VISUALIZATION OF STUDENTS' ACTIVITIES TO SUPPORT TEACHERS AND STUDENTS IN PERSONAL LEARNING ENVIRONMENTS

Bo Yu
Shanghai Jiao Tong University
Haoran Building, 6/F, 1954 Hua Shan Road, 200030 Shanghai, China

Carsten Ullrich
Shanghai Jiao Tong University
Haoran Building, 6/F, 1954 Hua Shan Road, 200030 Shanghai, China

Ruimin Shen
Shanghai Jiao Tong University
Haoran Building, 6/F, 1954 Hua Shan Road, 200030 Shanghai, China

ABSTRACT

A recent trend in distant education is the concept of Personal Learning Environments, a combination of services and tools controlled and assembled by the individual user. While offering advantages over traditional Leaning Management System, PLEs bring their own challenges, as they make it difficult to assess the users' interaction with the system. In this paper we present a graph-based visualization of the students' activities, and show how it can be used by both teachers and learners to achieve a better understanding of how the learning takes place.

KEYWORDS

Visualization, Personal Learning Environment, Activities, Feedback, Awareness

1. INTRODUCTION

A recent trend in distant education is the concept of Personal Learning Environments (PLEs). A PLE is a combination of services and tools that are controlled and assembled by individual users (Wilson, 2006). Rather than integrating different services into a centralized system, the idea is to provide the learners and teachers with a variety of different services and hand over control to them to select and use the services the way they deem fit (Chatti, 2007). PLEs can be realized in a number of ways, for instance as mash-ups, similar to iGoogle and Netvibes. A PLE stands in contrast to traditional Learning Management Systems such as Blackboard, because the control does lie in the hand of its user (a student or a teacher), not in the hand of an organization. The user decides which services to integrate.

The flexibility of a PLE can lead to better learning, especially self-directed learning (Law, 2008), but comes with its own problems: the average learner can quickly become overtaxed by the multitude of features potentially available, and requires appropriate guidance and motivation (Ullrich, 2010). Also, teachers face difficulties in assessing the students' learning progress, since the students might use a vast number of different tools. A traditional LMS can provide detailed statistics on the students' interactions since the available tools are know beforehand. In a PLE, the specific configuration will differ between students.

The paper presents a solution to the problem of understanding the students' activities in a flexible environment. Based on an existing activity capturing mechanism (CAM), it describes how to collect activities, and how to process them so that they can be visualized in a way that helps both the teachers and the students to see and interpret their activities. The activities are visualized as a graph whose vertices consist

of tools and whose edges show the order of interaction. The example in Figure 1 centers on a tool training a German exercise ("The present tense of haben"). The graph shows that this tool is connected to, e.g., ("nicht oder kein"), a different exercise.
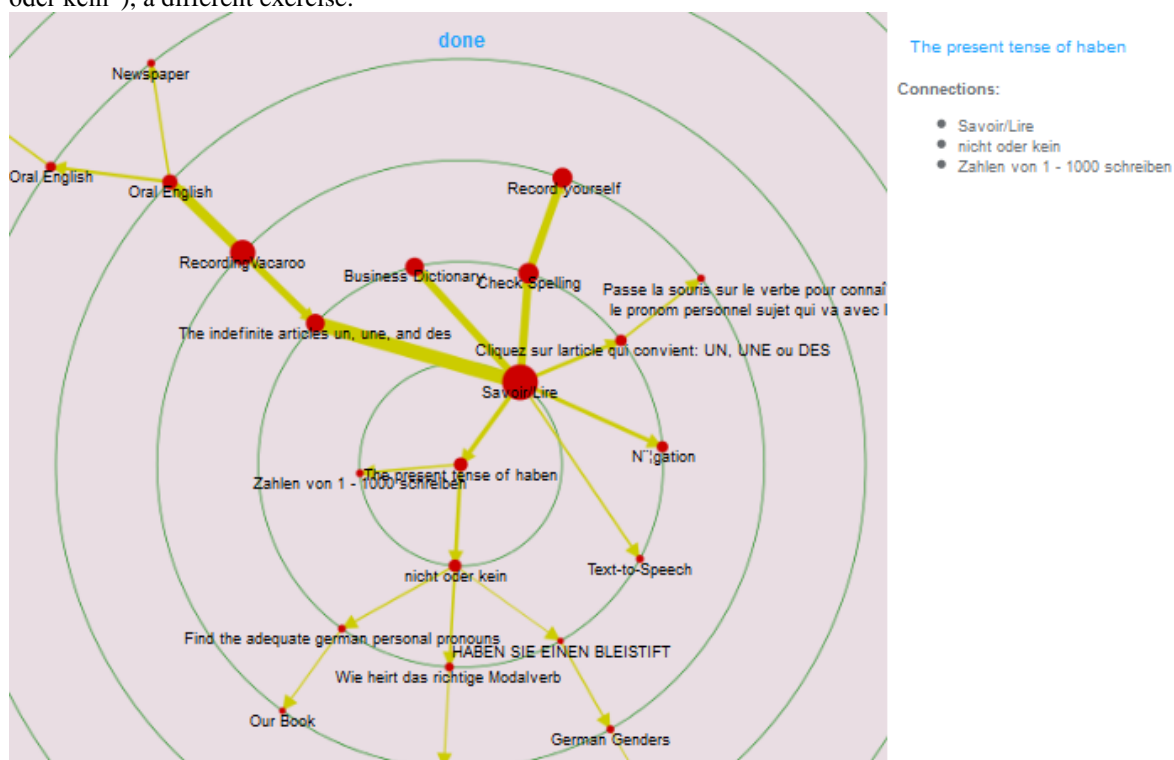

Figure 1. The user interface of the visualization

The paper is structured as follows. It first describes related work (Section 2) and presents the context of this work, both from a learning and technological perspective (Section 3). Section 4 describes the design of and implementation our visualization framework. Section 5 describes the meaning of the visualization. The paper closes with a conclusion and an outlook on further work.

2. RELATED WORK

Educational data mining and visualizations are two frequently used methods to analyze user activities and to provide feedback in a PLE.

Educational Data Mining (EDM) is "an emerging discipline, concerned with developing methods for exploring the unique types of data that come from educational settings, and using those methods to better understand students and the settings which they learn in" (Baker, 2010). A key area of EDM is mining log files of student performance. It is based on four different methods, namely accounting, clustering, classification and data analysis (Baker, 2010). Accounting counts numeric information to provide feedback to students, for example, by measuring the students' time spend while using the learning environment to enable comparison with peers. Clustering groups students and tools according to similar properties. This is used, for instance, to divide students into different groups by the examination scores to give them personalized support. Classification maps tools and students to predefined tags, and groups tools and students belonging to a specific tag. This can be used, for instance, to indicate that some of the students have the same learning habit. Data mining algorithms such as social network analysis and outlier analysis are employed to find the central tools which were used most frequently by the students.

Visualization of educational data transforms data collected during a learning experience, e.g., the usage of a LMS or a PLE, into a manipulable graphical representation. Two main methods of visualization can be distinguished (Soller, 2005). The first method provides a basic view of the data to improve awareness about

the performed user activities. An example is bar charts that show the statistics of the user activities. These bar charts allow teachers and students to identify quickly the participation frequency and the accuracy rate when for problems solving activities (Govaerts, 2010). The second method aggregates data in a high-level designed algorithm to indicate some conclusion. Usually it joins some data together to indicate high level results to be returned as the feedback. For instance, analyzing participation rates with word counts and message reply delays, and relating this back to an ideal model (Govaerts, 2010).

Examples of earlier work in data visualization for education are CourseVis (Mazza, 2004) and Moodog (Zhang, 2007). However, these tools are designed for a specific learning management system and cannot be used in a PLE setting that is configurable by the users. Also, they need to be installed and configured, and cannot be added "on-the-fly".

Recent work by Govaerts et al (2010) addresses data collection and visualization in a PLE setting. Their tool generates charts based on the global course statistics on time expenditure and document usage. The horizontal axis shows the dates and the vertical axis displays the total time spent during the course. In our work, we decided to focus on a new way of the graph visualization instead of the two dimension X - Y chart because we wanted to make the visualization more intuitive, understandable and interactive.

3. SETTINGS

In this section we describe the educational context in which our work was developed and provide information on the technology our work is based on.

3.1 Educational Context

This work has been done in the context of adult distant education. Such a setting is characterized by adult learners with limited time available for study, as they typically have a full-time job and a family. In addition, the digital literacy and motivation to work with new digital technology is often low, meaning that solutions have to be easy to use and understand, and the added value needs to be immediately visible (Ullrich, 2010).

Our testbed consists of teacher and students of the Shanghai Jiao Tong School of Continuing Education (SOCE), a higher education institution based on blended learning. There, students can attend lectures in person in the classroom, but also watch a live video stream. After the lecture has taken place, the video is available for replay. SOCE uses a self-developed LMS, which offers forums and downloadable media, such as slides. Our PLE offers additional learning activities, not covered in the school LMS, and is currently used during class but also for homework.
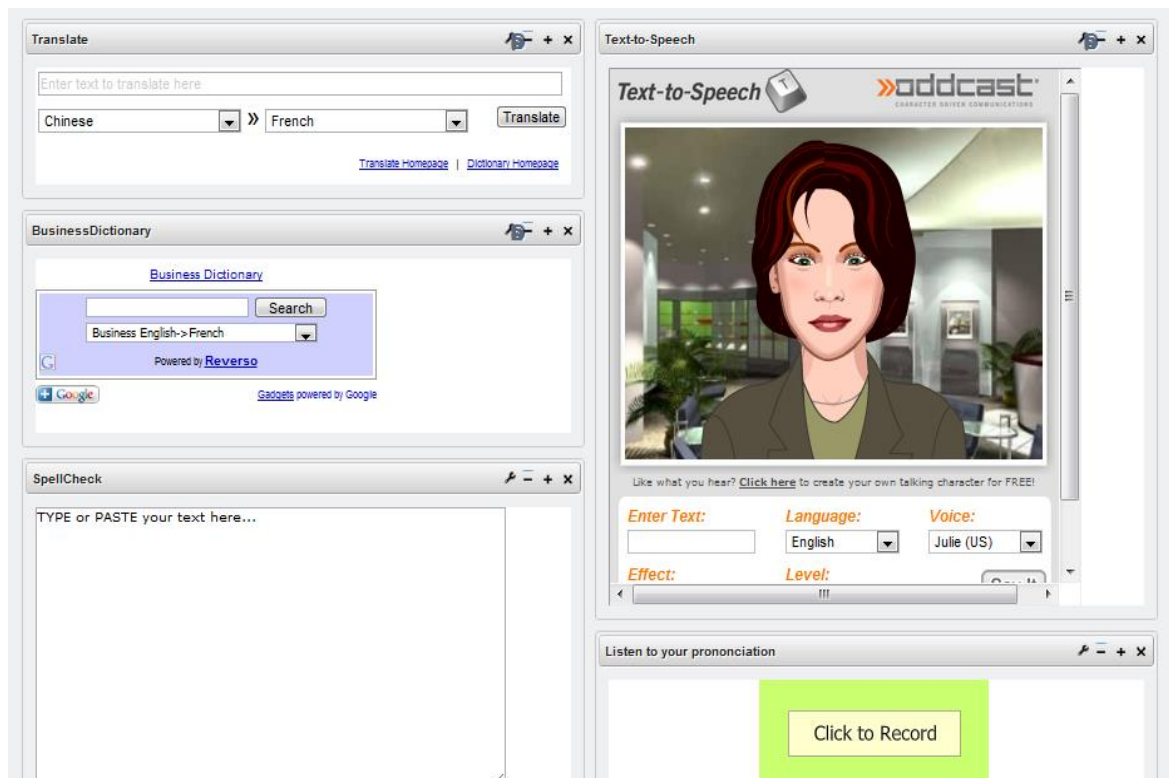
Figure 2. Example of the PLE

Figure 2 contains a screen shot of the PLE used in a French Course. In this example, the PLE consists of the tools Translate, Text-Speech, Business Dictionary, Spell Check and a pronunciation recorder. Students use Translate and Business Dictionary to look up the words. They use the Spell Check to check their French words spelling. They use the pronunciation recorder to record their oral French. Regarding the text-speech tool, it helped the students to listen to unfamiliar French pronunciation.

3.2 Technical Environment

Our PLE is based on Liferay (http://www.liferay.com/), an open-source portal environment that supports a large number of standards that enable extending the functionality of the portal through so-called gadget (dynamic web content that can be embedded on a web page).

Among others, Liferay supports the OpenSocial standard. OpenSocial is a set of common application programming interfaces (APIs) for web-based social network applications, developed by Google along with MySpace and a number of other social networks. It is based on HTML and JavaScript, as well as the Google Gadgets framework, which specifies how to encode gadgets using XML and javascript. A closely related technology is the W3C widget recommendation (http://www.w3.org/TR/#tr_Widgets). In our case, we use OpenSocial and Google Gadgets to implement the tools to be used by the student.

In addition to provide tools directly relevant to learning, it is important to capture the users' interaction with the tools. This can be achieved using CAMera (Schmitz, 2009), a tool for personal monitoring and reporting. CAMera uses Contextualized Attention Metadata (CAM) (Wolpers, 2007) that captures the interactions of uses with resources and tools and provides simple metrics. CAM can receive predefined events, store the data into database, and it also provide web services to query the CAM database by using SQL select statements. In our PLE, the captured user activities were sent to CAM, which stores all the activities. In the visualization we use the CAM data to visualize for teachers and students to help them quickly understand their status. The structure of CAM data is shown in Figure 3. In CAM, an "event" records the user activity, at a specific time. A "relatedenity" records data about the tool that the user accessed, the person who used the tool and the tools urls. All different types of records are discriminated by the role attribute in "eventrelatedentity" table. For example, the role attribute value is "used_widget" for the accessed tool information, and it will be "viewer"

for the personId who used the tool. The "session" table records the user login ip and their accessed domains. The "eventsession" table joins the "event" and "session" by eventid to capture each event's login ip and domain. The table "eventrelatedentity" joins "event" and "relatedentity" and thus captures which entities belong to a specific event. In our work, the "event", "relatedenity" and "eventrelatedentity" is the most important tables we used to record the user activities.
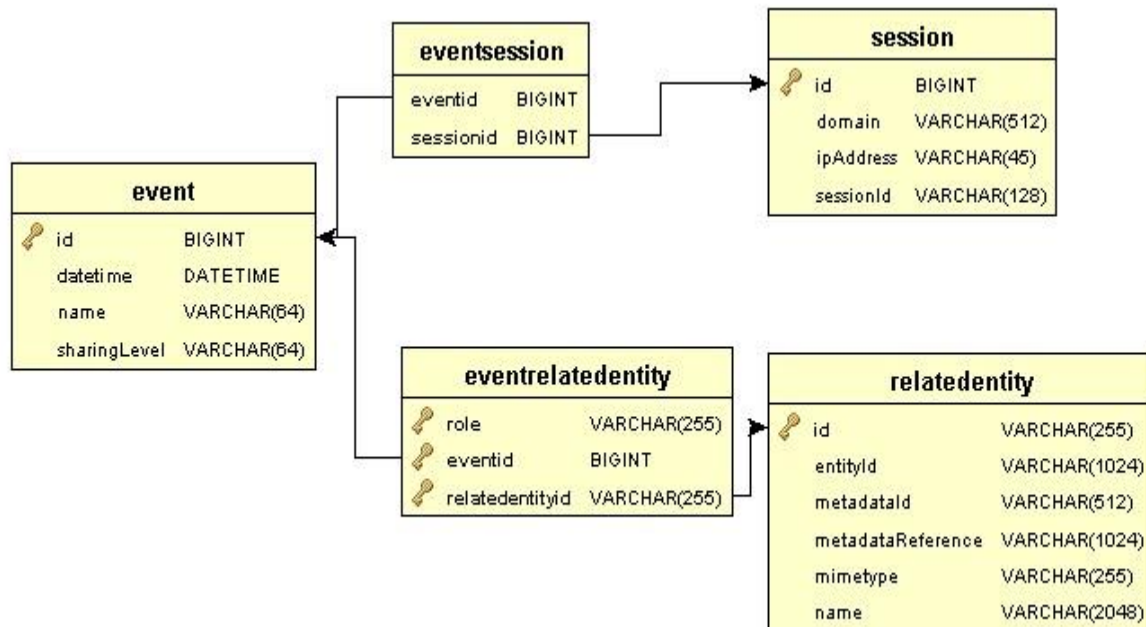

Figure 3. Tables in the CAM database

In our PLE, CAMera is integrated by a Google Gadget, which sends the user activities events to a CAM server, and displays the recent activities.

4. DESIGN AND IMPLEMENTATION

In this section, we describe the different steps required for generating the visualization, starting from data acquisition, over data transformation to the final visualization. We also present its interactive features.

4.1 Data acquisition

In general, data acquisition concerns how to collect the data that eventually visualized. In our case, interaction with tools is captured via events that is send out be the tools. This data is collected by the CAM widget.

In the current implementation, a tool sends out an event (called "usedGadget") every time it has been used for more than 5 seconds. While this is a rather imprecise way of measuring interaction, it has the advantage that existing tools can be extended very easily to send out this information (about two lines of code).

At the visualization generation time, our gadget connects to the CAM web service and retrieves data from the CAM database via SQL statements. Specifically, the gadget requires the following information: the time an event took place, the user's identifier, the tool's identifier and the tool's name. We retrieve the tool usage information of all the users, then group the results by the user identifier and sorted by time. This enables to identify the sequence in which a user used the available tools. The tool's name is later used in the graph for labeling the vertices.

The CAM web service returns a JSON representation of the results. This format now needs to be transformed into a graph.

4.2 Data processing / transformation

The goal of data processing/transformation is transforming the retrieved data into a format suited for visualization. Here, the goal is to capture the tools and exercises used by each user. CAM, however, captures

a significant amount of information not directly relevant to this goal, for instance very detailed interaction data (this depends on the specific gadget). In order to capture only the relevant activities for each student, the gadget analyzes the JSON data and keeps only those events which are of type "used". The result of this processing step is a list of activities of each student ordered by time that represents all the activities performed by the students during the course time.

In the second data processing step, we transform this list of activities of a student S_i into a directed graph $G_i = (S_i, V, E, W, R)$. In this step, we generate graphs for each student s_i and one graph for the activities of all the students. These graphs are defined as follows. Let t_1, t_2,..., t_n be the tools and exercises that were assembled in the PLE. For each tool or exercise t_i which is used by student s_i, a vertex v_i is used to represent it, so that $V = \{v_1, v_2,..., v_m\}$ is a set that contains all the tools and exercises being used by a student in the PLE. The radius of a vertex v_i depends on the number of times the tool t_i that n_i represents has been used. Let $n_i \in N$ be the number of usages of the tool t_i represented as v_i. Then the radius is defined as a function $R : V \rightarrow N$ as follows: $R(v_i) = n_i$. E is a set of ordered pairs of vertices $e = \{v_i, v_j\}$, for which the condition holds if t_i and t_j are the tools represented by v_i, v_j, then there was at least one occurrence of usage by the student where t_j was used by a student immediately after he/she uses t_i. Let $m_i \in N$ be the number of usages of the pair of tool t_i and t_j represented as $e_{i\,j}$. The weight of edge is defined as a function $W : E \rightarrow N$ as follows: $R(e_{i\,j}) = m_i$.

4.3 Visualization
The visualization presents the transformed data as a manipulable graphical representation. In our visualization, we use a Radial Graph (Ka-Ping, 2001) to visualize the PLE learning activities. A Radial Graph is a graph layout using concentric circles. The vertice is placed in the center and their descendants are placed on the next circle, etc.

The visualization also presents aggregated data to provide the high level feedback. The thickness of graph vertices, edges and the edge directions all have meanings. The radius of the vertex corresponds to the number of usages, bigger vertices are used more frequently than smaller vertices. The edges thickness corresponds to the number of the number of usages. Thicker edges are used more frequently than thinner edges.

The reason we used the radial graph visualization is that the radial graph allows us to focus on a specific vertex to discover the relation with the other vertices.

For the implementation, we use the JavaScript InfoVis Toolkit (http://thejit.org/).

4.4 Observation and interaction
The observation of and interaction with the graph enables the users to identify patterns that otherwise would be difficult to detect in the data. An example of visualization was given in Figure 1.

In the visualization in Figure 1, the chart shows the activities of all the students. For instance, the students used the tool ("SavorLire") first, then ("SavorLire"), followed by Text-to-Speech in the next step.

The visualization provides additional information about the center vertex of the chart. All the vertices which were accessed after it (based on the recording data) are listed on the right side of the page.

The visualization is interactive. When a user clicks on another vertex but the central vertex, this vertex will automatically change its position to the center, and also display all the adjacent tools which might be used. For instance, if the user clicks on the tool ("SavorLire"), then the graph shown in Figure 1 changes to the graph shown in Figure 4.
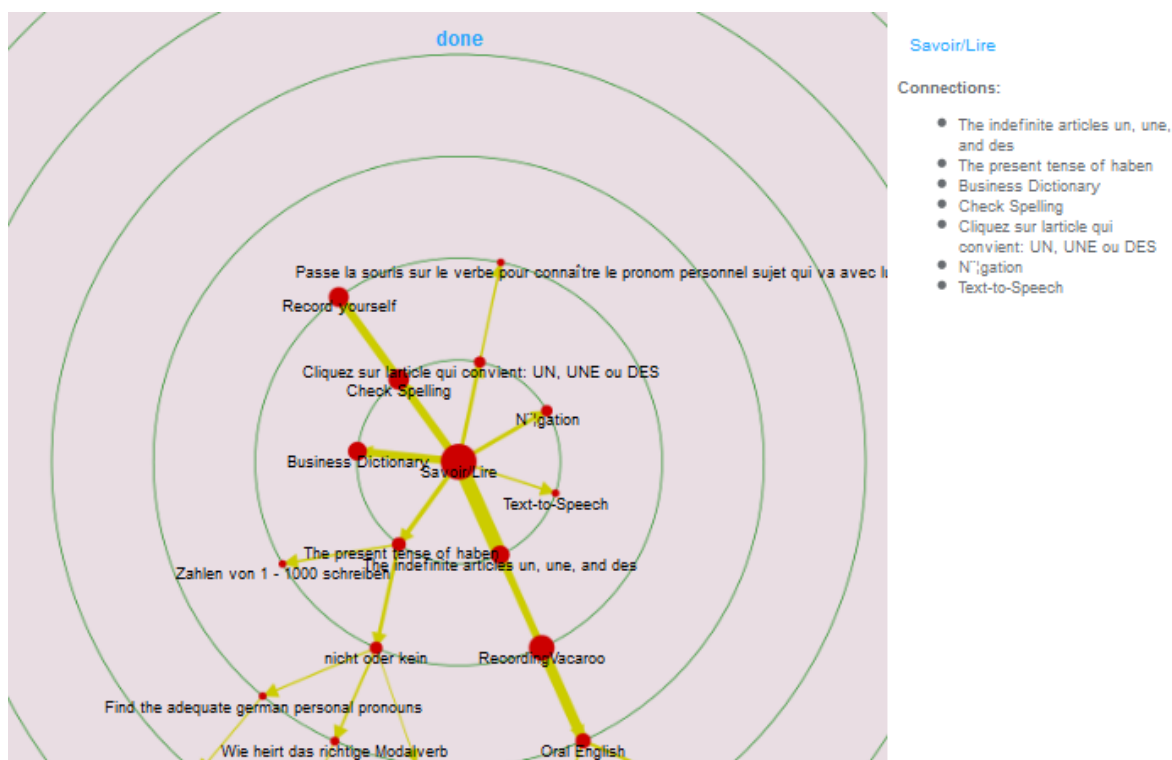
Figure 4. The animated graph

5. WHAT TO LEARN FROM THE GRAPH

The visualization enables the students and teachers to become aware of several features previously not easily discoverable from the raw interaction data.

First of all, the graph contains information about the tools. The users easily identify the tools most useful for the students during their learning process. For example, Figure 4 clearly shows that the tool ("SavorLire") is the most frequently used. ("SavoirLire") is actually an exercise, and thus the visualization shows that this exercise is central to the students' learning. Similarly, users can identify seldom used tools. For a teacher, this provides information that this tool or exercise should be revised or even replaced with a different tool.

Secondly, the graph contains information about the users. If a student seldom uses tools, the vertices size in the visualization will be quite small. The teacher can give feedback to the student and explain about the usefulness of the tool. Students can also compare their activities to other student, and see how their peers use the available tools. For example, a student can see that the text-to-speech tool is used by most of his classmate before recording their own voice. Using this information, he/she can then adjust his method of learning.

Thirdly, the directed edges presents each user's learning pattern. For example, in Figure 4 when the user used tool ("SavoirLire"), he/she would probably use the tool ("the indefinite articles un, une and des") in the next step. Thus, these learning patterns enable students' comparisons with peers.

6. CONCLUSION AND FUTURE WORK

In this paper, we presented a new visualization of the user activities in the PLE for teachers and students to get feedback of using PLE. The visualization presents the following advantages over existing work. First, the required event sending can be added easily to existing gadgets. The visualization itself works within existing PLE frameworks, e.g. ROLE (http://www.role-project.eu/). Extending it to other event systems is not difficult. Once deployed, it helps users monitoring their learning process. The graph-based visualization indicator enables teachers and students to focus on a specific tool used and see the usage of the related tools. The usage information is shown in the graph using an intuitive display based vertex size and edge thickness.

As future work, we will investigate the usage of graph analysis work on the weighted graph of user activities. For instance, we will analyze the learning pattern of each student to find a best learning pattern for everyone

to learn. We will also analyze the graph to find the isolated vertex which indicates the tool is not useful to students, so that it can be edited or replace. We expect such analysis to provide more meaningful feedback to teachers and students. We will also investigate how to use more elaborate information about the actual activities done within the tools.

ACKNOWLEDGEMENT

REFERENCES

Soller, A. et al, 2005. From Mirroring to Guiding: A Review of State of the Art Technology for Supporting Collaborative Learning. *In Artificial intelligence in Education* Vol. 15, No. 4, pp. 206-290.

Ullrich, C. et al, 2010. Not Yet Ready for Everyone: An Experience Report about a Personal Learning Environment for Language Learning. *In Advances in Web-Based Learning ICWL 2010.* Berlin, Heidelberg. Vol. 6483, pp. 269-278.

Bell, G. et al, 2009. *Total Recall: How the E-Memory Revolution Will Change Everything*. Dutton Adult Publishers, New York, USA.

Schmitz, H. et al, 2009, CAMera for PLE., in Ulrike Cress; Vania Dimitrova & Marcus Specht, ed. *In Proceedings of ICWL 2009: Learning in the Synergy of Multiple Disciplines*, Springer, pp. 507-520.

Zhang, H. et al, 2007. Moodog: Tracking Students' Online Learning Activities. *In Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications*, Vancouver, pp. 4415-4422

Yee, K. et al, 2001. Animated Exploration of Dynamic Graphs with Radial Layout. *In Proceedings of the IEEE Symposium on Information Visualization 2001* (INFOVIS'01), p.43

Law, L. et al, 2008. Fostering self-directed learning with social software: Social network analysis and content analysis. *In Proceedings of EC-TEL '08*, pp. 203-215.

Chatti, M. et al, 2007. The future of e-learning: a shift to knowledge networking and social software. *In IJKL* 3(4/5), pp. 404-420.

Chen, M. 2003. Visualizing the pulse of a classroom. *In Proceedings of the eleventh ACM international conference on Multimedia*, New York, NY, USA: ACM Press, November 2003, pp. 555-561.

Wolpers, M. et al, 2007. Tracking Actual Usage: the Attention Metada Approach. *In Educational Technology & Society* Vol. 10, No. 1, pp. 106-121.

Baker, R. et al, 2010. Data Mining for Education. In McGaw, B., Peterson, P., Baker, E. *In International Encyclopedia of Education (3rd edition)*, Vol. 7, pp. 112-118.

Baker, R. and Yacef, K. 2010. The State of Educational Data Mining in 2009: A Review and Future Visions. *In Journal of Educational Data Mining*, Volume 1, No. 11, pp. 3-17.

Mazza, R. and Dimitrova, V. 2004. Visualising student tracking data to support instructors in web-based distance education. *In Proceedings of the 13th international World Wide Web conference*, New York, USA , pp. 154-161.

Govaerts, S. et al, 2010. Visualization Activities for Self-reflection and Awareness. *In Advances in Web-Based Learning ICWL 2010*, pp. 91-100.

Wilson, S. et al, 2006. Personal learning environments: Challenging the dominant design of educational systems. *In CEUR Workshop Proceedings*, Vol. 213, pp. 173-182.