Human Resource Allocation to the Credit Requirement Process, A Process Mining Approach

Omid Mahdi Ebadati*
Department of Operation and Information Technology,
Kharazmi University,
Tehran, Iran
ebadati@khu.ac.ir

Mohammad Mehrabioun
Department of Operation and Information Technology,
Kharazmi University,
Tehran, Iran
mehrabioun@khu.ac.ir

Shokoofeh Hosesini
Department of Operation and Information Technology,
Kharazmi University,
Tehran, Iran
hoseini.shokoofeh@gmail.com

2022 13th International Conference on Information and Knowledge Technology (IKT) | 979-8-3503-2062-6/22/$31.00 ©2022 IEEE | DOI:10.1109/IKT57960.2022.10039030

***Abstract*—In the banking industry, process mining has been used in different case studies with promising results. Process execution efficiency can be improved by correctly allocating human resources, which also increases the revenue and satisfaction of clients in organizations. To gain access to these approaches among the resources, banks are trying to streamline their processes. For this reason, it is essential to take into consideration the prerequisites for an accurate view of the sessions. Nevertheless, few works have tended to process mining as it is connected to the banking industry. This work presents a new approach to solving human resource allocation problems in the credit requirement processes that combines process mining, statistical techniques, and the Theory of Constraint (TOC) method. The statistical techniques are used to build a simulation model. Moreover, to obtain more reliable results from the simulation, in this study, four scenarios are designed to simulate reality. By using it, the bottleneck in the process can be identified. At the end, TOC techniques have been used to tackle the identified bottleneck. This study suggests a combination of resources and activity within fewer iterations compared to the literature and proposes an online process to collect documents which is a second stage of the credit requirement process to overcome the existing bottleneck.**

Keywords*— *Business Process Management, Banking Industry, Metaheuristic Algorithm, Process Mining, Credit Requirement

I. Introduction

The credit requirement process represents a series of logically related activities with the number of employees to perform such a desired activity that is assigned to them [1]. The organization must perform the process in an improved manner that reduces costs, improves competitive advantages, and stays in the market [2]. Consider educating the employees to minimize activity execution times. It can often be concluded that growing the number of resources will enhance the process's average cycle time before the minimal cycle time is attained. Thus, deciding to train one's personnel to conduct their duties more quickly, several cost factors need to be taken into account. If the improvements in cycle time are not large enough, they will not justify the training cost. Rearranging tasks between employees is also possible, but careful consideration must be taken because it is not predictable whether the new assignee will be able to succeed or not. This emphasizes the importance of human resource allocation to a specific activity in the process. Nowadays, continuous improvement in banking industries is an important element of existing [3]. This research reveals the role of process mining in the credit requirement process as the main process in the banking business and points out how process mining can raise the efficiency and efficacy of the credit requirement process.

II. Literature

A. Process Mining in the Banking Industry

Financial applications, more specifically applications of the banking industry, are one of the top six case studies in process mining, according to [4].

Many applications with their underlying study areas in this domain exist regarding different processes in the banking sector. We discovered that papers exclusively discussing process mining methods in credit applications account for nearly 37% of this industry's overall number of papers. The second most relevant application is a loan reaching 30% and 8 papers. The third relevant application is related to payment and customer accounts at 7%, followed by other applications at 4%, such as ATM application, risk management, and multichannel application. Results indicate that the top two research topics account for more than half of banking application contributions. Fig 1 illustrates the distribution of banking applications in the field of process mining between 2007 and 2022.

B. Credit Application of the Banking Industry

A credit application is one that a potential borrower submits to a credit provider. The bank cannot move forward with a credit

979-8-3503-2062-6/22/$31.00 ©2022 IEEE

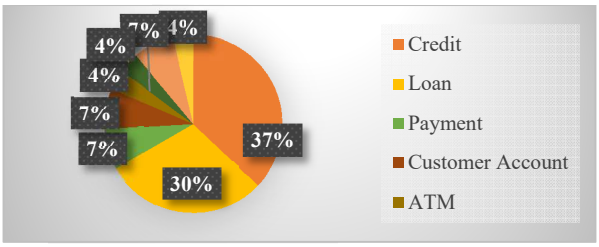

Fig. 1. Distribution of Papers by Application Type in the Banking Industry from 2007 to 2022

application without all the necessary information being provided in the credit application [5]. For processing credit applications, certain lenders could charge a fee. The prospective borrowers must submit a credit application regardless of whether they are applying for a house loan, a personal loan, an education loan, a car loan, a loan against collateral, a credit card, or any other type of loan. Nowadays, the majority of credit applications are made online. This section aims to introduce the study area of credit application. Most study area in recent year has been considered.

Figure 2 provides the most studied area of credit application in process mining. Fraud mainly aims to measure the difference between the process model and the execution of an entire process recorded in event logs at 42% in the credit application [6] [7] which conformance checking techniques must be applied to the process [8]. The second topic is related to resource allocation [4] [9] [10], which aims to optimize the process in terms of cost and time at 33% in the banking industry during recent years. Finally, the credit process is at 25%, mainly aiming at understanding activities flow [11].

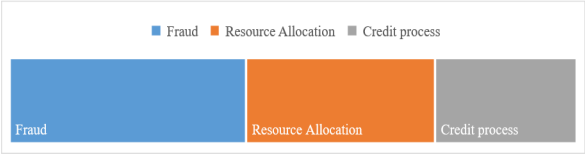

Fig. 2. The Most Study Areas in Credit Application

Process mining has three main perspectives [12]: ***process***, ***organizational***, *and* ***case*** *perspectives,* also the three major forms of the process: **discovery**, **conformance,** and **enhancement** [13]. The discovery type seeks to mine process models using process discovery methods; conformance aims to monitor and audit compliance; and enhancement attempts to improve and extend processes by providing attention, information, and KPIs (key performance indicators). In the following, research has been categorized under the process mining discipline.

From Table I, it can be inferred that the study of process mining in credit applications began in 2009. Reference [14] simulated the demands for a credit card during peak days. They employed discovery techniques along with three process mining perspectives. By including an organizational perspective, they designed a more realistic model.

The second study was conducted in 2012 by [11]. In order to improve the whole process through simulation and proper resource allocation, which provided a workflow strategy for operational decision support using all three process mining perspectives. They designed multiple resource allocation plans to propose the best combination of the resources and the activities of the credit requirement. Focused on the case perspectives of the credit requirement process by considering the specific properties of each request that influence the acceptance of the credit request called decision points is studied by [15]. They completed the task via discovery techniques.

The first study on fraud detection of credit applications using process mining techniques in the banking industry is done by [16]. They proposed attributes of fraud, focusing on the case and organizational perspectives through the conformance-checking task of process mining. Reference [6] modeled the standard business process with the objective of conformance checking via event logs. They claimed their method could detect anomalies in the process.

The first study on resource allocation was conducted in 2016 by [10], which improved the credit requirement process by proper resources allocated to each activity in the process using statistical techniques along with discovery techniques of the process mining approach. Moreover, they asserted that their approach is ideal for simulating processes and finding people who carry them out in the most comparable ways. Using an organizational perspective and a process perspective helped identify complex steps in the process.

Reference [7] used Hidden Markov Models and activity data from the event log. Depending on the event log, a hidden Markov model is utilized to estimate the likelihood of fraud. The approach is presented by considering the case and organizational perspective of process mining. [17] focused on the process perspective of the credit requirement and proposed resource allocation method, which can identify bottlenecks in the process by simulation.

III. Methodology

A. Defining the fitness function

The credit requirement process is improved by minimizing the fitness function. The defined KPI (Key Performance Indicator) to consider an improvement to the process which has been introduced in the fitness function is as follows:

$$\min_{x_i}\{f = a.\sum_{i=1}^{n} k_i w_i(x_i) + \sum_{i=1}^{n} x_i c_i \ \} \quad (1)$$

$$\alpha_i \le x_i \le \beta_i, \ \ \alpha_i, \beta_i \ \in N \quad \forall i = 1, \dots, n$$

Where x_i is the number of resources for $i-th$ activity in the process that the maximum (β_i) and minimum (α_i) Value of

TABLE I. COST AND THE TIME PREVIOUS STUDIES OF THE CREDIT APPLICATION OF PROCESS MINING

Research	Discovery	Conformance	Enhancement	Process Perspective	Case Perspective	Organizational Perspective	Simulation	Study Field
Rozinat et al., 2009 [14]	✓			✓	✓	✓	✓	Operational Decision
Liu et al., 2012 [11]			✓	✓	✓	✓	✓	Resource Allocation
De Leoni & van der Aalst, 2013 [15]	✓				✓			Decision Points of CR
Huda et al., 2014 [16]		✓		✓		✓		Fraud Detection
Sarno & Sinaga, 2015 [6]		✓		✓		✓		Anomaly Detection
Djedović, Žunić, & Karabegović, 2016 [10]	✓			✓		✓		Resource Allocation
Rahmawati et al., 2017 [7]	✓			✓	✓			Fraud Detection
Djedovic et al., 2018 [17]	✓			✓			✓	Resource Allocation

resources for $i-th$ activity in the process has also been taken into account. w_i is the waiting time for the $i-th$ activity, and the length of the delay is determined by the number of human resources available $w_i = w_i(x_i)$. The weight coefficient k_i is introduced to differentiate the priority of waiting time for $i-th$ activity in the process , and c_i is the unit cost of each activity in the process, which cannot be overlooked. Because the resources' spent waiting in lines are merged into one fitness function, multiplier a is added. The multiplier may be estimated using the delay value, which will result in the loss of user and cost the company money by causing the customer to leave.

B. *Methodology steps*

The methodology has 4 main steps: applying the process mining discovery algorithm, statistical analysis, simulation, and applying the TOC (Theory Of Constraint) thinking process. The first three steps were adopted from [17], except that in this research standard differential evolution (DE) algorithm has been used, and the LIFO (Last-In-First-Out) concept of the queue has been proposed to allocate resources in each stage. Table II indicates the purpose of each methodology step.

Step1: Discovering the workflow of a process with a process mining technique is employed. The Heuristic miner is used to achieve the credit requirements process. Fig 3. represents the credit requirement process. ProM 6.9 is used to extract the actual process model. The credit application comprises seven major processes, as illustrated in Figure 3. The credit card application is requested initially, followed by the acceptance of the relevant requests. Following that, the requests are collected and checked for completeness. When all of the documentation is completed,

Fig. 3. Workflow of the Credit Requirement Process

TABLE II. METHODOLOGY STEPS AND ITS PURPOSE

Step	Name of Step	Algorithm	Purpose
1	Process Mining	Heuristic Miner	To detect the workflow of the process
2	Statistical Analysis	Maximum Likelihood Estimator, Method of Moment Estimator, Cullen and Frey's graph, Anderson-Darling, Kolmogorov-Smirnov, and Carmén-von-Mises statistic test	Find Real-Time Execution of Process
3	Simulation	Differential Evolution	Optimal Human Resources Allocation
4	TOC Thinking Process	Current Reality Tree, Future Reality Tree	To resolve the identified bottleneck

the next step is to verify the applicant's creditworthiness (creditworthiness check), followed by a collateral check (Collateral check). If all goes well, the credit committee (Credit committee) determines the credit demand and evaluates requests that may be granted or refused. The final step is to go over the requirements (Requirements review). This process considers completed steps, and the total number of samples for each step is 10035.

Step 2: The second step of the methodology is the statistical analysis. This study used statistical analysis to find the distribution function and reveal the information hidden in the event log. The accumulated statistic is useful for identifying user behaviors in the process and significantly for simulation, which is explained earlier in this chapter. As the table indicates, two methods of estimation, namely "Maximum Likelihood" and "Method of Moment" has been applied to estimate the parameters of distribution in order to justify the results. Then Cullen and Frey's graph is used to identify the actual distribution function. In the final statistical analysis step, several tests were conducted to check whether the selected distribution fits our data [3]. Table III indicates the distribution probability of each activity followed by the parameter's value acquired from MLE and MOM estimation methods.

TABLE III. PROBABILITY DISTRIBUTION OF EACH ACTIVITY EXECUTION USING THE MLE AND MOM ESTIMATION METHODS

Activity Name	Probability Distribution	MLE Method	MOM Method
Instance generation	Lognormal	mean=1.136; standard deviation=1.089	mean= 1.026; st. deviation= 1.448594
Acceptance of requests	Uniform	max=19.967; min=9.017	max= 19.817; min= 9.017
Collection of Documents	Weibull	shape=98.708; scale=3.290	shape=88.895; scale=3.288
Completeness check	Uniform	max=89.933; min=45.850	max=89.710 ; min=46.245
Creditworthiness check	Normal	mean=82.813; standard deviation=7.548	mean= 82.931; st. deviation= 7.567
Collateral check	Uniform	max=74.383; min=20.133	max=74.109; min=20.619
Credit committee	Lognormal	mean=3.796; standard deviation=0.436	mean= 3.769; st. deviation= 0.463
Review of Requirements	Uniform	max=14.917; min=5.033	max=14.867; min=5.121

Step3: Simulation is the third step and is designed to allocate human resources to each stage of the process automatically. The simulation parameters are discussed by [17]. Fitness function must be considered in the simulation to guarantee optimal allocation. After analyzing the log with statistical techniques, simulation is possible. Differential Evolution (DE) is employed for this purpose.

The DE algorithm is an evolutionary optimization approach that is based on population. The population in the DE algorithm is made up of units and is the generation number. The first generation is generated randomly, with subsequent generations obtained by mutation and crossover processes. The differential evolution algorithm guides its search in the solution space by using the difference between individuals. It primarily consists of mutation, crossover, population initialization, and selection operations. The DE's main idea is to scale and differentiate two different individual vectors in the same population, then add a third individual vector to this population to generate a mutation individual vector, which is crossed with the parent individual vector with a certain probability of generating an attempted individual vector. Finally, the greedy selection is applied to the attempted individual vector and the parent individual vector, and the better individual vector is saved for the following generation [18].

There are many ways to improve business processes [19]; this study will observe different scenarios. The impact of improvement will be shown through different scenarios, which will be considered the waiting time for all activities in the process. LIFO (last-in-first-out) queue will be considered in the credit requirement process based on APQC 1 standards. Scenarios are defined by [17].

IV. DISCUSSION OF THE RESULT

A. Compare the result for each scenario

The tables below compare the execution of different scenarios using the MLE (Maximum Likelihood) and MOM (Method Of Moments) parameter estimation methods. As illustrated in Table III, the parameters of the distribution functions of each step are specified, then the main algorithm (DE) executes 10 times for each scenario to find the best solution according to the fitness function.

TABLE IV. COMPARING DIFFERENT SCENARIO'S EXECUTION WITH THE **MLE** METHOD ESTIMATION

Scenario	Best Iteration	Best Fitness	Best solution
Scenario 1	2	2041800	[8 9 8 6 8 4 3]
Scenario 2	2	4471000	[10 10 10 10 8 5 10]
Scenario 3	5	3896600	[7 9 10 9 10 5 4]
Scenario 4	8	4283000	[10 10 5 10 9 5 2]

TABLE V. COMPARING DIFFERENT SCENARIO'S EXECUTION WITH **MOM** METHOD ESTIMATION

Scenario	Best Iteration	Best Fitness	Best solution
Scenario 1	2	4094200	[5 4 8 9 10 5 8]
Scenario 2	24	4479500	[10 10 5 9 10 4 8]
Scenario 3	4	3902100	[9 9 10 10 8 3 4]
Scenario 4	14	4245100	[10 10 10 9 10 4 4]

When the best results from scenarios 1 and 2 are compared, it is possible to determine that a bottleneck occurs during the second activity in the process. A maximum number of users

[1] American Productivity & Quality Center

were allotted by setting the values $k_1 = k_2 = 2$, resulted in an increase in the number of users on the second activity. The same situation applied to Scenario 4. In scenarios 3 and 4, on the other hand, the method assigns a limited number of users to the last activity since the allocation of resources is modest when the delay of that activity is disabled by setting $k_7 = 0$.

Since distribution functions randomly generate numbers, using different strategies to determine the parameters of the functions has produced the same results. For this reason, the result obtained from implementing different scenarios is reliable.

B. *Applying TOC*

This study applies TOC to the credit requirement process in order to offer a solution for the identified bottleneck based on the data analysis and the results generated by the process mining methodologies are discussed in this section. Current Reality Tree (CRT) and Future Reality Tree (FRT) based on TOC thinking process concepts are employed to propose a solution. As the second activity identifies the bottleneck, the tools mentioned offer solutions to overcome this bottleneck.

Based on the TOC thinking process, the Current Reality Tree (CRT) is the chain of cause-and-effect relationships to identify the major issues in the second activity of the credit requirement process. To assess the current reality, the data analysis of the second activity is reviewed, which helps to answer the question, "What to change?."

The illustration in fig 4 reveals the probable chain of cause-and-effect of the existing bottleneck on the second activity, "Collection of Documents". The bottleneck is caused by the activity's sophistication (Collection of Documents). The second activity's difficulty might be due to the untrained employees assigned to it or the non-electronic collection of documents. Employees' lack of experience and abilities may be the root cause of unqualified staff. Employees may not have received adequate training. The lack of an electronic or online mechanism to gather documents may be the reason for non-electronic document collecting. The main reason for this might be an inadequate infrastructure for document collection.

It can be concluded that the likely underlying cause of the present block (undesired effect) in the second activity of the credit requirement process is unskilled staff or a lack of suitable technological infrastructure to gather documents. There is no other evidence than the data analysis and the outcome generated by the proposed method to find the root cause of the problem. Considering this as a research constraint, the figure depicts reality.

From fig 5, it can be concluded that having trained personnel may leads to having qualified personnel, and designing a proper infrastructure to collect documents online will eliminate the bottleneck. In this way, the injections made in one step of the

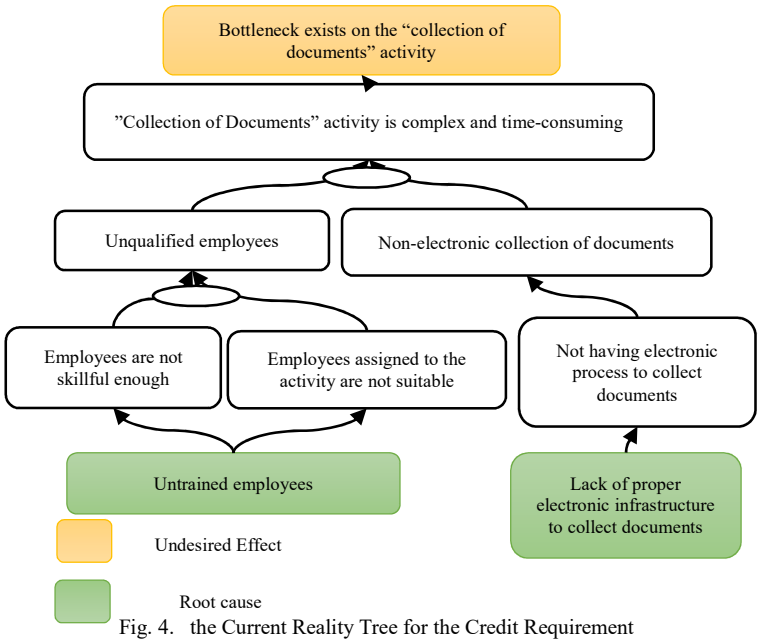

Fig. 4. the Current Reality Tree for the Credit Requirement Process_Document Collection

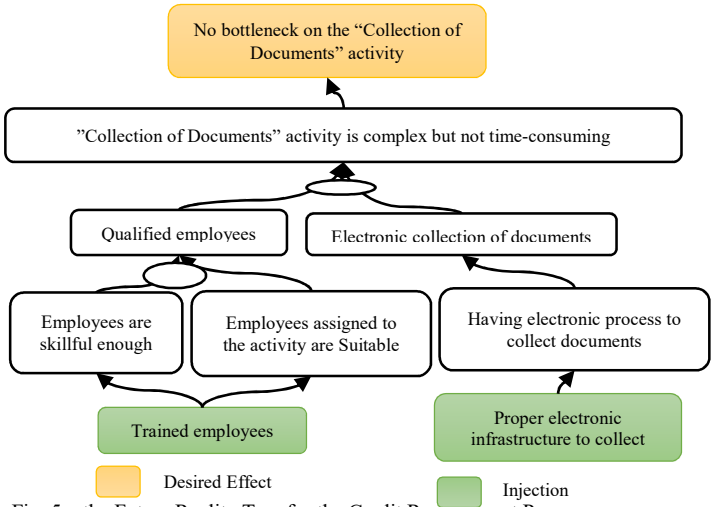

Fig. 5. the Future Reality Tree for the Credit Requirement Process, Collection of Documents

credit requirement process, which data analysis has shown to be a complex activity, will improve the process as a whole.

C. *Suggestion to Improve "Document Collection"*

Considering the future reality tree, one of the injections is "Proper Electronic Infrastructure" for document collection. Three main online processes are suggested for gathering information: 1. Identification information, 2. Incoming information, 3. judicial information. The following section; proposes the application and the services needed for developing an appropriate architecture to collect documents online.

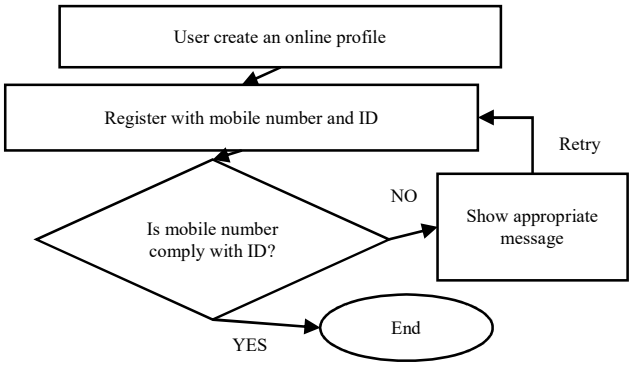

Fig. 6. Online Profile for Document Collection

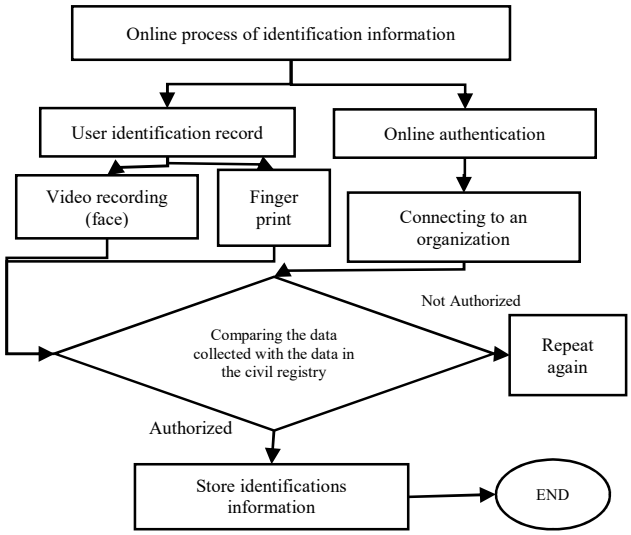

Fig. 7. Collecting User Identification Information

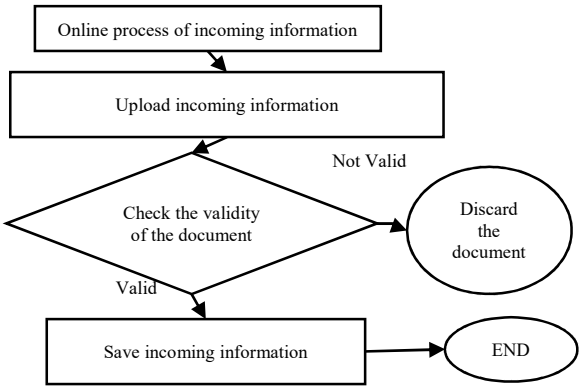

Fig. 8. Collecting User Incoming Information

V. Comparison

Resource allocation with the process mining approach is a new discipline; thus, little literature exists. Table VI represents the previous studies with their applied methods.

According to table VI, Liu et al., 2012[11] used workflow logs that provide detailed information about the process perspective, cas perspective, and organizational perspective of business processes to create event graph models. Although they proposed the resource allocation plan by defining various plans for distributing resources over the credit requirement process, the model cannot identify the bottlenecks in the process. The contribution of this study to the mentioned research is that the proposed methodology can identify as well as resolve bottlenecks. [3] applied the genetic algorithm as a resource allocation method. The contribution to this research is that in this study, the Differential Evolution (DE) algorithm is used instead as an evolutionary method for resource allocation and to find the optimal allocation of resources in fewer iterations rather than the GA.

TABLE VI. Comparison of the Literature

Research	Methods	Research Innovation Compared to the Mentioned Research
(Liu et al., 2012) [11]	Event Graph	Using Event Graph to model the credit requirement process via integrated event log information but unable to identify the bottlenecks. In contrast, this study proposed a model to identify bottlenecks and how to resolve bottlenecks in the process.
(Djedović, Žunić, Avdagić, et al., 2016) [3]	Genetic Algorithm	The genetic algorithm succeeded in locating the best answer, but as the complexity of the problem climbed, so did the number of iterations and the time required for the algorithm to complete its task. In contrast, the DE algorithm (employed in this study) can find the optimal solution in fewer iterations, so the algorithm execution time is also reduced.
(Djedovic et al., 2018) [17]	IDE	Although the IDE algorithm considered noise into account and identified the bottleneck via simulation, it was not able to resolve the bottleneck. In this study, TOC is combined with the DE algorithm to remove the bottlenecks in the credit requirement process.

It can be seen from the table that [17] used Improved Differential Evolution (IDE) and proved that the noise had been considered into account. They employed the FIFO (First-In-First-Out) concept of the queue to allocate resources to each activity. The proposed model can identify bottlenecks but not resolve that. While in this study, the LIFO concept is considered, and combine TOC thinking process to resolve the bottlenecks. Considering the research limitation into account, the result presented in this study can find the optimal solution in fewer iterations of the DE algorithm. In almost every scenario, the optimal allocation of resources is found in less than 20 iterations with a minimum value of fitness function.

VI. Conclusion

There are several strategies for allocating human resources in a process. This research adopted the LIFO approach which is new compared to the literature. Moreover, statistical analysis

with different strategies for estimating the parameters of the distribution function is designed to create a simulation model. Simulation manages to find bottleneck in the process. In addition, combining process mining with the TOC thinking process is novel for this research which helps to enhance improvement in the process. Adopting the GA (Genetic Algorithm) as a feature selection in the preprocessing stage and considering the Priority Queue concept is proposed for further research. It is also proposed to use system dynamics theory instead of TOC.

CONFLICTS OF INTEREST

The authors declare that they have no conflicts of interest.

REFERENCES

[1] Djedović, A., et al. *A rule based events correlation algorithm for process mining*. Springer.

[2] Taskesenlioglu, S., N. Ozkan, and T.G. Erdogan, *Identifying Possible Improvements of Software Development Life Cycle (SDLC) Process of a Bank by Using Process Mining.* International Journal of Software Engineering and Knowledge Engineering, 2022. **32**(04): p. 525-552.

[3] Djedović, A., et al. *Optimization of business processes by automatic reallocation of resources using the genetic algorithm*. in *2016 XI International Symposium on Telecommunications (BIHTEL)*. 2016. IEEE.

[4] dos Santos Garcia, C., et al., *Process mining techniques and applications–A systematic mapping study.* Expert Systems with Applications, 2019. **133**: p. 260-295.

[5] Hanafizadeh, P. and M. Mehrabioun, *A Systemic framework for business model design and development-Part B: Practical Perspective.* Systemic Practice and Action Research, 2020. **33**(6): p. 639-674.

[6] Sarno, R. and F.P. Sinaga. *Business process anomaly detection using ontology-based process modelling and multi-level class association rule learning*. in *2015 International Conference on Computer, Control, Informatics and its Applications (IC3INA)*. 2015. IEEE.

[7] Rahmawati, D., et al. *Fraud detection on event log of bank financial credit business process using Hidden Markov Model algorithm*. in *2017 3rd International Conference on Science in Information Technology (ICSITech)*. 2017. IEEE.

[8] Adams, N., et al., *On the Role of Process Mining in Business Process Compliance.* Available at SSRN 4081558.

[9] Huang, Z., X. Lu, and H. Duan, *Mining association rules to support resource allocation in business process management.* Expert Systems with Applications, 2011. **38**(8): p. 9483-9490.

[10] Djedović, A., E. Žunić, and A. Karabegović. *Model business process improvement by statistical analysis of the users' conduct in the process*. in *2016 International Multidisciplinary Conference on Computer and Energy Science (SpliTech)*. 2016. IEEE.

[11] Liu, Y., et al., *Workflow simulation for operational decision support using event graph through process mining.* Decision Support Systems, 2012. **52**(3): p. 685-697.

[12] de Medeiros, A.K.A., A.J. Weijters, and W.M. van der Aalst, *Genetic process mining: an experimental evaluation.* Data Mining and Knowledge Discovery, 2007. **14**(2): p. 245-304.

[13] Van Der Aalst, W., *Data science in action*, in *Process mining*. 2016, Springer. p. 3-23.

[14] Rozinat, A., et al., *Workflow simulation for operational decision support.* Data & Knowledge Engineering, 2009. **68**(9): p. 834-850.

[15] De Leoni, M. and W.M. van der Aalst. *Data-aware process mining: discovering decisions in processes using alignments*. in *Proceedings of the 28th annual ACM symposium on applied computing*. 2013.

[16] Huda, S., et al. *Identification of process-based fraud patterns in credit application*. in *2014 2nd International Conference on Information and Communication Technology (ICoICT)*. 2014. IEEE.

[17] Djedovic, A., et al., *Innovative Approach in Modeling Business Processes with a Focus on Improving the Allocation of Human Resources.* Mathematical Problems in Engineering, 2018. **2018**.

[18] Deng, W., et al., *An improved differential evolution algorithm and its application in optimization problem.* Soft Computing, 2021. **25**(7): p. 5277-5298.

[19] van der Aalst, W.M., *Process mining: a 360 degree overview*, in *Process Mining Handbook*. 2022, Springer. p. 3-34.