International Journal on Advanced Science Engineering Information Technology

Vol.10 (2020) No. 6
ISSN: 2088-5334

Student Performance Based on Activity Log on Social Network and e-Learning

Agusriandi[a], Imas Sukaesih Sitanggang[b,1], Sony Hartono Wijaya[b,2]

[a]*Universitas Muhammadiyah Enrekang, Enrekang, South Sulawesi, 91711, Indonesia*
E-mail: agusriandi595@gmail.com

[b]*Computer Science Department, Institut Pertanian Bogor, Bogor, 16680, Indonesia*
E-mail: [1]imas.sitanggang@apps.ipb.ac.id; [2]sony@apps.ipb.ac.id

***Abstract*— Learning activities in social networks and e-learning platforms bring massive activity log in the database, making it challenging to measure students' performance. Data mining technique and social network analysis provide some benefits in the field of education in discovering knowledge from hidden information of student's activities on e-learning and social network environment. This study aims to identify dominant students on social network group based on centrality values and to analyze log data from the activities on e-learning using process mining technique. Centrality value was measured by analyzing data quality or data pre-processing, creating the network, measuring the network, and highlighting degrees and layouts. The process mining technique included data pre-processing, discovering process, and conformance checking. This study found that dominant students were identified from a high hub score and authority. This study also found a free-rider student. The presence of dominant students and free-riders made the collaboration of social network group are weak. This study also found that student performance on e-learning has been discovered where the student's activity, namely, the course module viewed and course viewed, were more frequent than other activities. On the other hand, an optimum fitness value was obtained, i.e., 0.94 on all the processes of e-learning. This study provides insights that can be used to improve student collaboration and to enhance online learning activities.**

***Keywords*— activity log; e-learning; performance; process mining; social network.**

I. Introduction

Program for International Student Assessment (PISA) 2015 reports that the performance of Indonesian students in the fields of science, mathematics, and reading skills is still far below the average, which ranked the 62nd out of 70 countries [1]. Smart students who have outstanding achievements and who are literate in using technology in education are the most crucial part of education development [2]–[4]. Simultaneously, 99% of higher education institutions in some countries, including the United States, have implemented e-learning [5]. They do implement not only e-learning platforms but also learning technology tools involving social networking or Social Network-based learning (SN-Learning) [6].

Learning activity which involves social networks and the e-learning produces a massive volume of activity log in the database. These activities log can be utilized as the strategic knowledge to understand student performance [5]. Student performance can be analyzed based on data [2], social network analysis [7]–[9], and activity log analysis [10]. Social network analysis (SNA) is a structural analysis that employs graph theory [11]. In this context, SNA aims to study the pattern of social collaboration among users in education, including student, teacher, or educational institution, through structural analysis that reflects the network rather than an attribute or actor property [12]. While e-learning is a tool for conducting distance education [13], it allows students to work in e-learning platforms that produce log data in the database. Log data analysis in the e-learning platform is one way to transform the raw data into strategic knowledge [14].

Log analysis of activities on social networks aims to study patterns of social relations between actors in the context of education in students, teachers, or institutions through structural analysis that reflects the network rather than the attributes or properties of actors [15]. The use of social networks in collaboration often creates problems so that other students do not contribute equally. Therefore, identifying problems in student collaboration becomes an essential part so students can achieve common goals. Students are identified using a measurement of the degree of centrality. The degree of centrality is obtained by two approaches, global and local methods. The global method emphasizes all aspects of actor interaction (betweenness centrality), while the local method focuses on the position of the actor (degree centrality) [16].

Log analysis of activities in e-learning is one of the processes of transforming data lines into strategic knowledge that can be followed up to gain insight into business processes [17]. One of the objectives of activity log analysis is conformance checking to obtain strategic knowledge uses the ProM Framework to obtain knowledge from the activity log. The latest algorithm in the ProM Framework is Inductive Miner or IM [18].

The use of e-learning and social network in learning activity has some limitations that can trigger problems. Some students might dominate the group, while other students do not have the opportunity to contribute. While log data in e-learning has an enormous volume, it requires special techniques to obtain knowledge and data about the behavior of students regarding their learning activities. One way to identify students' activity in social networks is by using the degree centrality measurement [16], and analyzing student activities log on e-learning using process mining technique [19]. This study has two objectives (1) identifying students dominating the social network group using centrality values, (2) analyzing student performance in e-learning using process mining technique.

II. MATERIALS AND METHOD

A. Place, Data, and Tools

The data of this study were collected from social network groups, e-learning, and the grade of Applied Data Mining course offered to postgraduate students on Computer Science IPB University. The data consisted of student activities log on social networks (WhatsApp) and e-learning in the even semester of the 2017/2018 academic year.

B. Student's Analysis

The main objective of learning analysis is to understand and improve student learning and educational institutions [20]. Various analytical techniques have been used to determine student performance, such as e-learning analysis, data analysis, and social network analysis [21]. Generating predictive models is the main objective of various analytical techniques that are generally used for classifying student performance. Explained that building prediction models have tasks such as classification, regression, category, and the most commonly used is predicting students with classification [22].

C. Modeling Analysis

Learning analysis is a particular branch of academic analysis that focuses on collecting data produced by students and using predictive models. The main objective of learning analysis is to understand and improve student learning and educational institutions. The fundamental goal in the field of Educational Data Mining (EDM) is to model students. There are four paradigms of computational models for analyzing data such as statistical, Artificial Intelligence (AI), temporal, and Machine Learning [20].

The function of analysis, in general, is to make a model analysis, infrastructure analysis, and operational analysis. Types of model analysis are statistics, predictions, or data mining models that empirically come from data using statistical methods that are generally accepted. The work to produce functional models requires several analytical strategies, namely, data quality analysis, descriptive, diagnostic, predictive, and prescriptive [23], [24].

Data quality analysis is an effort to obtain data quality by measuring objectively [25]. Failure to produce good quality data at the pre-processing stage will significantly reduce the accuracy of each data analysis job. There are four core dimensions of data quality, such as data completeness, data suitability, data validity, and data accuracy. Data quality analysis was used in this study before carrying out the stages of descriptive, diagnostic, and predictive analysis [26].

The descriptive analysis produces simple standard or periodic business reporting, ad-hoc or on-demand reporting, and dynamic or interactive reporting. Furthermore, descriptive analysis is used to look at current and previous organizational c analysis is the science of identifying things that happened in the past or that are happening now.

The diagnostic analysis includes understanding the impact of input factors and operational policies. In this section, two types of approaches are carried out in diagnostic, namely log activity analysis on social networks and e-learning [23].

D. Centrality Value

The first objective of this study is to identify students who dominated the social network using the centrality value. The stages of calculating centrality value include analyzing data quality or data pre-processing, creating the network, measuring the network, and highlighting degrees and layouts. Student's conversation on WhatsApp is presented in three columns. The first column stores the conversation date, the second column represents studentx_i, and the third column represents the student $y_i$as the interaction with thex_i student.

After data pre-processing, a network was created using the R with the i-graph package, a library for network analysis. There are two principal arguments for creating a network with Igraph, namely data frame and directed value. Frame data contain a list of symbolic edges stored in two columns. The directed argument contains Boolean values (true or false) to make the directed graph. If the value is false, a non-directed graph is formed.

Network measurement is denoted as the graph $G = (V, E)$, where V is a vertex (student), and E (edge) is the connector between x_i students who interact with $y_i$students by applying the degree function [8]. The degree of a vertex is a structural property that is divided into three types; out-degree (hub), in-degree (authority), and all-degree. In this case, each student i in the network has two non-negative scores, namely authority score and hub score.

The HITS algorithm calculated the authority score and hub score of the students in the base set *I* as follows [17]:

- For every student, $i \in I$, a_i and h_i are initialized to 1.
- Repeat the following calculation until a_i and h_i of every student $i \in I$ do not change further

For every student $i \in I$,

$$a_i = \sum_{i' \in O} h_{i'}, \qquad h_i = \sum_{i' \in T} a_{i'}, \tag{1}$$

where *O* is the set of students that are in the base set *I* and interact to the student, *i* and *T* is the set of students who are in the base set *I* and interacted-to by a student *i*. Next, a_i and h_i are normalized.

$$\sum_{i\in I} a_i = \sum_{i\in I} h_i = 1 \qquad (2)$$

The next stage was to create a highlight degree and layout using the Kamada-Kawai (KK) algorithm. KK algorithm is one widely-used algorithm in making a graph [18]. Therefore, the KK algorithm was used to visualize the hub and authority values of each student. The layout shows that the differences in each student are within the social network group in graphical form.

E. Process Mining Technique

Process mining recognizes sequential patterns represented as a workflow from the activities log [10]. The workflow of process mining produces a process model that is used as a reference for analyzing all activities [19]. The stages of process mining used in this study included data pre-processing, discovering process, and conformance checking. In data pre-processing, some irrelevant attributes were removed from for process discovery.

Process discovery results in behaviors that occurred in the past (history) originating from the activity log in the form of a process model. The process model is represented by Petri net notation [28]. Petri net is a process language representing workflow and process chain based on activities involving α algorithm [22]. Petri net (N) labeling refers to equation 3 [30].

$$N = \left(P, T, \text{Ƒ}, m_o, m_f, \lambda\right) \qquad (3)$$

where P is a collection of places (e-learning page), T is a collection of activities (P, Ƒ: (P x T) ∪ (T x P) → {0,1} is a relation diagram, m_o is the first marker, m_f is the final marker, and λ is an activity class label.

The model is automatically founded by α algorithm, which is a cycle displayed in graphical form as a representative of the most common behavior. The α algorithm as follows [29] Let *L* is the activity log of students $T.\alpha(L)$ defined as follows:

- $\mathrm{T_L} = \{\mathrm{t} \in \mathrm{T} \mid \exists_{\sigma\in \mathrm{L}}\ \mathrm{t} \in \sigma \}$,
- $\mathrm{T_l} = \{\mathrm{t} \in \mathrm{T} \mid \exists_{\sigma\in \mathrm{L}}\ \mathrm{t} \in = \mathit{first}\ (\sigma) \}$,
- $\mathrm{T_o} = \{\mathrm{t} \in \mathrm{T} \mid \exists_{\sigma\in \mathrm{L}}\ \mathrm{t} \in = \mathit{last}\ (\sigma) \}$,
- $\mathrm{X_L} = \{(\mathrm{A},\mathrm{B}) \mid \mathrm{A} \subseteq \mathrm{T_L} \wedge \mathrm{A} \neq \emptyset \wedge \mathrm{B} \neq \emptyset \wedge$
- $\forall_{\mathrm{a}\in\mathrm{A}}\forall_{\mathrm{b}\in\mathrm{B}}\ \mathrm{a} \rightarrow_{\mathrm{L}} \mathrm{b} \wedge \forall_{\mathrm{a1,a2}\in\mathrm{A}}\ \mathrm{a_1}\#_{\mathrm{L}}\mathrm{a_2}$
- $\wedge\ \forall_{\mathrm{b1,b2}\in\mathrm{B}}\ \mathrm{b_1}\#_{\mathrm{L}}\mathrm{b_2}\}$,
- $\mathrm{Y_L} = \{(\mathrm{A},\mathrm{B}) \in \mathrm{X_L} | \forall_{(\mathrm{A}',\mathrm{B}')\in \mathrm{X_L}}\ \mathrm{A} \subseteq \mathrm{A}' \wedge \mathrm{B} \subseteq \mathrm{B}'$
- $\Longrightarrow (\mathrm{A},\mathrm{B}) = (\mathrm{A}',\mathrm{B}')\}$,
- $\mathrm{P_L} = \left\{(\mathrm{p_{(A,B)}} \,\middle|\, (\mathrm{A},\mathrm{B}) \in \mathrm{Y_L} \wedge \mathrm{a} \in \mathrm{A}\right\} \cup \{\mathrm{i_L}, \mathrm{o_L}\}$,
- $\mathrm{F_L} = \left\{(\mathrm{a}, \mathrm{p_{(A,B)}} \,\middle|\, (\mathrm{A},\mathrm{B}) \in \mathrm{Y_L} \wedge \mathrm{a} \in \mathrm{A}\right\} \cup$
- $\left\{(\mathrm{a}, \mathrm{p_{(A,B)}}, \mathrm{b} \,\middle|\, (\mathrm{A},\mathrm{B}) \in \mathrm{Y_L} \wedge \mathrm{b} \in \mathrm{B}\right\} \cup$

  $\{(\mathrm{i_L}, \mathrm{t}) | \mathrm{t} \in \mathrm{T_l}\} \cup \{(\mathrm{t}, \mathrm{o_L}) \mid \mathrm{t} \in \mathrm{T_o}\}$, and
- $\alpha(\mathrm{L}) = (\mathrm{P_L}, \mathrm{T_L}, \mathrm{F_L})$.

The next step in process mining was to perform conformance checking. Conformance checking was done to monitor deviations between observed behaviors in the activity log and normative process models (discovery results)[28]. Based on this method, behavioral deviations or actions that do not match the model process can be identified and analyzed[31]. The quality of the conformance checking result was measured by the fitness function, as shown in equation 4[29].

$$\mathit{fitness}\ (\sigma, N) = \frac{1}{2}\left(1 - \frac{m}{c}\right) + \left(1 - \frac{r}{p}\right) \qquad (4)$$

where σ is a trace (path) that is passed by cases (students). There are four types of calculations: p (tokens produced), c (tokens used), m (problematic tokens), and r (remaining tokens). Tokens represent a material that will be processed and accompanied by its activities. Fitness values range from 0 - 1, where 0 means very poor, and 1 means perfect. The fitness value of 1 indicates that all activity logs can be read by the inductive miner algorithm. For example, if the fitness value (L*full, N1)* = 0.90, that means 90% of activities inside *(*L*full)* can be checked by the system correctly. A model that has a high fitness value can repeat more traces in the log (Aalst 2011).

III. Results And Discussion

A. Student Performance Analysis on Social Network

The first objective of this study is to identify students dominating the social network group using centrality values. The social network used in this study is the WhatsApp group used by students to interact and be stored in the activity log. The WhatsApp group stores an activity log, containing 1,734 rows set conversations. The data are represented in three columns, namely the date of students interact, student x_i and student y_i. In this stage, the initialization of student names is also carried out to maintain the privacy of the students.

There are several stages of analysis carried out in the student activity log, namely, descriptive analysis and diagnostic analysis. The descriptive analysis presents the frequency of students interacting in the WhatsApp group, ranging from 0 - 50 times, i.e., seven students. One student shows the highest frequency of interactions between 250 - 300 times. The summary of students' interactions in the WhatsApp group measured using the Pastecs library in the R is shown in Table 1.

TABLE I
SUMMARY OF STUDENTS' INTERACTIONS ON WHATSAPP GROUP

Students	Min	Max	Sum	Median	Mean
14	2	283	1,382	64	98

Table 1 shows that the sum 1,382 means that the second number of data is in the x_i student id column and y_i student id column. Frequent interactions made by the 5 top students occurred on different dates, as shown in Table 2.

TABLE II
INTERACTIONS OF THE TOP 5 STUDENTS

Date	Number of interactions
2/16/2018	90
2/24/2018	55
2/08/2018	43
5/02/2018	39
4/25/2018	34
7/20/2018	33

Further analysis after the descriptive analysis is diagnostic. Diagnostic analysis to identify dominating students was performed based on the hub and authority (equation 1 and 2). The hub and authority values were obtained using hub. Core and authority. Core functions in the i-graph library in R. The results of hub and authority calculation programing language students are presented in Table 3.

TABLE III
HUB, AUTHORITY, AND STUDENT'S GRADE

Student ID	Hub	Authority	Grade
34140	1.00	1.00	A
34340	0.73	0.84	A
34074	0.70	0.76	AB
34086	0.63	0.82	AB
34092	0.59	0.58	B
34107	0.76	0.39	AB
34073	0.28	0.42	B
34404	0.21	0.23	A
34085	0.10	0.16	B
34132	0.19	0.15	-
34070	0.17	0.13	B
34156	0.06	0.11	AB
34075	0.02	0.03	B
34078	0.00	0.01	-

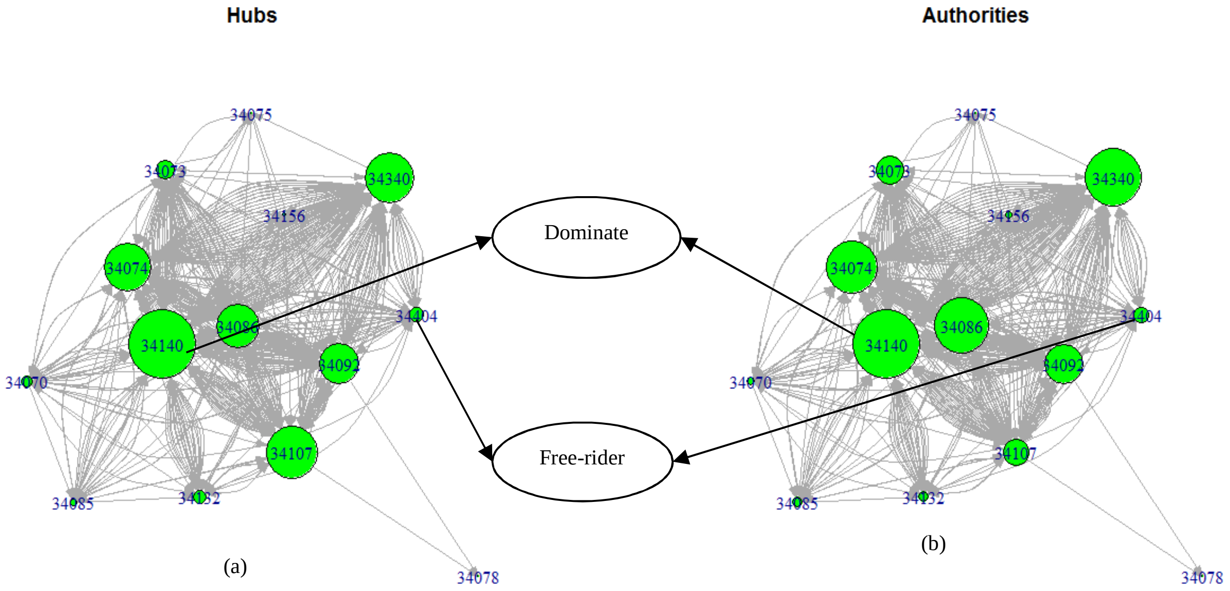


Fig. 1 Students interaction patterns on WhatsApp (a) hub of student (b) authority of student

The diagnosis shows that dominant students have higher hub and authority values. The results of the identification are displayed in highlight degrees and layouts, as shown in Fig. 1. Fig. 1 shows that student ID 34140 has the highest hub and authority values, indicating it was more dominant than other students in the WhatsApp group. The student ID 34140 got grade A on the Data Mining course (Table 3). It is assumed that the student understands the tasks presented in the e-learning platform and was able to share solutions or answer questions from members of the WhatsApp group. Thus, the dominant student dominant knew and explained the tasks. However, student ID 34404 had a hub value of 0.21, an authority value of 0.23, and grade A. This student is suspected of being a free-riders. Free-riders are those who get benefit from interactions within a group, but the member is only shared a few ideas [32].

High interaction on social networks occurs from February 6th to 22nd in 2018. The students' interaction is to know each other and ask the instruction on how to do the task. However, intense interaction was found at a certain time. It is understandable that student interactions increased on the deadline of assignment submission (close to deadlines), whereas, on other days, students only shared some information or greetings.

B. Student Performance Analysis on E-learning

The initial step in analyzing students' activities log on e-learning is pre-processing data. Activity log in e-learning contains 18,386 rows with nine columns; time, full user name, affected user, event context, component, event name, description, origin, and IP address. Data cleaning resulted in a dataset containing 2,235 rows. The data is represented in relation to three columns, namely time, user, and event name. At this stage, user initialization was carried out to protect students' privacy. Data quality analysis was conducted using RapidProM. The RapidProM describes the preliminary description of log data for the subsequent analysis process. The log data are described in Table 4.

TABLE IV
SUMMARY OF STUDENTS' ACTIVITIES LOG ON E-LEARNING

Description	Value
Number of students	14
Number of activities	2,235
Class of activities	16
Info log (start date)	23 January 2018
Info log (end date)	15 July 2018

The data used were data of students who obtained the grade A, AB, and B with amounting to 12 students, while data log contained 2,216 rows. Fig. 2 illustrates the result of the discovery process which includes the activities in e-learning; namely, a role assigned, user enrolled in course, course viewed, course module viewed, the status of the submission has been viewed, a submission has been submitted, submissions created, a file has been uploaded, a submission form viewed, and a user list viewed.

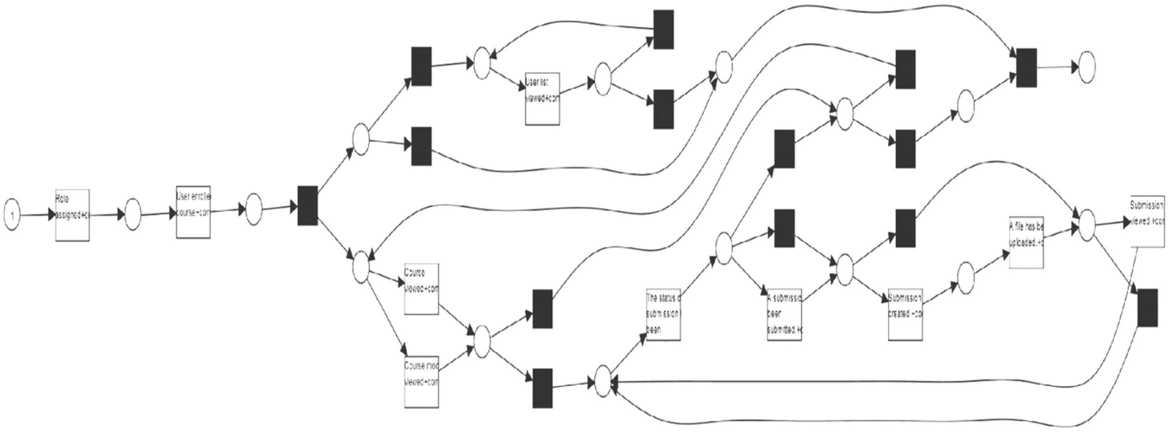
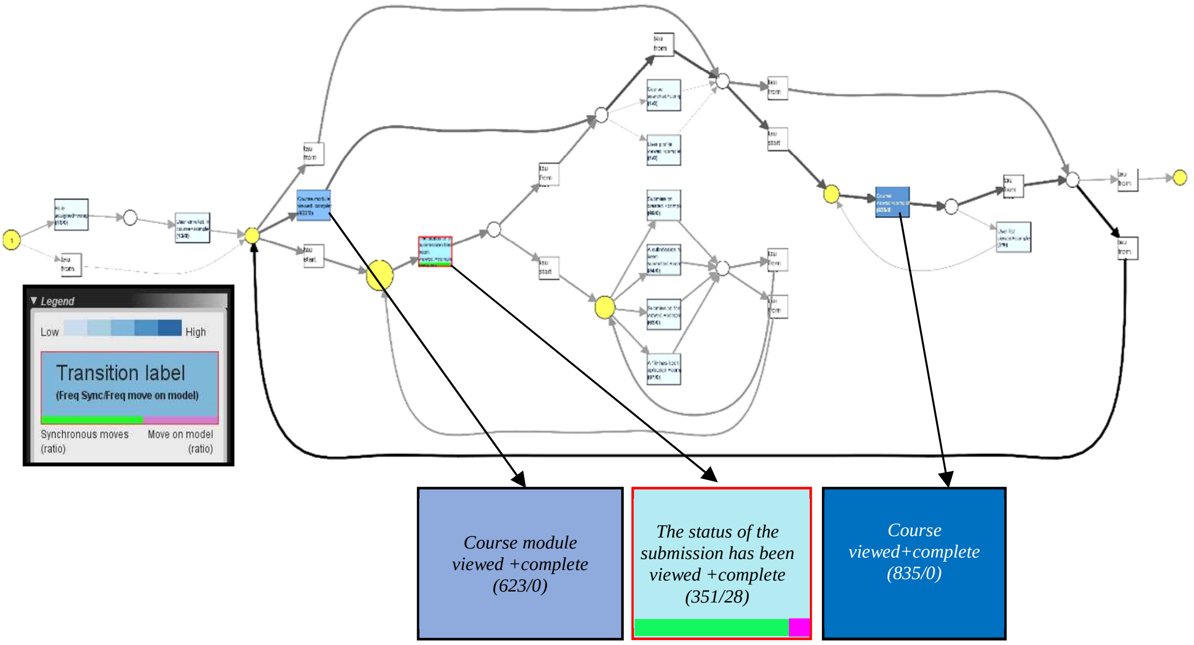


Fig. 2 The process model on e-learning



Legend
Low
High
Transition label
(Freq Sync/Freq move on model)
Synchronous moves (ratio)
Move on model (ratio)
Course module viewed +complete (623/0)
The status of the submission has been viewed +complete (351/28)
Course viewed+complete (835/0)

Fig. 3 The process model of conformance checking

Conformance checking was performed to the process model represented in the Petri net, which results are shown in Fig.3. The high frequency of student activity includes the course module viewed and the course viewed. Fig. 3 shows a synchronous activity (simultaneous), namely, the status of the submission has been viewed with a ratio of 351/28 (frequency synchronous/frequency move on model). The synchronous activity occurred as students submitted the assignments close to the deadline of submission. In this study, the fitness value for all log data is 0.94, which means that 94% of the activities in the process model could be correctly recognized by the algorithm of the inductive miner in the ProM Framework.

IV. Conclusion

This study has successfully identified a dominant student in the WhatsApp group based on hub and authority values. Besides, a free-rider was also spotted. The presence of the dominant student and free-rider has made collaboration and students' performance on social networks are weak. Student performance on e-learning has been analyzed, in which students' activities, namely, the course module viewed and course viewed, showed higher frequency than other activities. It revealed that activity log on social network and e-learning divide valuable information about the activities of the student. This information can be used to improve students' collaboration and to enhance students' activity in e-learning.

Acknowledgment

The authors would like to thank The Indonesia Endowment Fund for Education (LPDP), Ministry of Finance, the Republic of Indonesia for the research grant.

References

[1] PISA. (2015) Programme for International Student Assessment, "Pisa Results in Focus," *PISA*. [Online]. Available: https://www.oecd.org/pisa/pisa-2015-results-in-focus.pdf.

[2] A. Mueen, B. Zafar, dan U. Manzoor, "Modeling and Predicting Students' Academic Performance Using Data Mining Techniques," *Int. J. Mod. Educ. Comput. Sci.*, vol. 8, no. 11, pp. 36–42, Nov. 2016.

[3] M. Ciolacu, A. F. Tehrani, R. Beer, dan H. Popp, "Education 4.0—Fostering student's performance with machine learning methods," in *2017 IEEE 23rd International Symposium for Design and Technology in Electronic Packaging (SIITME)*, 2017, pp. 438–443.

[4] Y. Kim, "The Framework of Cloud e-Learning System for Strengthening ICT Competence of Teachers in Nicaragua," *Int. J. Adv. Sci. Eng. Inf. Technol.*, vol. 8, pp. 62–67, Feb. 2018.

[5] F. H. Wang, "An exploration of online behaviour engagement and achievement in flipped classroom supported by learning management system," *Comput. Educ.*, vol. 114, no.1 pp. 79–91, 2017.

[6] A. Krouska, C. Troussas, dan M. Virvou, "SN - Learning: An exploratory study beyond e - learning and evaluation of its applications using EV - SNL framework," *J Comput Assist Learn*, vol. 35, no.1 pp. 168–177, Oct. 2018.

[7] A. Singh, "Mining of Social Media data of University students," *Educ. Inf. Technol.*, vol. 22, no. 4, pp. 1515–1526, 2017.

[8] J.-H. Lam dan W. W. K. Ma, "When and how does learning satisfy? Working collaboratively online with a clear purpose," *Int. J. Innov. Learn.*, vol. 23, no. 4, pp. 400–415, 2018.

[9] A. E. E. Sobaih, M. A. Moustafa, P. Ghandforoush, dan M. Khan, "To use or not to use? Social media in higher education in developing countries," *Comput. Human Behav.*, vol. 58, pp. 296–305, 2016.

[10] A. Bogarin, R. Cerezo, dan C. Romero, "Discovering learning processes using Inductive Miner: A case study with Learning Management Systems (LMSs)," *Psicothema*, vol. 30, no. 3, pp. 322–329, 2018.

[11] N. F. Kolan, N. Jailani, M. Abu Bakar, dan R. Latih, "Trust Model Based on Islamic Business Ethics and Social Network Analysis," *Int. J. Adv. Sci. Eng. Inf. Technol.*, vol. 8, no. 6, pp. 2323, 2018.

[12] P. M. T. Crespo, "Social networks exploration for educational data mining," Instituto Superior Técnico, Lisboa (PT), 2013.

[13] F. Elghibari, R. Elouahbi, dan F. El Khoukhi, "Data Mining for Detecting E-learning Courses Anomalies: An Application of Decision Tree Algorithm," *Int. J. Adv. Sci. Eng. Inf. Technol.*, vol. 8, no. 3, pp. 980, 2018.

[14] I. Yurek, D. Birant, dan K. U. Birant, "Interactive process miner: a new approach for process mining," *Turk J Elec Eng Comp Sci*, vol. 26, pp. 1314–1328, 2018.

[15] J. G. Rigby, "Principals' conceptions of instructional leadership and their informal social networks: An exploration of the mechanisms of the mesolevel," *Am. J. Educ.*, vol. 122, no. 3, pp. 433–464, 2016.

[16] S. Ahajjam, M. El Haddad, dan H. Badir, "A new scalable leader-community detection approach for community detection in social networks," *Soc. Networks*, vol. 54, hal. 41–49, 2018.

[17] E. Rojas, J. Munoz-Gama, M. Sepúlveda, dan D. Capurro, "Process mining in healthcare: A literature review.," *J. Biomed. Inform.*, vol. 61, pp. 224–36, 2016.

[18] A. Bogarín, R. Cerezo, dan C. Romero, "Discovering learning processes using inductive miner: A case study with learning management systems (LMSs)," *Psicothema*, vol. 30, no. 3, pp. 322–329, 2018.

[19] J. Munoz-gama, *Conformance Checking and Diagnosis in Process Mining, Comparing Observed and Modeled Processes,* 1 ed. Chile: Springer, 2016.

[20] K. L. Vogt, "Measuring Student Engagement Using Learning Management Systems," University of Toronto, Toronto (CA), 2016.

[21] A. Brodsky, G. Shao, M. Krishnamoorthy, A. Narayanan, D. Menascé, dan R. Ak, "Analysis and optimization based on reusable knowledge base of process performance models," *Int. J. Adv. Manuf. Technol.*, vol. 88, no. 1–4, pp. 337–357, 2016.

[22] J. Hagerty, "2017 Planning Guide for Data and Analytics," *Gartner*, 2016. [Online]. Available: https://www.gartner.com/binaries/content/assets/events/keywords/catalyst/catus8/2017_planning_guide_for_data_analytics.pdf.

[23] R. Jugulum, "Importance of Data Quality for Analytics," in *Quality in the 21st Century*, Springer, 2016, pp. 23–31.

[24] D. C. Corrales, A. Ledezma, dan J. C. Corrales, "From Theory to Practice: A Data Quality Framework for Classification Tasks," *Symmetry (Basel).*, vol. 10, pp. 1–29, 2018.

[25] S.-H. Cheong dan Y.-W. Si, "Accelerating the Kamada-Kawai algorithm for boundary detection in a mobile ad hoc network," *ACM Trans. Sens. Networks*, vol. 13, no. 1, pp. 3, 2017.

[26] R. Conforti, M. Dumas, L. García-Bañuelos, dan M. La Rosa, "BPMN miner: automated discovery of BPMN process models with hierarchical structure," *Inf. Syst.*, vol. 56, pp. 284–303, 2016.

[27] W. M. P. Van der Aalst, *Process mining: data science in action,* 2 ed. London: Springer, 2016.

[28] B. Van Dongen, J. Carmona, dan T. Chatain, "A Unified Approach for Measuring Precision and Generalization Based on Anti-alignments," in *14th International Conference on Business Process Man- agement (BPM'16)*, 2016, pp. 39–56.

[29] A. Burattin, F. M. Maggi, dan A. Sperduti, "Conformance checking based on multi-perspective declarative process models," *Expert Syst. Appl.*, vol. 65, pp. 194–211, 2016.

[30] J. Siles-González dan C. Solano-Ruiz, "Self-assessment, reflection on practice and critical thinking in nursing students," *Nurse Educ. Today*, vol. 45, pp. 132–137, 2016.