2017 Fifteenth International Conference on ICT and Knowledge Engineering

Using Inductive Miner to Find the Most Optimized Path of Workflow Process

Woraphan Pulsanong
Graduate School Information Technology
Siam University
Bangkok Thailand
woraphan_p@siam.edu

Parham Porouhan
Graduate School Information Technology
Siam University
Bangkok Thailand
parham@siam.edu

Sompong Tumswadi
Graduate School Information Technology
Siam University
Bangkok Thailand
sompong@siam.edu

Wichian Premchaiswadi
Graduate School Information Technology
Siam University
Bangkok Thailand
wichian@siam.edu

***Abstract*—Nowadays dealing with the concept of customer's loyalty and customer's satisfaction has allocated significant importance to itself. All businesses try to make their customers as satisfied as possible in order to create and increase the extent of their brand awareness and brand image/value. Accordingly, they consistently keep integrating their businesses with technology to help them improve their business through data mining or performance checking techniques and approaches. In this study, the main focus of the target organization is to create a web-based application with the intention of managing job-related tasks and linked data to another web-based application in order to reply customer inquiries about the progress status of the offered service in real-time. Even if a sample Company A has supported applications, the company still faces problems and challenges from those departments in which they are dealing with job-related tasks especially customer service department. The customer service department cannot always follow and track the exact job-related status of the tasks on a web-based application in a real-time manner. Typically, there would be too much waste of time for an employee to find a job status manually (i.e., such as making a call and searching for those employees whose responsibility on that job is accomplished one by one) when receiving a call from customers. To find out what happens within the workflow processes, the researchers in this study have intended to use process mining techniques so as to analyze an event log previously collected from a web-based application which supposed to manage and control job-oriented tasks. After contemplating on various methods and algorithms supported by process mining approaches, the authors of the study decided to choose Inductive Mining technique (i.e., Petri net-based graphs and models) to find out the workflow of the ongoing tasks as well as investigating the performance of each flow in detail.**

***Keywords*—Process mining, Inductive mining, Patri net**

I. INTRODUCTION

Nowadays, most of organizations try to create their own information technology to support and facilitate the processes in their business processes and organizational structure. The in-house information technology has been created for supporting business processes, which usually make the data more accurate while reducing the total response rate to deal with customers' demands, questions, concerns, difficulties and Q&A, etc. Even though organizations try to use the state-of-the-art technology to support their business progress the organizations yet face problems getting along with call center tasks, customer complaints and so on challenges. Subsequently, technology is not necessarily able to reduce the total response time needed in order to address customer satisfaction unless the business is endangered by less satisfied and more aggressive type of customers who might be troublesome for the reputation and image of any organizations.Process mining [1] techniques can be helpful to analyze work flow processes of transactions within an event log generated through an Information system. Different algorithms can lead to generation of different results that depending on the type and specifications of processes. The results and outcomes of the process mining techniques can be used to show the existing errors or redundant processes that fixing them can significantly reduce the efficiency of the workflow process in terms of time and performance.

However, the data that is used as an input to a process mining platform need to conform with at least one of these 3 contributions: (1) It should include "Activity Type", since it is crucial to know what activity has been excecuse and performed in each process, and as a result set and define it as "Case ID". (2) Date and time of the occurrence of the relevant activity, since it is important to know when the activities have happened in each process, in what order or sequence they have been executed, and as a result set and define them as "Time Stamp". (3) The

978-1-5386-2117-2/17/$31.00 ©2017 IEEE

person or system who has been responsible or in charge of performing or creating each relevant activity, and eventually define and set it as "Resource".

This research presents a novel approach to apply process mining through Inductive Miner [2] (Petri net-based) algorithm to simulate workflow process of what happened in a medical system by comparing each trace of the event log with the master process model in order to find any violence of the rules or any bottlenecks and problems within the existing system. In this research an event log file from a private organization was used and the results showed that process mining can precisely help the administrators of the organization to analyze real-time workflow processes of what happened using a web-based application and finally addressing the problems or discrepancies with the purpose of reducing the total response time needed to offer a service to customers. The proposed approach in this study resulted in increasing the extent of customers' satisfaction toward the section in that organization, as well as better management of the personnel needed or ignored in different parts of the outpatient service center.

II. LITTERATURE REVIEW

A. *Processmining[1]*

Process mining is a technique which is capable of creating models from event logs. The models can be discovered based on the real processes captured by an information system. Depending on the type of the event logs and the type of the business environment or situation, process mining supports a wide variety of algorithms in which they can be applied on varied types of data and sources. The most interesting aspect of the process mining is that the event logs can be used in order to create and generate process models/graphs that are compatible with the real-life situations and scenarios to some degree depends on the type and accuracy of the algorithm applied within the process discovery process.

B. *Disco*

One of the most popular and commonly used process mining platforms is Disco, which is an open source software developed by Fluxicon company. Disco has a user-friendly and straightforward environment, making it understandable for most of the users who have ever tried it. Disco uses Fuzzy Miner algorithm in order to analyze the imported event-logs leading to generation and simulation of the fuzzy process models in which can be investigated in terms of: (1) frequency or number of times and actions has been performed/executed, (2) duration of the time allocated between the performed tasks and actions. However, most of the time, researchers use Disco in order to pre-process, cleanse, or filter the data based on their preferences and demands. Another popular feature of Disco is ability to convert the type of the imported data into other formats such as XES and MXML supported by ProM, which is another popular process mining platform.

C. *Inductive Miner[2](Patri net)*

Inductive Miner is an improvement on the Alpha Miner and Heuristic Miner algorithms that aim to produce a sound process model. The Inductive Miner approach is based on finding a prominent split in an event log firstly, and then by detecting how these splits are related to each other, secondly, and then by connecting the splits logs together, thirdly.

III. RESEARCH METHODOLOGY

In this phase, researchers usually get the event logs previously captured and collected from a versatile variety of web-based applications in different formats or structures such as .CSV. In this study, after importing the collected CSV file into the Disco environment, a pre-processing (so-called cleansing) step was done in order to select the appropriate type of the data needed throughout the process discovery stage, while getting rid of the unwanted and redundant types of the existing data within the event logs.

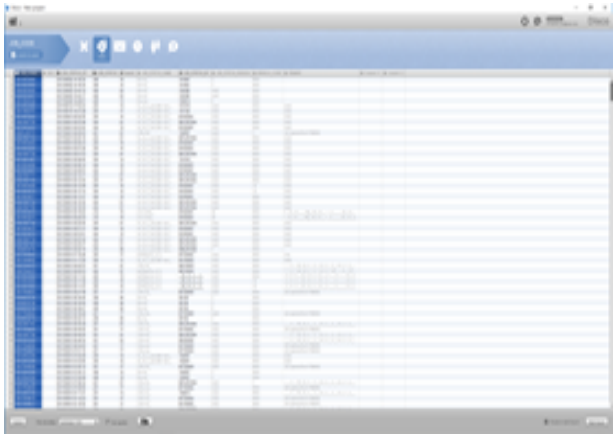

Figure 1. Importation of CSV file into Disco environment

Accordingly, after the data importation and cleansing steps, the Case ID, Activity Type, Timestamp and Resource columns were identified and selected.

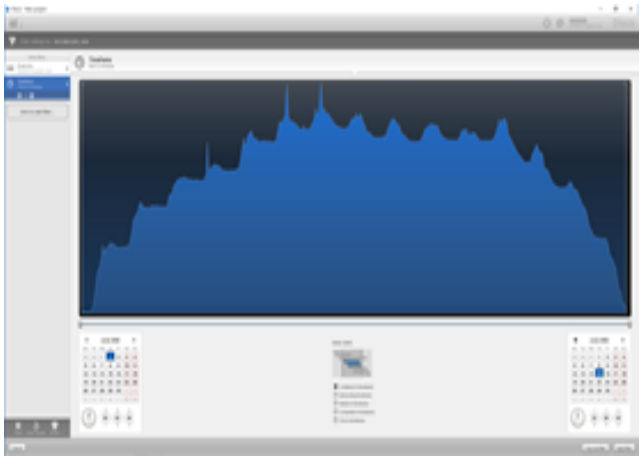

Figure 2. Timestamp filtering

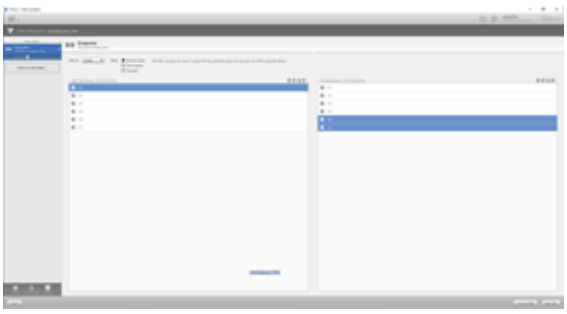

Figure 3. Endpoints filtering

Accordingly, 2 types of filtering (i.e., based on Timestamp and Endpoints) were run in Disco in order to investigate and keep in track only the desirable and useful parts of data prior to exporting theminto the Rapid Miner studio environment.

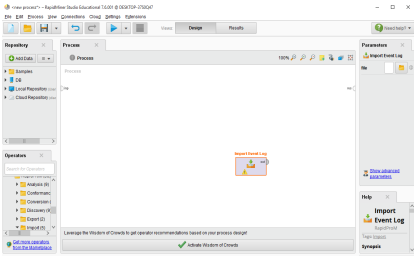

Figure 4. Importing XES file into Rapid Miner environment

Figure 5. A screenshot of the generated model obtained from Inductive Miner algorithm in RapidMiner platform

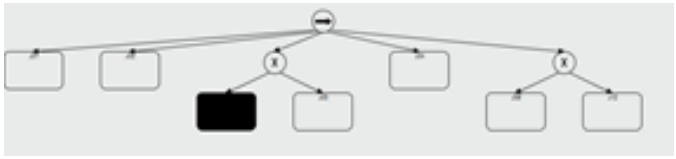

Figure 6. Model from Inductive Miner algorithm(Peri net)

The results obtained from the Inductive Miner algorithm showed the existence of four different paths of workflow. By comparing the generated paths with the performance graphs generated in Disco, more insightful information was discovered about the workflow processes in detail.

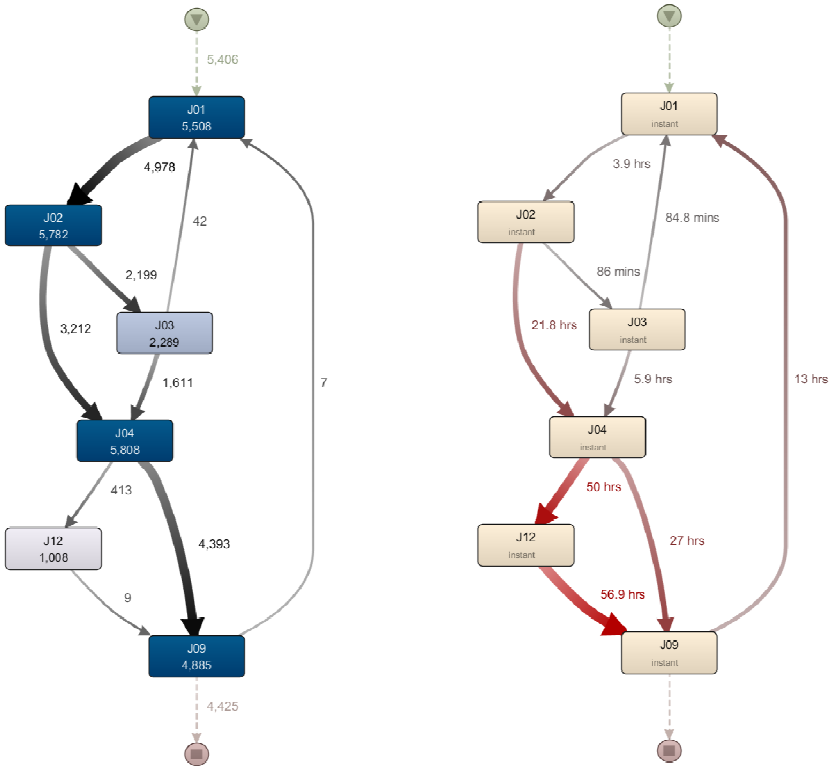

Figure 7. Workflow simulation of the processes generated by Disco fuzzy miner algorithm (left: time performance based, right: frequency performance based) as shown in the following box

1.1 J01 > J02 > J03 > J04 > J09	1.2 J01 > J02 > J03 > J04 > J12
average J01 > J09 time is 38 hours	average J01 > J12 time is 68 hours
average J01 > J04 time is 11 hours	average J01 > J04 time is 11 hours
1.3 J01 > J02 > J04 > J09	1.4 J01 > J02 > J04 > J12
average J01 > J09 time is 52.7 hours	average J01 > J12 time is 82.6 hours
average J01 > J04 time is 25.7 hours	average J01 > J04 time is 25.7 hours

By contemplating on customer's service in token "J04", it is obvious that the assigned tasks have been completed in a very slow and time consuming manner

IV. Comparasion of result

The infrastructure of the web-based application was designed in such a way that be capable of showing the links of data between two running web-based applications in real-time. The web-based application also was capable of controlling the ongoing works and tasks using mobile technology in order to keep the status of the jobs/tasks updated in real-time (i.e., the mobile technology identified and marked the status of the ongoing jobs and tasks with "J03" throughout the process workflow). In other words, the developed web-based system allowed the administrators to skip the need to follow up the latest status of tasks/jobs manually in case of the mobiles are run out or in case any loss of the signal was happening. The designed web-based application allowed the administrators to follow up the recent updates in regard to the status of "J03" automatically by adding job details directly without syncing the job from the mobile devices. Accordingly, one of the advantages of the developed system is that each user can keep manually using their mobile device, why the administrators from the headquarter operating office could track and monitor each movement and behavior of the users in an online and real-time based behavior. Overall, the data collected data from the developed web-based application included 5,508 cases in a month while only 1,611 cases showed tendencies to follow up the status job "J03". Accordingly, many of the cases have skipped the "J03" process.

The results of the study could manifest and discover the performance of the activities in each workflow in a more detailed and straightforward approach. Based on the findings of the research, the those kind of the routes of the workflow that passed through "J03" always had a shorter period than the ones in which were skipped. Based on the organization's policies and service standards all of the tasked need to include the process "J04" and need to be completed in 24 hours. If a task follows a path that includes mobile application, then the following jobs/tasks status would be updated in real-time by the help of the developed web-based application in the organization's call center. By applying this approach, the problems of the customer service department will be decreased and the duration of the response time to each customer will be much shorter and more satisfying. Obviously, such improvements will affect and improve the organization's brand's image and reputation in a more positive way.

V. Conclusion

According to the results of the study, it was understood and discovered that overall the system workflow is well related to each job-related employee, therefore the degree of support was quite reasonable. The only problem and challenge ahead was that due to the management's resistance to change, it is not always an easy task to convince the managers and administrators of the organizations and companies to use and apply a novel system similar to the one was developed and proposed in this study. Unfortunately, managers are willing to be stuck to using traditional systems and approaches, and therefore, in order to motivate them to use a new system like the one discussed in this paper, there is a need to show them the real advantages and benefits of the novel approaches to track and monitor the behaviors of the personnel in real-time situations and scenarios.

Consequently, the results of the research showed that the performance of the existing workflows will surely get better if every worker, who is in charge of an assigned job, followed every step to complete that job in a more timely manner. Flexibility of the developed system should be used only for urgent cases (J03 Should be skip in urgent case such as when got problems with mobile server, the weak of mobile signals, or ect.).If all of the users carefully and attentively follow every step in the designed system, they can decrease the duration of the long waiting time by at least 56% leading to a better efficiency and higher customer retention and satisfaction rates, as well as improvement of the brand image. Following all of the mentioned steps in job-tasked controlling in a web-based environment would make real-time status to call center web-based and decreased competing with those departments. If they work following the right path even it has 1 step more than

another it can decrease time to filled data in system. Job status will appear in another web-apps in real-time then workers need not to followed the job/task status by manual.

It will less conflict between 2 department. As a result, there is a need to enforce a more intense policy in order to force the users to fully use the system as much as possible for improving the performance of the company by not wasting any investment budget on development of lousy or impotent applications.

REFERENCES

[1] W. van der Aalst. Process Mining: Discovery, Conformance and Enhancement of Business Processes. Springer-Verlag, Berlin, 2011.

[2] Inductive Miner [online] https://www.futurelearn.com/courses/process-mining/0/steps/15642[Cited: October 7, 2017.]

[3] W. Premchaiswadi, 2015,Process Mining, Engineering Journal of Siam University, Vol.16, Issue 1, No.30, pp.1-3.

[4] Porouhan, P., & Premchaiswadi, W. (2017). Pattern mining and process modelling of collaborative interaction data in an online multi-tabletop learning environment. International Journal of Knowledge Engineering and Data Mining (IJKEDM) , 4 (2), 1-31.

[5] Porouhan, P., & Premchaiswadi, W. (2017). Process Mining and Learners' Behavior Analytics in a Collaborative and Web-Based Multi-Tabletop Environment. International Journal of Online Pedagogy and Course Design (IJOPCD) , 7 (3), 1-25.

[6] Premchaiswadi, W., & Porouhan, P. (2015). Process modeling and bottleneck mining in online peer-review systems. SpringerPlus (Computer Science) , 4 (441), 1-18.

[7] Premchaiswadi, W., & Porouhan, P. (2015). Process modeling and decision mining in a collaborative distance learning environment. Decision Analytics , 2 (6), 1-34.