



A model for collecting and analyzing action data in a learning process based on activity theory

Kyungrog Kim¹ · Nammee Moon¹

Published online: 23 December 2017
© Springer-Verlag GmbH Germany, part of Springer Nature 2017

Abstract

Research on learning analytics in technology-enhanced learning has recently been on the rise with the intent to support learners' achievements, develop personalized learning environments, and improve learning methods. Learning analytics is a data analysis method intended to help understand learners' tendencies toward activities and the significant aspects of such activities manifest in their teaching–learning process. In learning, an activity is comprised of a series of actions and then gives a couple of examples for different actions. To better understand a teaching–learning activity, the data from the stream of actions taking place need to be analyzed. To that end, this paper proposes a model for collecting and structuring the teaching–learning action data based on activity theory. The proposed model is designed to identify activities based on a series of learner actions over time. Among the components of activity theory, the model focuses on subjects, objects, and tools to collect data, which elucidates the use of tools that serve as the media between subjects (i.e., learners) and learning activities. The model offers insight into the continuity and persistence of objects or teaching–learning activity systems to better understand teaching–learning activities.

Keywords Learning analytics · Activity theory · Learning activity · Data model · Stream data

1 Introduction

With the advancement of the information and communications technologies (ICT) environment, data analytics concerning behavior and consumption by ICT users is extensively investigated (Cho and Moon 2015). Similarly, a range of efforts is exerted in technology-enhanced learning (TEL) to analyze instructional data and to apply the findings to the betterment of learning environments. The United Nations Educational, Scientific and Cultural Organization (UNESCO) in 2012 categorized instructional data analytics into learning platform analytics, predictive analytics, adaptive learning analytics, social network analytics, discourse analytics, and assessment using ICT (Shum 2012). Thus, the scope of instructional data analytics was extended.

Notably, diverse approaches have been taken in learning analytics to delve into teaching–learning activities and to characterize individual learners' interests and traits on the grounds that learning analytics is conducive to highlighting information that improves learners' achievements, develops personalized learning environments and services, helps learners maintain and improve attention and concentration, and improves learning methods or content (Drachsler and Kalz 2016).

A teaching–learning activity consists of continuous actions. To better understand teaching–learning activities, it is necessary to analyze the data from a stream of actions arising in teaching and learning environments and thus to characterize such activities. Hence, this paper proposes a data model based on activity theory to facilitate the collection and structuration of action data arising from teaching–learning activities. The proposed model allows analysis of the relationship between learners and learning media and highlights the continuity and persistence of learning activities.

This paper covers the following: Section 2 describes the theoretical rationale. Section 3 deals with the design of the teaching–learning data model based on activity theory. Sec-

Communicated by G. Yi.

✉ Nammee Moon
mnm@hoseo.edu

Kyungrog Kim
it4all@hoseo.edu

¹ Hoseo University, Asan, Chungcheongnam-do, Republic of Korea

tion 4 applies the designed model to actual data analysis. Finally, Sect. 5 presents the conclusion.

2 Literature review

2.1 Learning analytics

Learning analytics has been well documented. Still, the definition of learning analytics varies among researchers. Siemens (2010) defined learning analytics as using analysis models in order to make use of and predict intelligent data and learner-generated data, inform learning from such data, and discover informative and social connections. Long and Siemens (2011) defined learning analytics as collecting, measuring, analyzing, and reporting data about learners and learning contexts for the purpose of understanding the learning environments, optimized learning, and learners (Kim 2016).

Other definitions of learning analytics in the literature can be summarized as follows: Learning analytics is a series of processes for collecting and analyzing data about learners and learning contexts with the intent to understand the learning environments, as well as learning itself, and to provide optimal learning environments (e.g., personalized services, achievement prediction, and instruction) (Jeong 2016; ECAR-ANALYTICS Working Group 2015). Based on the foregoing definitions, this paper redefines learning analytics as a data-based decision-making process, and a cyclic process involving the extraction, storage, and analysis of data, visualization of analysis findings, prediction of future behaviors, and application of findings with a view to offering personalized and more effective learning experiences.

Table 1 outlines several frameworks established for learning analytics. Each model defines a series of processes for analyzing and applying the data arising from teaching–learning activities. In brief, the analytical models come down to a process involving Select, Capture, Aggregate & Report, Predict, Use, Refine, and Share (Jeong 2016; Lias and Elias 2011).

Data models are standardized in tandem with research on the frameworks for learning analytics. That is, the frameworks for learning analytics involving the collection, analysis, and management of data arising in the systems, content, learning tools, and users in heterogeneous environments are standardized. The standardization is led by the ISO/IEC and IMS Global (Jeong 2016; Pardos et al. 2016; Kim and Moon 2014).

- ISO/IEC JTC1 SC 36 (Information Technology for Learning, Education and Training) works on the standardization of the reference model for learning analytics by deriving use cases from the perspectives of learners, instructors, and institutions.
- The IMS Global Learning Consortium (GLC) published Learning Measurement for Analytics for learning analytics and the Learning Measurement Framework (IMS Caliper) and worked on their sophistication.

Despite the foregoing efforts, learner-oriented personalized services reflecting a learner’s characteristics (e.g., learning style) call for a data model for representing and analyzing the information of teaching–learning activities as actions. Also, the persistence and iteration of certain activities should be represented across diverse types of learning activities, which will enable the learning analytics that takes into account learners’ characteristics.

2.2 Activity theory

Activity theory is a theoretical framework intended to help understand human interactions by analyzing the use of tools and media. In activity theory, a human action system is defined as an activity. The relationship between subjects and objects is viewed as a unit and mediated by the presence of tools (Hashim and Jones 2007; Choi 2016; Uden 2006).

An activity system consists of three steps: activity, action, and operation. Specifically, a human activity exists in the form of actions. That is, actions underlie the human activ-

Table 1 Comparison of analytics frameworks and models

Knowledge continuum	Five steps of analytics	Web analytics objectives	Collective applications model	Process of learning analytics
Data	Capture	Define goals Measure	Select Capture	Select Capture
Information	Report		Aggregate	Aggregate & report
Knowledge	Predict		Process	Predict
Wisdom	Act Refine	Use	Display	Use Refine
		Share		Share

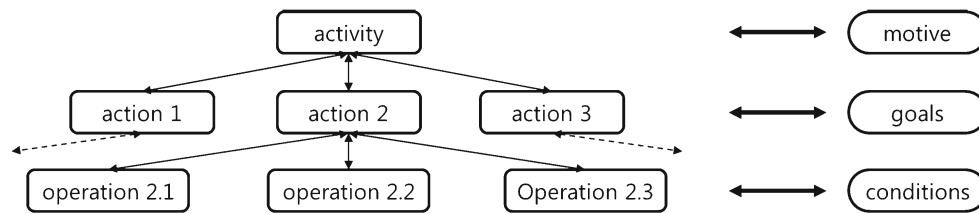


Fig. 1 Activity system

ity. Basically, actions belong to and constitute an activity. Actions convey meanings in the context of an activity. Exploring actions within an activity enables the conceptualization of human actions from the perspective of wider contexts. Also, an operation is a method of realizing an action.

According to Hashim and Jones (2007), a series of operations constitutes a higher-order action. Actions are linked to individuals' skills and knowledge. Engaging in an activity is to carry out a limited series of actions. Actions can be classified into individual or collective actions. The highest-order activity is relevant to goals and motivation. Actions are associated with goals. Operations directly depend on the conditions for the achievement of goals. Figure 1 shows how these three notions are related (Hashim and Jones 2007).

Literature on learning analytics based on activity theory in TEL environments mostly applies activity theory to lay a qualitative foundation for analysis (Xing et al. 2014; Mwalongo 2016).

Gifford and Enyedy (1999) argued that activity theory would be an appropriate framework for knowledge-generation models in computer-based collaborative learning activities. They asserted that activity theory could establish the characteristics of collaborative activities and how people could engage in social interactions based on technology, allowing the effective use of computer-supported collaborative learning activities (Hashim and Jones 2007).

Scanlon and Isroff (2005) measured the outcome in higher education by analyzing data using a framework based on activity theory on the grounds that the outcome (i.e., more learning) is underpinned by an organic combination of the components of activity theory. They defined the components of activity theory and the attributes of each component as follows: tools (learning technology), subjects (students), objects (tasks or learning situations), rules (ethics as appropriate), communities (higher education institutions), and division of labor (who controls what) (Hashim and Jones 2007).

Zurita and Nussbaum (2007) drew on activity theory to propose a conceptual framework and design for analyzing a mobile computer-supported collaborative learning system. In approaching computer-supported collaborative learning, they divided (components of) activity theory largely into roles and rules, networks, and collaborative activities. Roles

and rules involve the rules from activity theory. The networks component involves tools, subjects, and communities. Collaborative activities involve objects and division of labor. They applied these components to understand the subjects and the activities arising in relevant contexts. That is, they intended to analyze the interrelationship between the components based on activity theory (Zurita and Nussbaum 2007).

Liaw et al. (2007) investigated learners' attitude factors in e-learning systems based on activity theory. They elicited four components of learner attitudes, i.e., the learner autonomy environment, the problem-solving environment, the multimedia learning environment, and instructors' roles as supporters. They considered activity theory as an appropriate approach to address the challenges associated with e-learning systems and environments, and moreover, as a positive element for problem-solving in e-learning environments (Liaw et al. 2007).

Florian et al. (2011) proposed an activity-based learner model to monitor learners. Their activity-based learner model was intended to provide data for competency assessment applicable to adaptive learning support, evaluation, competency analysis, and recommendation. To that end, a method of building a Moodle activity–learner model was proposed based on an activity theory model and an actuator–indicator model, where social roles arising in a teaching–learning process are viewed as a principal component to provide learners and instructors with the prospect of adaptation to learning (Florian et al. 2011).

2.3 Teaching–learning activities and stream data

Engagement actions constitute a teaching–learning activity. Actions are continuous, where data arise in the form of a stream. Also, teaching–learning engagement actions tend to be iterative and continuous. Such continuous data are generated in the form of a stream, as well. To derive meaningful information from teaching–learning activity data, activity indicator components are needed. That is, to analyze a teaching–learning activity process, both activity indicators and stream data of actions are necessary. Teaching–learning activity indicators include things like the frequency and duration of a learner's access to a learning source and the learner's gender, etc. (Brooks et al. 2015).

Previous studies on teaching–learning activity indicators for learning analytics in a TEL environment are discussed below.

Ruipérez-Valiente et al. (2015) proposed an extended-learning analytics model to better understand the learning process. They built on learning analytics to support the Khan Academy platform. Specifically, their model expanded the features of the tools for visualization of the results from learning analytics, supporting the analysis of entire classes and individual learners. Ultimately, their model supported instructors and learners with decision making about learning processes. They categorized the indicators for learning analytics into six groups: total use of the platform, correct progress on the platform, time distribution for the use of the platform, gamification habits, exercise-solving habits, and affective states (Ruipérez-Valiente et al. 2015).

Mukala et al. (2015) used a fit linear model to analyze learning patterns in massive open online course (MOOC) environments. A learner visits a MOOC during a semester, eagerly engaging in video instructions and quizzes. The learner clicks to search instructional videos or quizzes, leaving a trace of click events before logging out. Such actions constitute a click stream. Click streams constitute learning via videos and quizzes. So do page views. These data indicate how the learner interacts with the instructional videos or quizzes (Mukala et al. 2015).

Gašević et al. (2016) drew on learning analytics to investigate the instructional conditions influencing the predictors of successful learning. Based on the Moodle, they classified data for learners into 12 types in order to extract the following data from each course: assignments, books, chats, course logs, feedback, forums, lightbox galleries, maps, quizzes, resources, Turnitin, and virtual classrooms. Based on these data, they performed learning analytics (Gašević et al. 2016).

3 Designing a teaching–learning data model based on activity theory

3.1 Overview

Teaching–learning in TEL environments involves participants (instructors/learners), learning objects (learning sources/learning tools), interactions, and learning outcomes. That is, learning is achieved via student–student, student–teaching, and student–learning sources/content, student–learning tools, teaching–learning sources/content, and student–system and teaching–system interactions. Such interactive actions and activities underlie the implementation of learning. The likelihood that a learner will complete the learning is predicted based on the analysis of their relationships with each component. In TEL environments, a learner often logs onto the system more than once daily. Such actions should

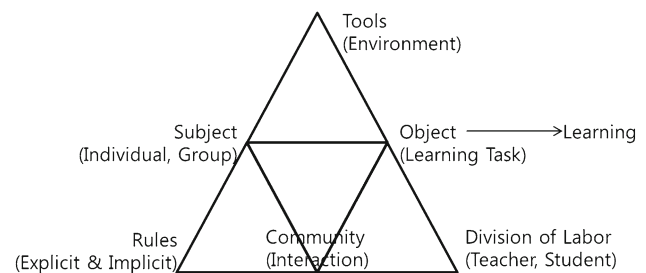


Fig. 2 Components of activity theory

be represented for learning analytics. That is, activities performed by the learners over a certain period of time (as well as each time they access the system) should be represented. To that end, this paper is based on activity theory to approach the model for data collection.

The components of activity theory applied to the proposed model include tools, subjects, objects, rules, communities, and division of labor, as shown in Fig. 2 (Hashim and Jones 2007; Xing et al. 2014, 2015; Zurita and Nussbaum 2007; Liaw et al. 2007).

Subjects are motivated to achieve goals, use tools, and engage in a learning process (objects). Here, the learning process is completed by an activity, which is comprised of continuous actions, as shown in Fig. 3 (Carvalho et al. 2015). In other words, a learning process consists of a series of actions that constitute an activity, through which learning goals are achieved.

3.2 Defining relationships between activity theory and teaching–learning activities

Data should indicate whether a learner participates in a learning activity and the extent of engagement over time. Thus, the teaching–learning data model based on activity theory starts from the relationship between activity theory components and teaching–learning components, as defined in Table 2, where primary teaching–learning attributes are linked with each component of activity theory defined (Xing et al. 2014, 2015).

Subjects are motivated to achieve learning goals by engaging in learning activities (objects) using tools. Therefore, this paper proposes the Activity-Subject, Object, Tool (A-SOT) model involving subjects, objects, and tools among other components of activity theory.

First, the term subject means the characteristics of participants in learning. Learners have their own style of learning. Styles of learning vary with the methods of perceiving information and those of processing information. The styles of learning in terms of the perception of information are divided into user groups preferring visual components (such as pictures, photos, tables, charts, and graphs) and groups

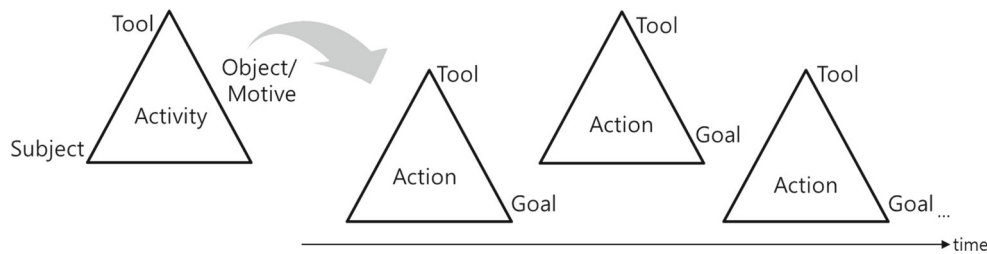


Fig. 3 Learning process based on activity theory

Table 2 Components of activity theory relevant to those of the learning activity

Division	Definition	Attribute
Tools	MOOC, LMS, resources	Online tools
Subject	Individual or group motivation for engagement, effort, and learning styles	Learning style, motive
Object	Activities performed to achieve learning goals	Learning task
Rules	Rules of learning activities	Deadline
Community	Interactions arising in learning activities	Communication tools
Division of labor	Relationships between participants and contribution	Teacher, student

preferring verbal components (such as words, sentences, and descriptions). Also, in terms of information processing, user types are divided into groups preferring practical components or active activity components (e.g., participation in a group or discussion) and groups preferring reflective or passive activity components (e.g., individual activities, attentive listening, and thinking). Classification helps to better understand users (Kim et al. 2015).

Next, objects are comprised of activities, actions, and operations to represent the system of activities. That is, a teaching–learning activity consists of a series of actions. An action may be represented as a time frame (start and end times) so that the start and end times for each action can be extracted as an action.

In TEL environments, a learner uses a range of learning sources and tools. Tools imply specific features of an LMS/MOOC in a TEL environment. Tools refers to sources or instruments for learning, e.g., instructional materials, videos, discussions, quizzes, blogs, and reports, used by a learner as the media in learning activities. Based on the tools, learning actions continuously take place. That is, a teaching–learning process is a series of actions. To understand the tools, indicators for learning tools need to be set up. In this paper, the components of the indicators for tools are limited to Notice, Video/Text Content, Grade, Group, Forum, and Blog. Figure 4 shows the A-SOT data model based on subject, object, and tool components.

To save the model in Learning Record Store (LRS) format for data, the relationship should be represented in detail. To

that end, the data are represented as a Tuple: <Attribute: Value>.

First, to extract data over time, Duration is set as Day, Week, or Month. Therefore, each time frame is stated as in formula (1).

- Time Frame = <Timestamp: Starttime-Endtime, Duration: D> (1)

The sub-components defined in the foregoing class diagram are defined in formulae (2)–(4).

- Subject = {Individual Learning Style, Group Learning Style}
Learning Style Type = {Visual(Vi)/Verbal(Ve), Active(A)/Reflective(R)} (2)

- Tool = {Notice, Video/Text Content, Grade, Group, Forum, Blog} (3)

- Object = {Activity, Action, Operation}
Action = \sum Operation = O1, ... On
Activity = \sum Action = A1, ... An (4)

Thus, the learning activity data associated with a learner’s use of tools are represented in formula (5). The components of activity theory and those of learning activities are represented in the <Attribute: Value> format, which can be used to state whether, and how often, a learner engages in teaching–learning activities.

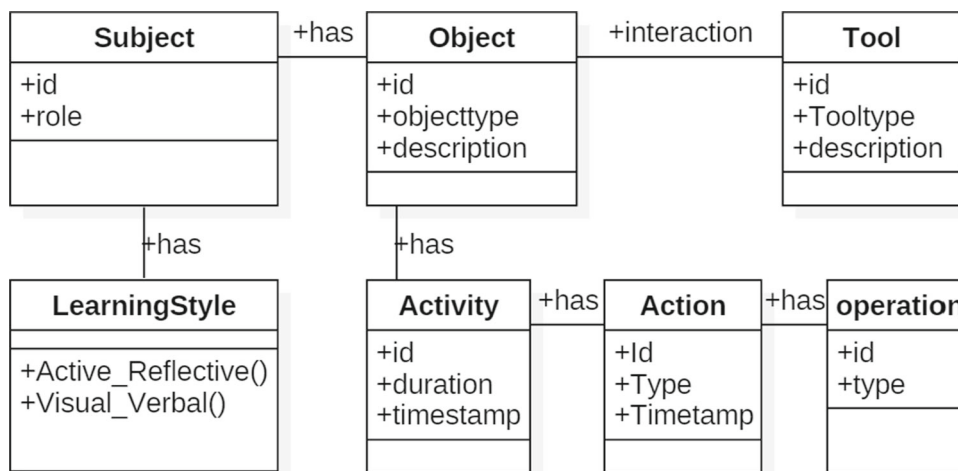


Fig. 4 Class diagram in the A-SOT data model

- <Subject: individual, Tool: Notice, Object: Number of operation, Duration: D> (5)

For example, the following string indicates the data about user 1 engaging in learning three times a day using a video object: <Subject: student1, Tool: Video Content, Object: 3, Duration: D> where teaching–learning activities are represented as a series of actions based on the data generated in a teaching–learning process. Hence, the relationships between a learner or a subject, and a tool or a medium for a learning activity, as well as the continuity of an object, or a teaching–learning activity system, are identified.

4 Application of the teaching–learning data model based on activity theory

4.1 Overview

When the teaching–learning support systems (e.g., MOOC and LMS) are used in TEL environments, a huge amount of data is generated in the teaching–learning process. Such data are not suitable for immediate use because of the differences in the types of data. Thus, collected data should be processed before the attributes of activity-based teaching–learning components can be identified quantitatively for analysis.

To analyze the data based on activity theory, the data collection process should be separated from data analysis. First, data collection is based on teaching–learning activity indicators. A learner’s click-stream data are collected. A data set is constructed based on the click-stream data. To begin with, the duration of learning is set as Day, Week, or Month to process the data. Also, the tools used are classified. Next, the analysis of data is classified into analysis of actions and

analysis of activities (Leslie et al. 2016). Analysis of actions involves analyzing a series of actions per teaching–learning tool. Analysis of activities involves cluster analysis, where comparable clusters are formed, and the use of tools is characterized per cluster.

4.2 Experimental setting

In the experiment for this paper, teaching–learning action data generated in an engineering class for junior undergraduates at H University using a supplementary teaching–learning support system are collected and analyzed. The data concern a series of learning actions via the medium of the tool in the teaching–learning support system. The daily and weekly click-stream data are processed and used as the action data.

Based on the stream data collected from the participants, the frequency and the duration (days) of engagement are preprocessed. Table 3 shows the data collected from the 47 participants.

The experimental data are standardized with the max–min method, which is widely used to compare each component, as in formula (6) Kim et al. (2015):

$$\bullet \text{ Standardization of Max – Min} = \frac{(x - \text{Min})}{(\text{Max} - \text{Min})} \quad (6)$$

4.3 Data analytics

The proposed A-SOT model is used to characterize the subjects (individuals/groups), tools (notices, content, grades, groups, forums, and blogs), and objects (actions and activities).

Figure 5 shows the results of analyzing the user action data relevant to the tool components as the media for teaching–learning. Based on the characteristics of tools, individual-activity components are separated from group-

Table 3 Experimental data set

Division	Number of actions			Dates of participation			Weeks of participation		
	Min	Max	Avg	Min	Max	Avg	Min	Max	Avg
Total actions	172	813	461.36	12	47	33.04	8	16	14.77

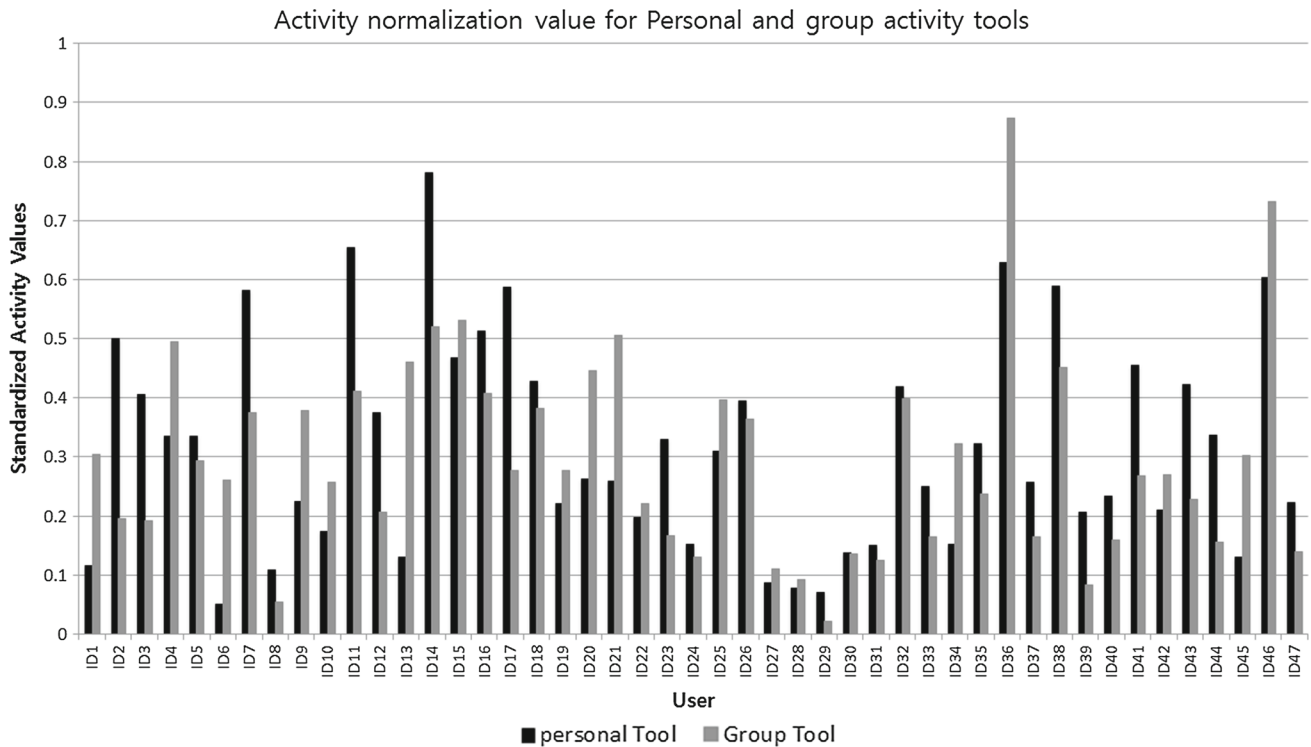


Fig. 5 Standardized data values of learning tools used by each user

activity components. The individual-activity components include the notices, content, and grades, while the group-activity components include the groups, forums, and blogs. In Fig. 5, ID29 is less active, whereas ID36 is more active in using all learning tools.

Now, user characteristics are clustered based on the tool components. The clustering methods include hierarchical and non-hierarchical methods. First, to derive the most suitable number of clusters, hierarchical clustering is used. In hierarchical clustering, the Ward method is used to determine the variance in correlation coefficients. Here, a value with a large variance is chosen. That is, $K = 5$ (K value = 46-41), meaning the ideal number of groups is 5. Next, the k -means method is used for non-hierarchical clustering.

Table 4 shows the results of clustering by the tools component for individual activities, together with the median values of sub-components in each cluster. In Table 4, Cluster 2 most actively uses individual-activity tools, such as grades, notices, and content.

Also, Table 5 shows the results of clustering by tool component for group activities together with the median values of sub-components in each cluster. In Table 5, Cluster 4 most

actively uses group-activity tools, such as groups, forums, and blogs. Cluster 3, which has the most members, uses the tools for group activities the least.

Next, the relationship between individual- and group-activity tool clusters is identified to characterize the subjects (individuals/groups) or learners who engage in learning activities using tools. Figure 6 shows the correlation between the individual- and group-activity tool clusters. Cluster 3 in the tools for group activities shows the lowest level, while playing a central role in the relationship with the individual-activity tool cluster. Cluster 1 in individual-activity tools and Cluster 5 in group-activity tools actively use both individual- and group-activity components.

Next, daily and weekly data are used for clustering so as to identify the teaching-learning activity systems (objects). This is intended to derive comparable groups based on teaching-learning activity systems. Table 6 shows the results of clustering by daily frequency of engagement. Cluster 4 shows the most active daily engagement and participates in learning several times a day.

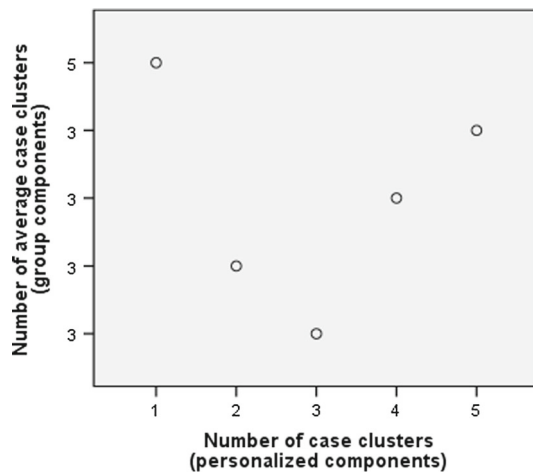
Table 7 shows the results of clustering by duration of engagement in learning in terms of the number of days. Clus-

Table 4 Clustering by tool component for individual activities

Division	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
My grade	.42308	.77472	.22727	.11111	.46923
Notice	1.00000	.34066	.32867	.15385	.18461
Content	.32337	.74884	.53162	.17286	.35326
Member	1	7	11	18	10
Action	High	Very high	Mid	Very low	Low

Table 5 Clustering by tool component for group activities

Division	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
Group	.70068	.20918	.20408	.73469	.22157
Forum	.25683	.80328	.19541	.84426	.25059
Blog	.28053	.25743	.12396	.83169	.63225
Member	9	4	25	2	7
Action	High	Low	Very low	Very high	Mid

**Fig. 6** Correlations in individual- and group-activity clusters

ter 3 regularly and continuously engages in learning activities most actively.

Based on Tables 6 and 7, the relationship between the frequency and the duration of engagement among the clusters is

Table 6 Clustering by frequency of engagement in learning

Division	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
Daily engagement	.072	.699	.480	.961	.291
Member	5	7	20	3	12
Action	Very low	High	Mid	Very high	Low

Table 7 Clustering by duration of engagement in learning (number of days)

Division	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
No. of days	.689	.517	.886	.217	.000
Member	9	20	12	5	1
Action	High	Mid	Very high	Low	Very low

determined, as shown in Fig. 7. Cluster 2 is placed in the middle in terms of duration (number of days), playing a central role in clustering by frequency of engagement. Cluster 4 in terms of duration (number of days) of engagement, and Cluster 1 in terms of frequency of engagement show the lowest levels.

In addition, the clustering relationship based on weekly data between the number of days (duration) and the frequency of engagement is shown on the left side in Fig. 8. Cluster 2 in terms of the weekly duration of engagement (number of days) is placed in the middle, playing a central role in relation to the frequency of engagement.

As mentioned before, the relationship between teaching–learning subjects, teaching–learning objects, and teaching–learning media is analyzed with clustering. That is, individual and group activities are clustered based on the characteristics of learning tools used by learners. Also, based on daily and weekly cycles and the frequency of learning activities, learning-action groups are clustered. Moreover, based on weekly cycles and frequency of learning activities, learning activity groups are clustered.

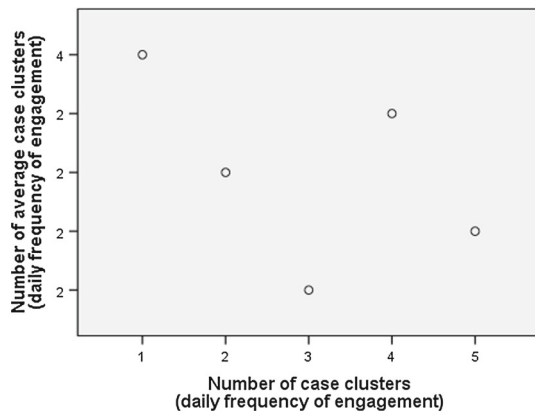


Fig. 7 Correlation between clusters by frequency and duration of engagement

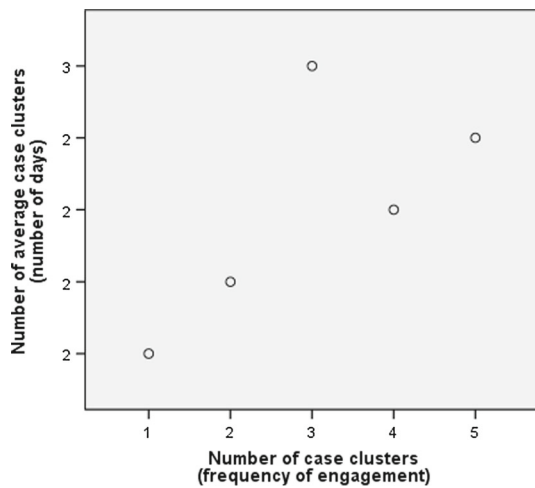


Fig. 8 Correlation between clusters by weekly frequency and duration of engagement

Users ID29 and ID36 are characterized in Table 8. ID29 hardly uses the tools for individual and group activities and rarely engages in learning, whereas ID36 actively uses the tools for individual and group activities and frequently engages in learning daily. Yet, the latter’s weekly engagement in learning is about average.

5 Conclusion

In a TEL-based teaching–learning activity, each action generates data, which are automatically accumulated in the system. The accumulated data can be restructured or repurposed. That is, such data may be used to improve learner achievements and to develop personalized learning environments and services through learning analytics, which requires efficient analytic models for the data’s collection and structuration. Hence, this paper proposes the A-SOT model (based on activity theory) for data collection. The proposed model is

Table 8 Tool–subject–object clusters by user

Division	General use of learning	Individual- activity tools	Group- activity tools	Daily engagement (frequency)	Daily engagement (# of days)	Weekly engagement (frequency)	Weekly engagement (# of days)
ID29	Lowest	Cluster 4 Very low	Cluster 3 Very low	Cluster 1 Very low	Cluster 4 Low	Cluster 1 Very low	Cluster 2 Low
ID36	Highest	Cluster 2 Very high	Cluster 4 Very high	Cluster 4 Very high	Cluster 1 High	Cluster 5 High	Cluster 1 Mid

intended to identify and use a teaching–learning process that is iteratively and continuously implemented through activities and actions.

The A-SOT model represents a teaching–learning activity as a series of actions based on data generated in the process. That is, subjects for learning have tendencies toward individual and group activities. Six components of learning tools are derived. Activity systems are defined as having operation, action, and activity steps. With the proposed model and definitions, not only the relationship between learners (subjects) and tools (media for learning activities) but the continuity and persistence of teaching–learning activity systems (objects) are also determined.

The proposed A-SOT model is applicable to clustering and sheds light on learners' tendencies toward individual and/or group activities, as well as the frequency, iteration, and continuity of learning activities. Also, the proposed model is conducive to explicating individual learners' tendencies toward learning and to predicting their achievements in learning activities.

These findings are meaningful on the following grounds. First, as a model based on activity theory and intended for data collection, the A-SOT model represents the interactions arising in TEL environments and allows extraction of the generated data in tuple format over time. Second, stream data are collected and analyzed by deriving and applying teaching–learning activity indicators. Third, a learning activity is approached and characterized in terms of action and activity. Also, learners are characterized collectively and individually in light of their tendencies toward frequency, cycle, and continuity of learning activities.

The present findings should be reinforced by further research on standardization of collected teaching–learning data, data-storage models, and teaching–learning activity indicator–profile models in TEL environments, and through replication of the proposed model. In addition, the proposed model will be upgraded continuously so it can be applied to a range of teaching–learning models for online instruction, offline supplementary instruction, and online–offline instruction. Furthermore, it will be made more sophisticated in order to predict the collective and individual achievements of learners.

Acknowledgements This work was supported by the National Research Foundation of Korea (NRF) Grant funded by the Korea government (MSIP) (No. 2016015499). Also, this work was supported by Research Projects for Senior Researchers through the Ministry of Education of the Republic of Korea and NRF (National Research Foundation of KOREA) (No. 2017-2017008886).

Compliance with ethical standards

Conflict of interest The authors declare no conflict of interest.

References

- Brooks C, Thompson C, Teasley S (2015) A time series interaction analysis method for building predictive models of learners using log data. In: Proceedings of the fifth international conference on learning analytics and knowledge. ACM, pp 126–135
- Carvalho MB et al (2015) An activity theory-based model for serious games analysis and conceptual design. *Comput Educ* 87:166–181
- Cho YS, Moon SC (2015) Recommender system using periodicity analysis via mining sequential patterns with time-series and FRAT analysis. *J Converg* 6(2):9–17
- Choi S (2016) Understanding people with human activities and social interactions for human-centered computing. *Hum Centric Comput Inf Sci* 6(1):1–10
- Drachler H, Kalz M (2016) The MOOC and learning analytics innovation cycle (MOLAC): a reflective summary of ongoing research and its challenges. *J Comput Assist Learn* 32(3):281–290
- ECAR-ANALYTICS Working Group (2015) The predictive learning analytics revolution: leveraging learning data for student success. ECAR working group paper Louisville, CO: ECAR, pp 1–23
- Elias T (2011) Learning analytics: the definitions, the processes and the potential, vol. 2016. pp 1–23
- Florian B et al (2011) Activity-based learner-models for learner monitoring and recommendations in Moodle. In: European conference on technology enhanced learning. Springer, Berlin, pp 1–14
- Gašević D, Dawson S, Rogers T, Gasevic D (2016) Learning analytics should not promote one size fits all: the effects of instructional conditions in predicting academic success. *Internet High Educ* 28:68–84
- Gifford BR, Enyedy ND (1999) Activity centered design: towards a theoretical framework for CSCL. In: Proceedings of the 1999 conference on Computer support for collaborative learning. International Society of the Learning Sciences, p 22
- Hashim NH, Jones ML (2007) Activity theory: a framework for qualitative analysis. In: The 4th international qualitative research convention (QRC), pp 1–20
- Jeong E-S (2016) Learning analysis techniques and standardization trends. *TTA J ICT Stand Certif* 164:76–82
- Kim K (2016) Learner activity modeling based on teaching and learning activities data. *KIPS Trans Softw Data Eng* 5(9):411–418
- Kim K-R, Moon N (2014) Content modeling based on social network community activity. *J Inf Process Syst* 10(2):271–282
- Kim K et al (2015) Teaching-learning activity modeling based on data analysis. *Symmetry* 7(1):206–219
- Leslie H et al (2016) Measuring completion and dropout in MOOCs: learner-centered model. In: EMOOCS, pp 55–67
- Liaw S-S, Huang H-M, Chen G-D (2007) An activity-theoretical approach to investigate learners' factors toward e-learning systems. *Comput Hum Behav* 23(4):1906–1920
- Long PD, Siemens G (2011) Penetrating the fog: analytics in learning and education. *EDUCAUSE Rev* 46(5):30
- Mukala P, Buijs JCAM, van der Aalst WMP (2015) Uncovering learning patterns in a MOOC through conformance alignments. Technical report, Eindhoven University of Technology, BPM Center Report BPM-15-09, BPMcenter.org, pp 1–20
- Mwalongo AI (2016) Using activity theory to understand student teacher perceptions of effective ways for promoting critical thinking through asynchronous discussion forums. In: Activity theory in education, sensepublishers, pp 19–34
- Pardos ZA, Whyte A, Kao K (2016) moocRP: enabling open learning analytics with an open source platform for data distribution, analysis, and visualization. *Technol Knowl Learn* 21(1):75–98
- Ruipérez-Valiente JA et al (2015) ALAS-KA: a learning analytics extension for better understanding the learning process in the Khan Academy platform. *Comput Hum Behav* 47:139–148

- Scanlon E, Issroff K (2005) Activity theory and higher education: evaluating learning technologies. *J Comput Assist Learn* 21(6):430–439
- Shum SB (2012) Learning analytics. UNESCO Institute for Information Technologies in Education Policy Brief 1–12
- Siemens G (2010) What are learning analytics. *ELEARNSPACE: learning, networks, knowledge, technology, community*. <http://www.elearnspace.org/blog/2010/08/25/whatare-learning-analytics/>
- Uden L (2006) Activity theory for designing mobile learning. *Int J Mob Learn Organ* 1(1):81–102
- Xing W, Wadholm B, Goggins S (2014) Learning analytics in CSCL with a focus on assessment: an exploratory study of activity theory-informed cluster analysis. In: Proceedings of the fourth international conference on learning analytics and knowledge. ACM, pp 59–67
- Xing W et al (2015) Group learning assessment: developing a theory-informed analytics. *Educ Technol Soc* 18(2):110–128
- Zurita G, Nussbaum M (2007) A conceptual framework based on activity theory for mobile CSCL. *Br J Educ Technol* 38(2):211–235