Applying Process Mining to Analyze the Behavior of Learners in Online Courses

Poohridate Arpasat, Nucharee Premchaiswadi, Parham Porouhan, and Wichian Premchaiswadi

***Abstract*—The most critical challenge in analyzing the data of Massive Open Online Courses (MOOC) using process mining techniques is storing event logs in appropriate formats. In this study, an innovative approach for extraction of MOOC data is described. Thereafter, several process-discovery techniques, i.e., Dotted Chart Analysis, Fuzzy Miner, and Social Network Miner, are applied to the extracted MOOC data. In addition, behavioral studies of high- and low-performance students taking online courses are conducted. These studies considered i) overall behavioral statistics, ii) identification of bottlenecks and loopback behavior through frequency- and time-performance-based approaches, and iii) working together relationships. The results indicated that there are significant behavioral differences between the two groups. We expect that the results of this study will help educators understand students' behavioral patterns and better organize online course content.**

***Index Terms*—Process mining, event log, fuzzy miner, social network, dotted chart analysis.**

I. Introduction

Currently, many countries have recognized the importance of information and communication technology (ICT) in education and online courses). ICT has been implemented in education to help teach concepts, phenomena, and theories. Technology has broadened the way education is delivered. Traditional approaches have been supplemented or replaced by various technologies [1]. Technological developments have made it possible for teachers and students to access specialized material in multiple formats without any significant time- or space-related constraints [2]. In addition, data related to student behavior while participating in online courses will be generated and recorded in event logs.

However, there are various problems associated with the behavior of students taking online courses. For example, some students do not read all the instructions. They do not follow the intended order of the course and jump directly into the main tests or tasks. In addition, some students may face time management issues and are unable to complete the assigned projects on time. Examples of such problems and a discussion can be found in a recent study by M. Černý [1].

Furthermore, awareness of styles of teachers and students may substantially help both teaching and learning processes. Typically, approaches for statistical analysis such as that applied by Man *et al*. [3] only consider and use frequency counts to analyze the statistical behavior of students, where the data is obtained through a student query, based on a query statement with visualization output. However, the method proposed by Man *et al*. [3] and their results could not depict the learning by students from a process-centric perspective.

Accordingly, the central questions of an appropriate research study would be as follows: "What constitutes an approach that can be proposed and implemented such that both teachers and students can track all student behavior during the assigned online course?" "How using new techniques, algorithms, and platforms such as process mining can benefit both students and teachers to better self-reflect on the behavior of students both individually and collectively?"

Process mining techniques are used to extract information from event logs. Recently, event logs have become omnipresent, thereby enabling end-to-end process discovery [4], [5]. The starting point for process mining is an event log. Each event in an event log refers to an activity related to a particular case [6]. Mukala *et al*. [7] explored students' learning behavior based on the techniques of Dotted Chart Analysis and conformance diagnostics. They modeled and analyzed different study parts obtained from Massive Open Online Courses (MOOC) data, using a case study from the Eindhoven University of Technology. They could successfully locate and categorize behavioral differences among groups of students. Van den Beemt *et al*. [8] analyzed structured learning behavior in MOOC. This study provided critical insights into students' overall learning behavior and its impact on performance.

Previous studies have investigated online learning using the Disco Fuzzy Miner [9] and the Social Network Miner [10] to discover the extent to which students work together.

Using process mining to investigate learning behavior is a relatively new. To the best of our knowledge, few studies have provided results related to process mining. However, the closest one we found was a study by Bogarń *et al*. [11] that used a content assessment approach in an e-learning scenario. Their study assessed student acquisition of a core skill that played a crucial role in self-regulated learning. In other words, the main objective of their work was to identify and visualize students' self-regulated study processes throughout an e-learning course using process mining techniques and the Inductive Miner algorithm. Their research used datasets collected from 101 university students taking courses supported by the Moodle 2.0 learning platform. The dataset was collected over a single semester and contained 21,629 traces. Although the approach proposed by Bogarń *et al*. [11] and our proposed approach have many similarities

Manuscript received January 27, 2021; revised June 12, 2021.

Poohridate Arpasat, Parham Porouhan, and Wichian Premchaiswadi are with Graduate School of Information Technology, Siam University, Bangkok, Thailand (e-mail: yoisakstudio@gmail.com, parham@siam.edu, wichian@siam.edu).

Nucharee Premchaiswadi is with College of Creative Design and Entertainment Technology, Dhurakij Pundit University, Bangkok, Thailand (e-mail: nucharee@dpu.ac.th).

doi: 10.18178/ijiet.2021.11.10.1547

concerning self-regulation modeling, applying process mining, and online learning platforms, the two studies have significant differences.

Cerezo *et al*. did not apply a Social Network Miner graph/model. No time-performance analysis of the interval gaps between the performed and executed activities was applied and discussed. Their study focused on the Inductive Miner algorithm/analysis. Our study focuses on the Fuzzy Miner algorithm/analysis to investigate behavioral differences between student groups in terms of performance.

Another paper [7], used several process mining techniques, such as the Fuzzy Miner algorithm and Dotted Chart Analysis, to analyze a Coursera MOOC dataset to improve quality and delivery. However, that paper [7] did not apply any Social Network Analysis or Time-Performance Analysis of the MOOC data.

This paper aims to analyze and compare high- and low-performance study behavior using process mining based on online courses at a University in Thailand. The remainder of this paper is organized as follows. In Section II we provide background information on MOOC and process mining. In addition, we discuss our objectives and outline our approach. In Section III, we discuss applying online course events logs to process mining. Conclusions, recommendations, and suggestions for future work are presented in Section IV.

In this study, the datasets included 23 students in their final academic year (52% male, 48% female) who were taking an online web programming course as part of their main curriculum over an entire semester. We recognize that the number of the students participating in this study was not high. However, the main objective of the research was to propose and implement a new platform and approach for behavioral analysis students taking an online course; thus, the number of students involved was sufficient.

II. MOOC and Process Mining

MOOC refers to free online courses available via the Web, thereby providing learning opportunities for an unlimited number of participants In general, MOOCs have many advantages compared with existing traditional teaching/learning approaches. MOOCs can provide more interactive courses by establishing user-friendly forums or discussion areas among learners and instructors, and by providing quick feedback to online assignments and tests [12].

Process mining is a relatively new approach that can be used to improve behavior analysis based on event log data [4], [5]. Process mining is considered an excellent tool in the context of Business Process Intelligence. In general, process mining comprises three main parts, i.e., process discovery, conformance checking, and enhancements, as shown in Fig. 1. Prior to learners' behavior analysis, an event log needs to be captured from an information system in an authentic, real-world organization or business. In an educational context, the data is used to simulate students' learning activities throughout an assigned task activity to optimize efficiency and effectiveness.

Unfortunately, with existing online Learning Management Systems, instructors are not aware of students' real-time behavior. Most standard online learning systems are output-oriented, which means instructors only care about (and focus on) the correctness of the answers provided to the assigned tasks and activities. The approach proposed in this paper offers a holistic platform where lecturers can track and trace the learners' movement on each page or activity in real-time and in a more comprehensive manner. The advantages of the proposed approach are not limited to improving the quality of teaching and learning process; the proposed approach will also help identify the limitations of existing systems so that both lecturers and students can improve their performance.

In this study, our primary objectives were as follows. 1). To develop and propose an innovative approach for the extraction of appropriately formatted event logs from a MOOC system. 2). To apply process mining platforms and techniques to facilitate a process-centric analysis of the appropriately formatted and filtered data.

In this study, we were interested in using process mining to analyze student behavior based on the MOOC data format. Process mining is used to track and trace the events generated by the students during an ongoing learning process. The study also investigates the behavioral differences between groups with different performance levels, i.e., high and low performance. Various demographic metrics, e.g., students' backgrounds and learning goals [7], also can be used for behavioral analysis of learners Three process-discovery algorithms were used, i.e., i) Dotted Chart Analysis algorithm supported by ProM Software, ii) Fuzzy Miner algorithm supported by Disco Fluxicon Software [9], and iii) Social Network Miner algorithm, (based on the Working Together metric) supported by ProM Software [10], [13].

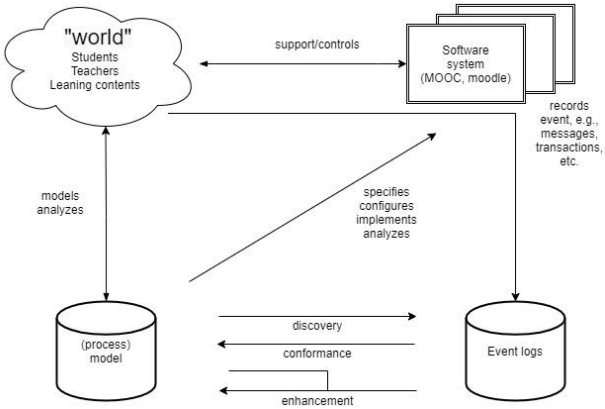

Fig. 1. Process mining model [4].

Fig. 2 shows how an ordinary school or university teacher can implement such a solution in the most popular MOOCs. As shown in Fig. 2, our proposed approach allows students to access any online course in the MOOC environment. The student behavior data in comma-separated values (CSV) format can be collected. Subsequently, the data are filtered and cleaned. Then, the prepared data are analyzed using process mining platforms, such as ProM and Disco Fluxicon, and through Fuzzy Miner, Social Network Miner, and Dotted Chart Analysis techniques.

Process mining involves capturing, collecting, and storing an event log as an input source. All process mining techniques assume that it is possible to record events

sequentially such that each event refers to an activity and is related to a particular case. Typically, a case is based on additional information, such as details about the resources, additional information about the activity type, the event timestamp, or further details about the data elements recorded in compliance with the event [14]. It is essential to capture and collect event logs based on the following conditions. i). Event logs should be saved automatically. ii) Event logs should be in an appropriate (structural) format. iii) Event logs should be adequately validated and reliable to make sure whether they contain the proper contents and values, in addition to whether they follow the appropriate structure and format. iv). Event logs should be saved in a safe/secure approach with the highest level of integrity and the least extent of bias/error. v). The highest degree of privacy should be maintained to prevent any breach of the sensitive/personal user data. vi). Event logs should contain appropriate attributes with reasonable meanings [14]. As mentioned earlier, the study's main emphasis is on behavior analysis of students through a set of previously captured event logs/data stored in the MOOC format (Fig. 3)

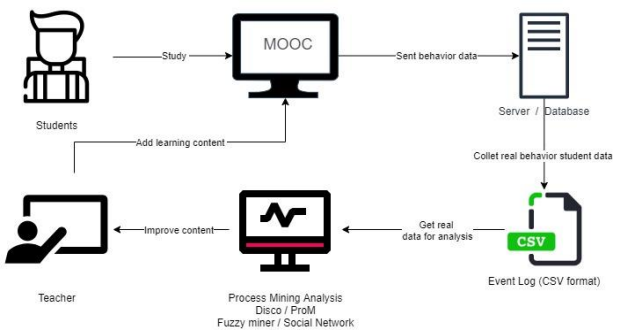

Fig. 2. Holistic view of the proposed approach showing what data are included in the reports on teaching.

USER (STUDENT)

ID	In_signature_track	Registration_time	Signature_registration_time
1	No	7 Oct '14 19:00	n/a
2	No	9 Oct '14 01:05	n/a
...			

CLICKSTRAM EVENT

ID	User_id	Event_type	timestamp	Lecture_id
25000	1	Pageview action	10 Nov '14 16:01	103
25001	1	Video action	10 Nov '14 16:03	103
25002	1	Pageview action	11 Nov '14 16:42	104
25003	2	Pageview action	11 Nov '14 02:05	103
25004	2	Pageview action	11 Nov '14 02:15	104
...				

Fig. 3. A sample event log collected from MOOC [7].

We have provided four tables that show what data stored in the MOOC format would look like.

Table I lists User (Student) Data, e.g., Student ID, First Name, Last Name, Level, Grade, Certificate, and Performance level (i.e., high or low). A sample view of the event log used in this study representing the MOOC format is shown in Fig. 3.

TABLE I: USER (STUDENT) TABLE

ID_ student	Name	Lastname	Level	Grade	Certificate
86	A student	A lastname	Senior	A	High-Performances
30	B student	B lastname	Senior	D+	Low-Performances
...	...	...	...	...	...

TABLE II: LESSON TABLE

ID_lesson	Name	Subject
Lesson1	Database	Web Programming
Lesson2	Web application	Web Programming
...	...	...

TABLE III: LEARNING EVENT LOG

ID *(Primary key)*	ID_student *(Case ID)*	Page *(Activity)*	Timestamps *(Timestamps)*	ID_Lesson *(Activity)*	Grade *(Resources)*	Certificate *(Other)*
1	30	Learning page1	28.01.2018 08:30:26	Lesson1	D+	Low- Performances
2	86	Learning page1	28.01.2018 08:29:50	Lesson1	A	High-Performances
3	86	Learning page2	28.01.2018 08:32:52	Lesson1	A	High-Performance
4	86	Learning page3	28.01.2018 08:33:47	Lesson1	A	High-Performance
5	86	Learning page1	11.02.2018 09:00:14	Lesson2	A	High-Performance
6	86	Learning page4	11.02.2018 09:01:18	Lesson2	A	High-Performance
...	...	...	...	...	...	...

TABLE IV: SOCIAL EVENTS LOG

ID_student *(Case ID)*	ID_Lesson *(Activity)*	ID_student *(Friend) (Resources)*	Timestamps *(Timestamps)*	Grade *(Case ID)*	Grade *(Friend) (Resources)*
86	Lesson1	62	28.01.2018 16:49:26	A	A
86	Lesson1	60	28.01.2018 16:49:50	A	A
30	Lesson1	29	28.01.2018 16:56:50	D+	C+
30	Lesson2	43	11.02.2018 16:25:13	D+	C+
...	...	...	...	...	...

Table II presents Lesson Data collected using the developed MOOC coursto capture learner-specific lesson-related information, e.g., Lesson ID, Lesson Name, and Subject.

Table III is related to Learning Event Logs to identify useful event data. The data in this table are essential for simulating active student behavior throughout the system (e.g., Student, Visited Page, Timestamps, Lesson ID Grade, and Certificate). A sample screenshot of the Events Table following the MOOC format is shown in Fig. 3.

Finally, Table IV is a Social Event Log. It recorded the students' ID who are working on the assigned weekly lessons with a friend in a collaborative manner.

III. Applying Process Mining to Online Course Event Logs

After generating and collecting the Learning Event and Social Event Logs, three process mining algorithms were used to simulate and model learner behavior. The Dotted Chart Analysis algorithm is supported by ProM Software [14] and the Fuzzy Miner algorithm is supported by Disco software [9]. The Social Miner Analysis algorithm, which is based on the Working Together metric, is also supported by ProM Software [10]. Subsequently, the students were divided into two groups, i.e., high- and low-performance students, to analyze student behavior associated with performance.

A. Dotted Chart

The Dotted Chart Analysis technique was applied to the Learning Event Logs to provide an overview of student behavior. Conventionally, the x-axis in the chart deals with the date on which a student has used the system, while the y-axis deals with the group type of the students. The color of the dots represents the name of each lesson, and the size of the dots represents the frequency with which a lesson has been viewed each week, as shown in Fig. 4. An analysis the generated dotted charts yielded the following results.

- When dealing with a set of lessons, students viewed the lessons in sequence and show a low tendency to review the lessons during the week.
- After completing a set of lesson but prior to taking the final exams, students reviewed the lessons. However, lessons were not review in sequence. Students tended to focus on the expected content of the final exam.
- The approach results show that the lessons' review of the general “study guidelines” before the learning lessons can affect student performance. Fig. 4 clearly illustrates the contrast analysis (frequency-based approach) between high- and low-performance students.

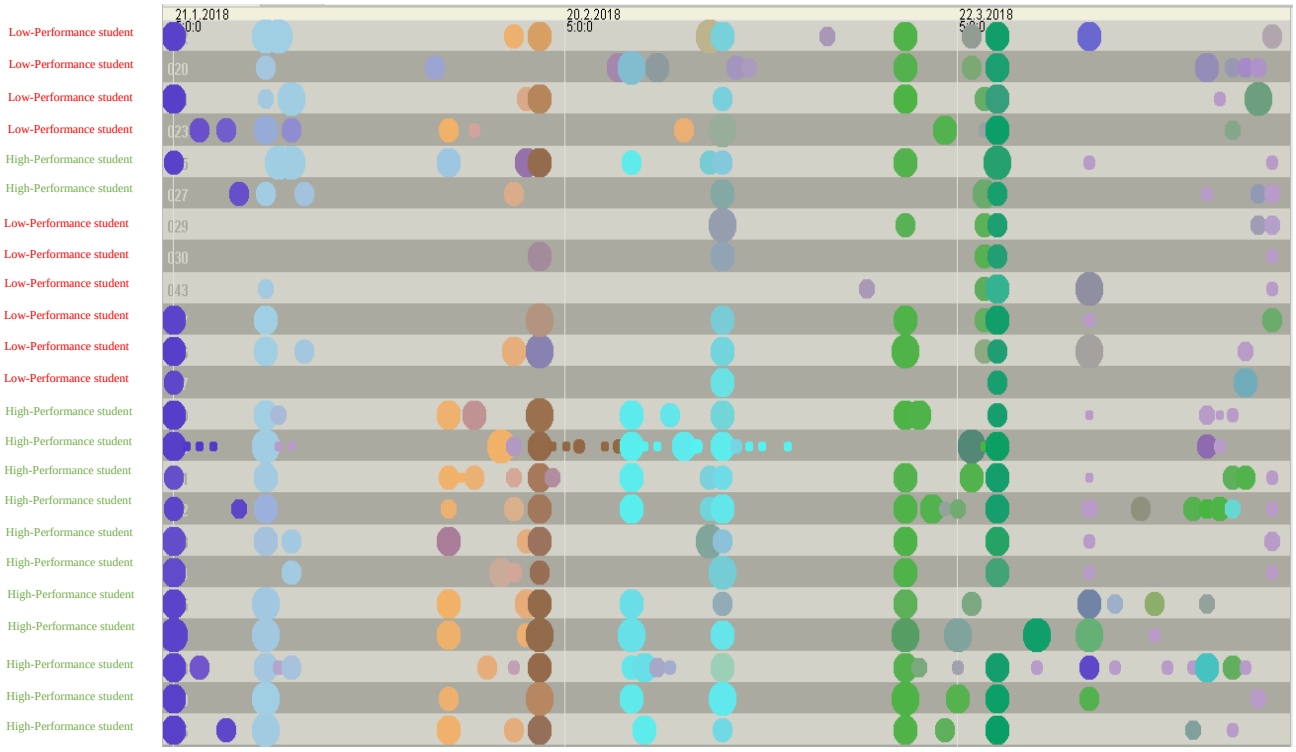

Fig. 4. The result of Dotted Chart Analysis graph to illustrate learner behavior.
Note: • Week 1 lectures • Week 2 lectures • Week 3 lectures • Week 4 lectures • Week 5 lectures • Week 6 lectures • Week 7 lectures.

B. Fuzzy Miner Model

During the learning experiment applied in this study throughout an online course taught via MOOC and after analysis of the collected/converted CSV datasets, the Fuzzy Miner algorithm supported by Disco software was used to reveal student behavior. The Fuzzy Miner algorithm was applied to Learning Event Log data to investigate event frequency and time performance (mean) for both high- and low-performance students. The results were compared and studied in detail.

Fig. 5a provides an overview of the behavioral flow (frequency-based) for high-performance learners.

Fig. 5b shows the loopback frequency for high-performance students, i.e., where a student reviews a lesson after the class. The figure indicates that students in the high-performance group frequently review lesson content after the class.

Fig. 5c shows the frequency of bottlenecks, i.e., where a high-performance student does not appear to have progressed in different topics covered in the course. The student appears

to study only specific issues rather than looking at all learning topics and materials. This may have happened because some high-performance students may have been confused by some of the content. However, they could eventually understand the problematic content and could move to the following learning content.

Overall, Fig. 6a shows a behavioral representation of high-performance learners' flow for both the frequency-based and time-performance analysis approach.

- Fig. 6b shows the average duration of the loopbacks. Based on the Fuzzy Miner results, it is evident that the loopback time for high-performance students is six days (mean). In other words, on average, it takes six days for students in that group to return and watch a lesson again.
- Fig. 6c shows the average duration of bottlenecks. Based on the Fuzzy Miner results, it is evident that the bottleneck time for the group of high-performance students is 2.2 min/page (mean). In other words, on average, high-performance students have difficulty understanding the content at a rate of 2.2 min/page (mean).

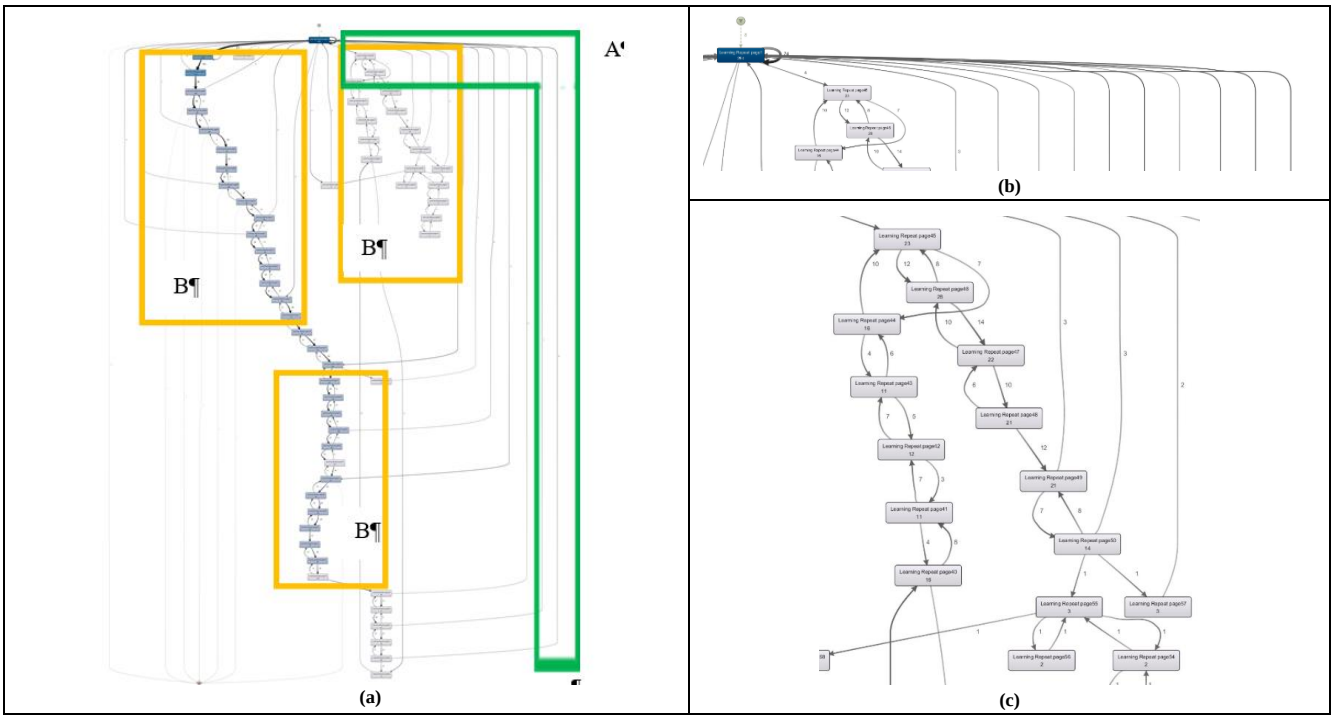

Fig. 5. (a) An illustration of the Fuzzy Miner model (frequency) of the high-performance groups of students. (b). Fuzzy Miner model (loopback frequency) of high-performance student learner behavior. (c). An illustration of the Fuzzy Miner model (bottleneck frequency) of the high-performance groups of students.

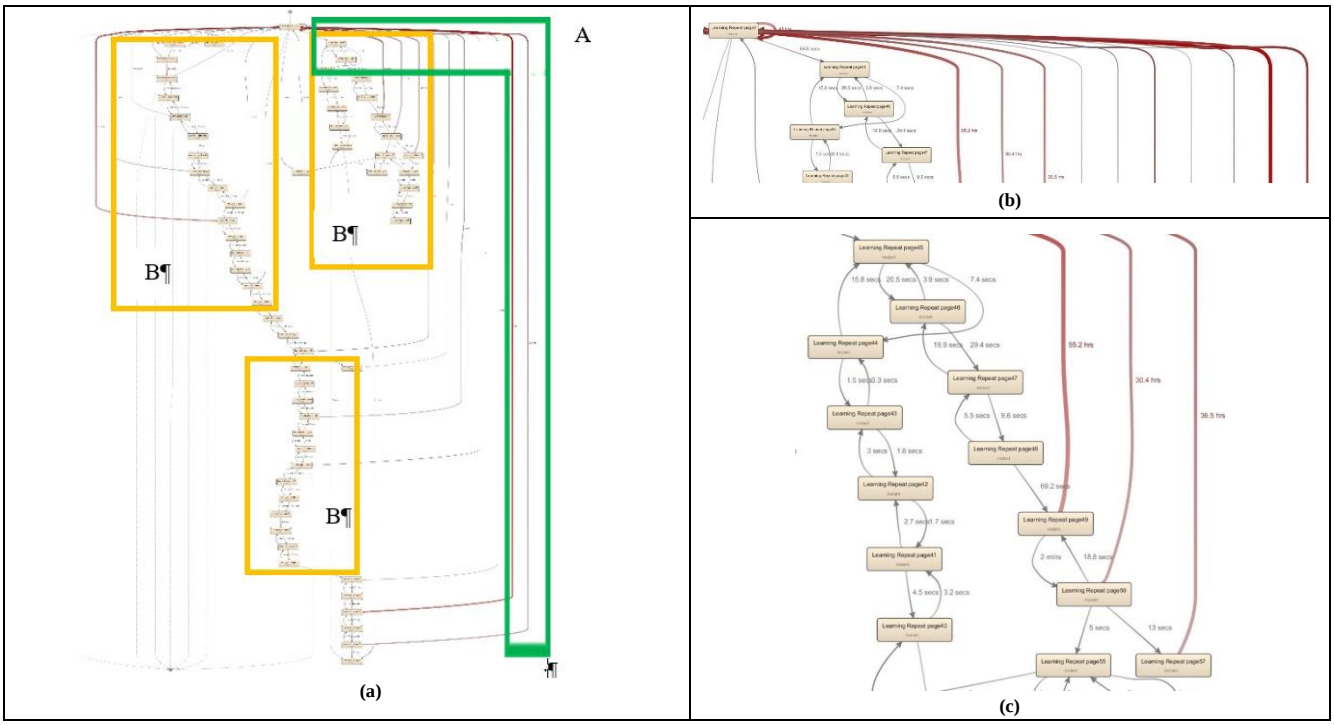

Fig. 6. (a) An illustration of the Fuzzy Miner model (time-performance) of the high-performance groups of students. (b) Fuzzy Miner model (loopback time-performance) of high-performance student learner behavior. (c) An illustration of the Fuzzy Miner model (bottleneck time-performance) of the high-performance groups of students.

Overall, Fig. 7 shows a behavioral representation of the learners' flow for both the frequency-based and time-performance analysis approach for the group of low-performance students.

- Fig. 7a shows the bottleneck frequency. As mentioned previously, this is a situation where students appear not able to progress in different topics covered in the course. The students appear to study limited/specific issues rather than looking at all the learning topics and materials. We can see more bottlenecks in the low-performance student results that the results for high-performance students.
- Fig. 7b shows that students in the low-performance group could not understand the contents of the study well. They took considerable time to complete a learning topic.
- Fig. 8a shows the average duration of the bottlenecks in the low-performance group. Bottlenecks indicates the students had difficult with the learning content. The resulting Fuzzy Miner flow indicates that, on average, low-performance students were stuck on a topic for 47 s/page (mean) during the learning process.
- Fig. 8b shows that low-performance students took an average of 33 s/page (mean) considerable time to accomplish a learning lesson/topic. Note that there was no evidence of loopback behavior among low-performance students.

Although categorizing students in high- and low-performance groups provided some insights regarding performance, we also investigated more granular categories. The dataset was divided into the following four groups (see Fig. 9).

- Group 1, students who obtained an A grade
- Group 2, students who obtained B+ and B grades
- Group 3, students who obtained C+ and C grades
- Group 4, students who obtained D+ and D grades

In this research, no student dropped out and no student failed the course.

The primary rationale behind grouping students according to grade was to investigate student performance based on four different criteria.

The data for this investigation was obtained from students at a private university in Thailand who were taking an online web programming course. In total, 23 students enrolled and passed the course within 22 weeks.

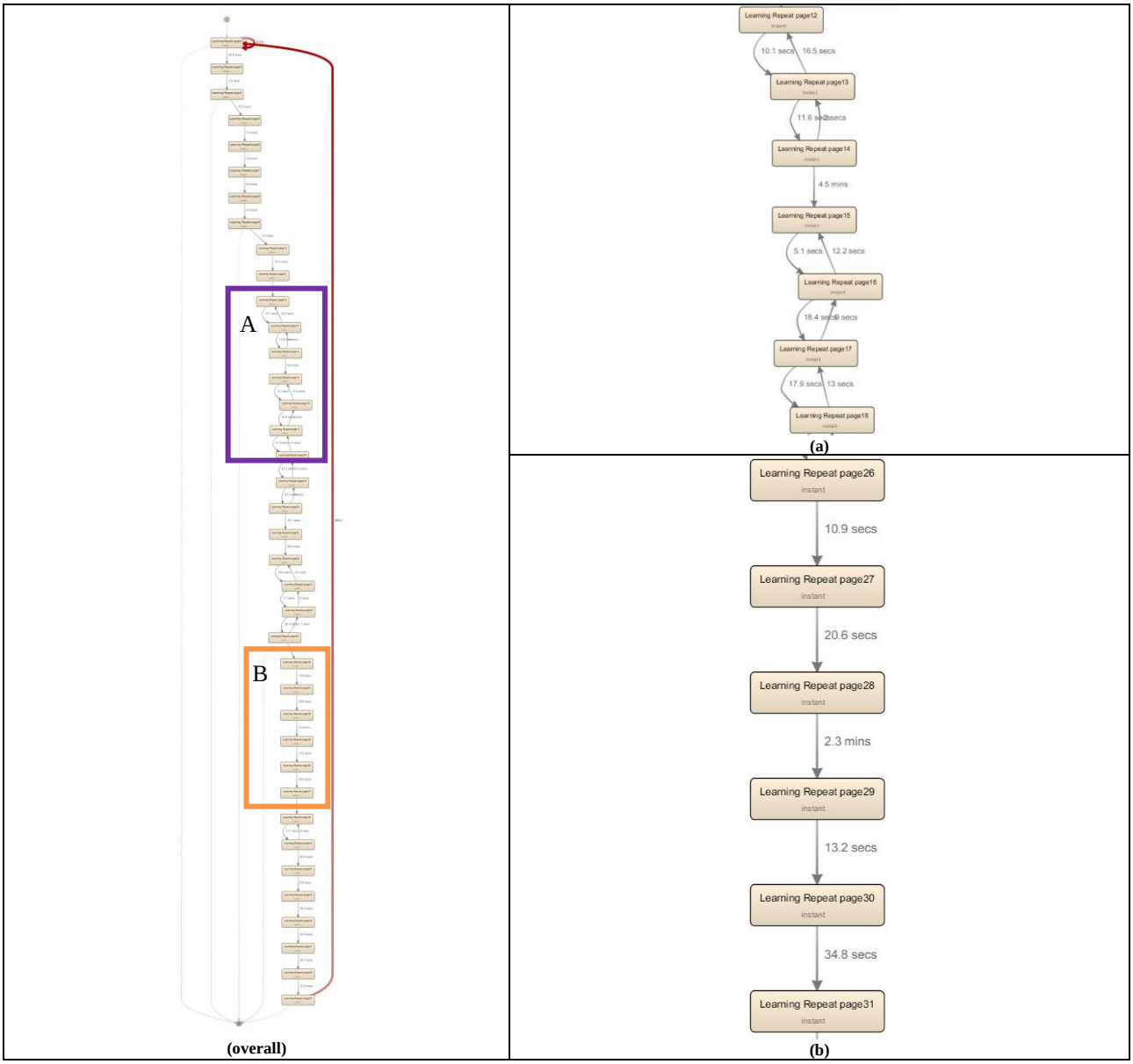

Fig. 8. Fuzzy Miner model (time performance) of low-performance student learner learners.

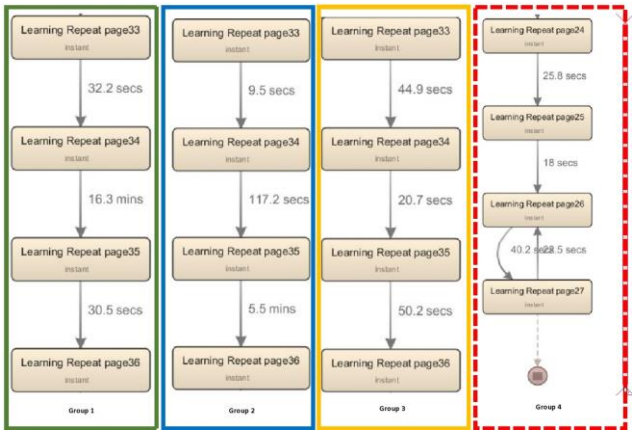

Fig. 9. A comparison and analogy between the four groups of the students based on the "time intervals" and the "sequential order" between the activities performed.

Fig. 9 contrasts and compares the four groups in terms of the "Total Duration" or "Time Interval Gaps" spent between the students' activities during the online course. In Fig. 9, Page 34 in the developed course refers to "Discussion Questions". Page 35 refers to the "Summary of The Week's Topic", and Page 36 is focused on the "Introduction to Next Week's Topic".

As shown in Figure 9, the time gap between Page 34 and Page 35 for Group 1 (i.e., students who got As) was 16.3 min. The time gap between the same two pages for Group 2 (i.e., students who got B+ and Bs) was 117.2 s, and, for Group 3 (i.e., students who got C+ and Cs) that time gap was 20.7 s. However, students in Group 4 (i.e., those who got D+ or Ds) did not read or study Page 34.

In other words, the students in Group 1 focused on the "Discussion Questions" 8.3 times more than students in Group 2 and 47.2 times more than students in Group 3. This implies that Page 34, as the "Discussion Questions" page, can be considered a key performance indicator for the students who got an A in the course. As an educator and instructor, the importance of "Discussion Questions" is evident and undeniable based on this research's findings and results. Interestingly, after studying Figure 9, it is evident that the students in Group 4 (i.e., D and D+ grades) never completed the "Discussion Questions" and "Summary of the Week's Topic", but they did not read or study the "Introduction to Next Week's Topic".

C. *Social Network Miner (Working Together)*

Finally, we used the Social Network Miner algorithm to discover working together relationships and the extent of the handover of tasks and dependencies) between high- and low-performance groups during the assigned learning course/topics (Fig. 10). In a generated Social Network Miner model/graph, the lines and arrows represent the relationships between the entities involved in the process. For example, suppose the direction of a line/arrow is from Node A to Node B. This means entity (A) has assigned/submitted a task/job to the entity (B), where this submission is based on the "Working Together metric". The circles' size representing the entities indicates the extent of the relevant entity's workload. For instance, the more workload an entity receives from another entity, the more Vertical the entity's circle will look. In contrast, the more workload an entity assigns or submits to another entity, the more Horizontal the shape of the entity's circle will appear. However, As illustrated in Fig. 10, the study results showed that high-performance student group established small teams so that fewer students can work in each group (i.e., 2–4 students/group). Also, based on Figure 10, exceot the amount of the workload, the handover of the work from one oerson to another person has been handled equally due to the circular shapes of the dots in the diagram. In contrast, the low-performance student group established a larger team so that more students can work in each group (i.e., 10 students/group). Therefore, based on the Social Network Miner findings, the size of the teams, i.e., the number of people involved in each group to accomplish an assigned learning work, affects the students' academic results.

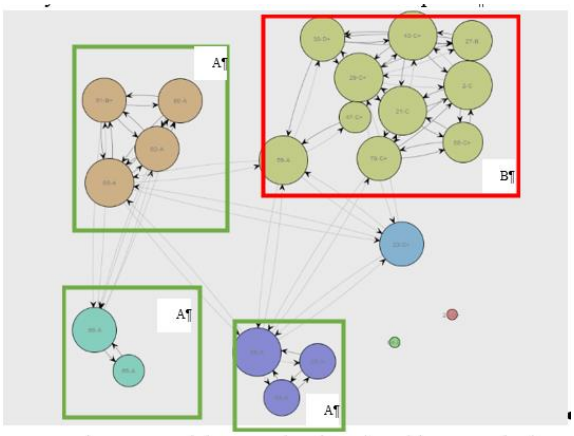

Fig. 10. Social network miner (working together). A high-performance group; B, low-performance group.

IV. Conclusions

In this study, we modeled and analyzed students' learning behavior based on MOOC format. Three process-discovery process mining techniques were used, i.e., Dotted Chart Analysis algorithm, Fuzzy Miner algorithm, and Social Network Miner (Working Together metric) algorithm. The study's main objective was to compare and study the behavioral learning differences between high- and low-performance students.

In general, the results and findings of this paper were compatible with the hypothesis of Mukala *et al*. [7], which suggested that the occurrence of loopbacks in groups of high-performance students is considerably more possible, while bottlenecks and deviations in groups of low-performance students is considerably more expected. In this work, we found that the high-performance student group showed less frequent bottlenecks during the learning process. The low-performance students showed different behavioral patterns in terms of sequence of activities performed and the time intervals/gaps between the activities executed by them. The group demonstrated a prolonged learning process through the next topics while spending considerable time on each learning content to accomplish a learning part. High-performance students established small working teams of 2–4 people/groups, while low-performance students established large working teams of 10 people.

The results of this study provide a solid basis for further and future research. The study's scope can be beneficial not

only to lecturers and students but also to curriculum developers, academic administrators, and team organizers to better track their students' real-time learning behavior, which will lead to improved efficiency and effectiveness for online learning platforms.

The paper's main emphasis was on behavioral analysis of student data collected from a MOOC environment. Therefore, the students' behavioral patterns were studied and analyzed in terms of frequency- and time-performance-based graphs, and social network graphs. The results show the relationship and dependencies among students. The dotted chart was used to illustrate the activities performed during the MOOC study. In future, we aim to show the effectiveness of the proposed approach in this study using other techniques, such as Association Rule Miner or Decision Tree Models.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

Poohridate Arpasat developed the main idea of the research and designed its test, and also contributed in the analysis of the data, Nucharee Premchaiswadi worked on the data collection and definition of the conceptual framework, Parham Porouhan contributed in the intreperetation of the findings and discussions of the study, Wichian Premchaiswadi contributed in conducting the research methodology and supervision of the research althorough the study.

References

[1] M. Černý, "Measurement of selected aspects of student's behavior in online courses," *E-Pedagogium*, vol. 19, no. 2, 2019.

[2] V. Batdi, A. Aslan, and C. Zhu, "The effect of technology supported teaching on students' academic achievement: A combined meta-analytic and thematic study," *International Journal of Learning Technology*, vol. 13, no. 1, pp. 44-60, 2018.

[3] M. Man, M. H. N. Azhan, and W. M. A. F. W. Hamzah, "Conceptual model for profiling student behavior experience in e-learning," *International Journal of Emerging Technologies in Learning*, vol. 14, no. 21, pp. 163-175, 2019.

[4] W. Van Der Aalst, "Process mining: Discovery, conformance, and enhancement of business processes," Heidelberg: Springer, vol. 2, 2011.

[5] W. M. P. Aalst, B. F. Dongen, J. Herbst, L. Maruster, G. Schimm, and A. J. M. M. Weijters, "Workflow mining: A survey of issues and approaches," *Data and Knowledge Engineering*, vol. 47, no. 1, pp. 237-267, 2003.

[6] W. Van Der Aalst, "Process mining: Overview and opportunities," *ACM Transactions on Management Information Systems (TMIS)*, vol. 3, no. 2, pp. 1-17, 2012.

[7] P. Mukala, J. C. A. M. Buijs, and W. M. P. Aalst, "Exploring students' learning behavior in MOOCs using process mining techniques," *BPM Report*, pp. 179-196, 2015.

[8] W. Aalst *et al*., "Analysing structured learning behaviour in massive open online courses (MOOCs): An approach based on process mining and clustering," *International Review of Research in Open and Distributed Learning*, vol. 19, no. 5, 2018.

[9] C. W. Günther and A. Rozinat, "Disco: Discover your processes," *BPM (Demos)*, pp. 40-44, 2012.

[10] W. M. P. Aalst, H. A. Reijers, and M. Song, "Discovering social networks from event logs," *Computer Supported Cooperative Work (CSCW)*, vol. 14, no. 6, pp. 549-593, 2005.

[11] A. B. Vega, R. C. Menéndez, and C. Romero, "Discovering learning processes using inductive miner: A case study with learning management systems (LMSs)," *Psicothema*, 2018.

[12] A. Christopoulos, N. Pellas, and M. J. Laakso, "A learning analytics theoretical framework for STEM education virtual reality applications," *Education Sciences*, vol. 10, no. 11, pp. 317, 2020.

[13] M. Song and W. M. Aalst, "Supporting process mining by showing events at a glance," *17th Annual Workshop on Information Technologies and Systems (WITS)*, pp. 139-145, 2007.

[14] W. Aalst, A. Adriansyah, A. K. A. De Medeiros, *et al*. "Process mining manifesto," *International Conference on Business Process Management*, pp. 169-194, Springer, Berlin, Heidelberg, 2012.

Copyright © 2021 by the authors. This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited (CC BY 4.0).

Poohridate Arpasat is a Ph.D. candidate in the Graduate School of Information Technology at Siam University. He received his master's degree from the Graduate School of Information Technology at Siam University, and his bachelor of business computer degree from Thonburi University. His current research areas are mainly focused on process mining applications in authentic situations and real-life scenarios.

Nucharee Premchaiswadi is an associate professor and advisor to the College of Creative Design and Entertainment Technology. She received her Ph.D. degree in electrical engineering from Waseda University, Tokyo, Japan. Her research interests are on the areas of process mining, optical character recognition, image processing, and decision support systems.

Parham Porouhan is a lecturer at the Graduate School of Information Technology, Siam University. He received his master's degree from the Graduate School of Master of Business Administration at Siam University. His areas of interest and expertise are artificial intelligence, data science, advanced data mining, process mining, process modeling, process simulation, and business process.

Wichian Premchaiswadi is a senior member of IEEE and Member of ACM, IEICE and ISACA. He is vice president of Digital Council of Thailand. He is currently the Dean and Professor of Graduate School of Information Technology at Siam University, Thailand. He also works as an assistant president at Siam University. He received his master's and Ph.D. degrees in electrical engineering from Waseda University, Tokyo, Japan. His research interests are in the areas of process mining, data mining, data analytics and high speed computing.