowski, D. A. (2002). Simulation With Arena. McGraw-Hill.

[14] Ringuest, J. L. (1986). A chi-square statistic for validating simulation-generated responses. Computers & Operations Research Volume 13, 379-385.