EVALUATING A MIXED-REALITY 3D VIRTUAL CAMPUS WITH BIG DATA AND LEARNING ANALYTICS: A TRANSVERSAL STUDY

Livia Ştefan¹
Florica Moldoveanu¹
Dragoș Gheorghiu²

¹ University POLITEHNICA of Bucharest, Romania
² Doctoral School, National University of Arts Bucharest, Romania
livia.stefan@itc.ro; florica.moldoveanu@cs.pub.ro; dragos_gheorghiu@yahoo.com

Keywords: 3D virtual campus; tracking system; Learning Analytics; pedagogical effectiveness; OpenSimulator.

There is a growing body of literature that recognizes the importance of the 3D multi-user virtual environments (3DMUVEs) with high relevance for prospective developments and improvements in both distant and in-class education. However, to be accepted on a larger scale by education practitioners more analytical studies are required to evaluate these environments against usability and pedagogical effectiveness criteria. The present paper proposes an automatic in-world monitoring system consisting in scripted tracking objects and Learning Analytics (LA) tools for an experimental 3D virtual campus.

The research was conducted within a transversal study using focus groups from two universities, an engineering and a vocational one. The results highlighted the validity of the automated system and LA tools, which can constitute the ground for more efficient teaching and learning in complex activity-based 3D educational settings.
1 Introduction

For many years the 3D multi-user virtual environments (3DMUVEs) are an important topic for research development concerning teaching and learning, especially in higher education. Several online 3D presences in Second Life (SecondLife, 2016) or OpenSimulator (OpenSimulator, 2016) platforms exemplify their research and educational potential (VUED, 2016). Improvements of learning, reduced costs compared with the traditional learning settings, and also mobility and ubiquitous learning are considered as the main advantages of virtual worlds.

From user’s perspective, 3DMUVEs are rich multimedia mediums (2.5D or 3D graphical components, audio, video, voice and text communications), i.e. providing valuable pedagogical affordances.

The 3DMUVEs are mostly explored for their immersivity and sense of presence which facilitates the grasp of abstract concepts or difficult phenomena, and also for their capability to support “communities of practice” (Wenger, 1998), fostering collaborative and social interaction (Girvan & Savage, 2010). Recently, the immersivity is extended with information from the real life, to create “mixed realities” (MR) or a continuum of real and virtual experience (Milgram et al., 1994). MR create pedagogical enhancements, e.g. real time transmissions of classes, fostering collaboration between real and in-world participants or blended–learning strategies for mixing face-to-face learning with different forms of digital and distant learning. As such, the 3DMUVEs support modern learning styles, e.g. contextual, exploratory and social learning.

The instructional design for these educational environments has the same importance as in the traditional ones, i.e. the activities have to be pedagogically significant (Mamo, 2011; Kickmeier-Rust, 2014) and validated against usability and effectiveness criteria.

If not properly designed, the complexity of a 3DMUVE can turn into distracting components and determine an inefficient usage of the environment or even a rejection. The evaluation of the 3DMUVEs against recognized pedagogical, usability and effectiveness criteria is crucial for their acceptance on a larger scale as educational systems.

The most frequent methods for measuring pedagogical effectiveness are based on surveys, but these are not well correlated with the learning process or knowledge transfer (Brown, 2005).

This research attempts to evaluate learning in an experimental virtual campus named “3DCampSim”, designed for collaborative and mixed reality (MR) activities by using several 3D Learning Objects (LO), i.e. 3D objects for teaching and learning. In our research, by designing dedicated functional components, i.e. the Learning Objects, the evaluation was correlated with the
usage of LO during well-established learning activities (courses, laboratories).

The virtual campus “3DCampSim” (the simulator) is an experimental implementation for research purposes, derived from a larger research project, proposed by the research team “Graphics and Virtual Reality” of the Faculty of Automatic Control and Computers, University POLITEHNICA of Bucharest (UPB) to become a Massive Multiplayers Online (MMO) service-based educational counter-part (Moldoveanu et al., 2014; 3DUPB Project, 2015). The virtual campus is implemented on OpenSimulator (OpenSimulator, 2016), an open-source application server for 3D virtual worlds, and hosted on a public server. “3DCampSim” is implemented in a standalone architecture, which currently supports circa concurrent 25 participants.

The present paper proposes an experimental tracking and monitoring system for multi-purpose evaluation of the “3DCampSim” using LA tools. The tools are applied on a high volume of data (“big data”) collected from two focus groups represented by voluntary students and teachers from two universities, under similar experimental conditions. The evaluation of the educational simulator is performed from two perspectives, i.e. the usage of the simulator and the pedagogical effectiveness. The data collection is automated with the help of dedicated tracking objects inside the simulator (“in-world”), which communicate with an external database. To verify the validity of the evaluation with the LA tools, the preliminary results are compared with ones provided by user surveys, evaluating students’ and teachers’ perceived usefulness of the LA tools.

Understanding how the simulator is used for teaching and learning can be leveraged for future improvements of the educational simulator for conducting teaching, learning and assessment activities. The information highlighted by the LA tools can further provide for an adaptive behavior and content of the simulator.

The paper has been organized in the following way: first it provides with an overview of the research topic and related work; then continues with the implementation of the tracking system and LA tools, case studies for evaluation of the “3DCampSim”, and the user surveys. The last sections present the findings of the research, focusing on the key research results, conclusions and future work.

2 The role of LA in complex educational systems

Learning Analytics (LA) has developed as a recent discipline supporting higher education and academic settings (Fulantelli & Taibi, 2014). LA borrows methods from business analytics for learning data collection, measurement and analysis with the purpose to understand and optimize the “learning and the environments in which it occurs” (Ibidem). LA is also connected and sometimes overlapped with methods and objectives of Educational Data Mining
(EDM) which utilizes data mining and machine learning for extracting patterns and unique features and support learning decision support systems. Other LA objectives are fostering reflection on the learning and teaching process (Yousef et al., 2015; Chatti et al., 2012).

Educause Learning Initiative (Educause, 2016) first proposed the study of this domain and defined the LA predictors and indicators and defined LA as “the use of data and models to predict student progress and performance, and the ability to act on that information”. Since 2011, the Horizon Report (Johnson et al., 2014), an annual publication of Educause Learning Initiative, stresses the importance of LA for the quality of the educational systems. George Siemens (Siemens, 2008), the innovator of cMOOC, asserts that the educational institutions need to have both technological means to analyze data and also the capacity to act accordingly. Current LA tools are focused on implementation of adaptation of content based on monitoring, analysis and prediction, and less for assessment, feedback and reflection, according to (Chatti et al., op. cit.).

An online virtual campus is a multi-user, multi-role 3DMUVE, e.g. teachers, students, administrative personnel, visitors and multi-level activities, resembling or complementing the ones from the real organization. Consequently, a virtual campus represents a 3DMUVE with enhanced complexity, requiring an evaluation from different perspectives.

Our research hypothesis is that LA along with other methods, such as formal methods, observations and survey-based evaluation, is a powerful tool for multi-purpose evaluation of teaching and learning in complex 3DMUVE. The multi-dimensional LA reports can contribute to better teaching and learning, thus streamlining the integration of 3DMUVEs in the current educational practice.

The proposed tracking system and the LA tools provide a first level evaluation regarding teaching, learning and assessment activities in the experimental educational simulator, in line with objectives identified by (Chatti et al., op. cit.), i.e. monitoring, analysis, assessment. A second level evaluation’s objectives are in line with other qualitative objectives (Yousef et al., op. cit.; Chatti et al., op. cit.) such as reflection, mentoring, and adaptation, and can be performed after a longer period of utilization.

The beneficiaries of the LA tools, i.e. the “stakeholders” as referred to in Chatti et al. (op. cit.) are teachers, students and researchers.

The analytics of the experimental educational simulator were applied on an “intensive data set” (Steiner, 2014), i.e. on a relative small number of participants with several tracked parameters: the distribution of the avatars in space and time inside the simulator, the activities they undertook (e.g. course attendance, access to resources, time spent), the communication preferences, the learning outcomes (e.g. results from quizzes and term projects). Our rese-
The research method differs from those applied by other researchers, which analyze the OpenSimulator server’s logs or the chat logs, as unstructured data. The input data is extracted from the “3DCampSim” with the help of the proposed tracking system, to collect more structured data than the existing OpenSimulator logs.

LA is a cyclical process consisting of data collection and pre-processing, analytics and action, and post-processing (Chatti et al., 2012). The pre-processing is an important stage, consisting in data integration, filtering and preparation for analysis purposes. Regarding the action stage that follows the analytical one, it can consist in either teacher and student reconsider their activities or in an automatic feedback or recommendation, by further development of the system towards an adaptive real-time learning environment. Post processing is defined as a “continually improving analytics, by refining analytics methods or using new methods, including new data sources” (Steiner, 2014). This is also possible with “3DCampSim”, as the databases and the processing services are external to the educational simulator, and can be integrated with data from Moodle LMS or other learning systems. When larger data sets will be available in the future, advanced statistics or data mining can be applied to implement predictive, machine-learning or decision making LA functions. Dedicated languages, such as R or Python, or Microsoft Azure Machine Learning algorithms can be employed to support these functions.

From the analytics methods usually applied in LA, descriptive statistics and data visualization are employed in the present research. The visualization tools are of crucial importance as they need the capability to communicate information according to a given purpose and audience. In (PTVM, 2016) a synthetic presentation of different visualization methods is given. For the purposes of our research 2D quantitative visualizations consolidated in dashboard presentations and 3D in-world reporting were implemented.

3 Related work

Traditional virtual learning environments, i.e. Learning Management Systems (LMS) or more recently, the Massive Open Online Course (MOOC) platforms (Yousef et al., 2015), support the transfer of knowledge and have collaborative and social learning capabilities, by means of forums, blogs, quizzes. These 2D virtual learning environments generate high volume of data and already offer simple reporting capabilities per user or course usage. Multi-dimensional analytical capabilities are possible with third-party implementations, such as Moodle Engagement plugin (Moodle, 2016) or by data export to third party processing software (e.g. Excel).

Using LAK Dataset, Fulantelli and Taibi (2014) identified more than 450 research papers on LA and EDM from 2011 to 2013. The main data sources are
the traditional Learning Management Systems. Recently the development of Massive Open Online Courses (MOOCs) (Yousef et al., 2015), created a new opportunity and also a challenge for implementing LA due to the high volumes of learning and social data collected from massive and distributed participants.

Up to now, far too little attention has been paid to the use of LA for virtual worlds. Research work of Camilleri et al. (2013) and Fernández-Gallego et al. (2013) have similar purposes to understand learning pathways of learners in virtual worlds.

In Camilleri et al. (op. cit.) the researchers conducted two study cases, one in academic environment for pre-service teachers and one for corporate training. The first one investigates the in-world engagement and predicts behavioral intentions of students in the 3DMUVEs by highlighting the types of collaborative tasks the students are undertaking in-world. The LA takes the form of statistical analysis and real-time dashboards for visualization of different categories of activities on the commercial Avaya Live Engage Platform. In Fernández-Gallego et al. (op. cit.) a framework for LA and EDM in an OpenSimulator-based 3DVW is presented, for validation of the learning units against the IMS LD specification for collaborative activities using a compliant engine and a formalization of events and interactions among the avatars using Petri nets.

In Kickmeier-Rust, 2014 smart software tools are presented as a support for teachers using an OpenSimulator-based island for language learning in a playful manner. Instead of dealing with “big data” from log files, LA real-time prediction algorithms are used for formative assessment and feedback purpose, using finite set of competences inferred from students’ performance data.

An interesting case is made by the educational games where the LA is embedded in the game play (Steiner, 2014) and not implemented as a separate system.

In “3DCampSim” different types of educational activities can be organized such as access to video and web resources, peer-assessment, teamwork, discussions and tutoring, but the main activity is represented by in-world live transmissions of the courses and laboratories. As we could identify, the majority of the research work evaluate the collaborative and communication activities, but less explore the mixed reality context (Liu, 2012).

4 Implementation of the LA-based system

4.1 Learning affordances of the 3D simulator

For our research it was designed an experimental virtual campus “3DCampSim”, consisting in two buildings and several distinct educational spaces: a) 2 virtual laboratories which model a Computer Graphics laboratory; b) a virtual
meeting space for private or small group meetings; c) 2 exhibition spaces for showcasing term projects, peer review and teacher evaluation; d) a “sandbox” space for collaborative 3D graphic and design activities; e) a gamified space for inter-university communication; f) administrative areas. Also an indoor and an outdoor lounge/meeting spaces for social interaction were designed. The virtual campus occupies an OpenSimulator mega region, so the movement to other regions is not monitored, but only the movement to the designated educational areas.

Several specialized 3D collaborative and adaptive LOs were designed with different purposes: presentation panels for course materials, media panels for websites and video materials (e.g. showcase term projects, reporting), quiz panel, notecards, interactive whiteboard.

4.2 Automated Data Collection

The tracking objects, i.e. scripted objects able to store data in an external Microsoft SQL (MSSQL) database by means of an implemented add-on module, have the role to monitor the usage of the “3DCampSim” to provide source data for the LA tools. The designated educational areas have identification name, also registered in the database. In order to be monitored and provide a common baseline for a future analysis, the identifications of the courses, activities, quizzes and project assignments are taken from the university’s Moodle LMS and connected with avatar’s Universal Unique Identifier (UUID).

A data integration and processing service connects MSSQL data with OpenSimulator’s MySQL and Moodle MySQL data (Fig. 1), and performs the necessary data pre-processing.

Fig. 1 - Integration between the 3D simulator and other data sources (e.g. Moodle platform)
4.3 The measured metrics

Four LA indicators were designed to target different stakeholders’ areas of interest. For university teachers and administrators, but also for researchers, two indicators were designed to measure the activities of the virtual campus:

1. in-world distribution of avatars in time and space, across the educational 3D spaces (A1);
2. in-world communication preferences (A2);
3. course rank by attendance.

For student personal analytics, an indicator was designed to measure their engagement, by taking into account a weighted combination of students’ average time spent in different classroom, quizzes’ results and participation to the Q&A session (S1).

For teacher personal analytics, an indicator was designed to measure their activity by number of participating students (T1).

5 Research design

Several experiments were conducted during one month, two hours per day, to preliminary evaluate teaching, learning and assessment activities in the campus simulator. Other research methods (short interviews, questionnaires) were also employed to strengthen the validity of the results from automated data collection, and were presented in Stefan (2016).

The “3DCampSim” was evaluated by 8 students and 3 teachers, previously instructed regarding the usage of the simulator, navigation, communication and the purposes of the research. A short interview highlighted that 4 students and one teacher had previous experience of using a 3D environments, such as simulations and educational games and that the majority of the students were not aware of the usefulness of a 3D campus simulator. The students accessed the simulator both from a laboratory computer and also remotely. The demographics of the study was the following: Students aged 20-25 (6 male, 2 female); Teacher aged 30-40 (1 male); Teachers aged 50-60 (1 male, 1 female), divided in two focus groups from two universities, an engineering (POLITEHNICA from Bucharest) and a vocational one (National University of Arts Bucharest). Each group made experiments during 2 weeks.

The in-world activities were designed after the ones familiar to teachers and students, therefore easier to follow, and spread across different days:

1. one-hour lectures or laboratories;
2. small group discussions and projects’ evaluations;
3. social interactions, communications, access to resources (e.g. videos).
An in-world lecture or laboratory consisted in a theoretical presentation of the main topic, a practical demonstration, a Q&A session, collaborative or individual tasks, quizzes and project assignment. The questions from Q&A sessions were taken individually from students and answered on a dedicated in-world communication channel. The educational content was presented under the form of Power Point documents, recorded or live video presentations, web sites.

The learning topics were computer graphics and interior design, respectively. Collaborative tasks were writing in a Google Docs document, 3D whiteboard or 3D building or design in a dedicated “sandbox” space. Students’ avatars could move outside the laboratory to other areas, e.g. the lounge, meeting room, exhibit room, or logout.

6 Results

The indicators are consolidated in dashboard visualizations, which at this stage of the research are implemented using Google cloud services (i.e. Google Sheets) and are in-world accessible from an integrated web browser and media objects. The “Simulator’s activity” dashboard (A1 and A2 indicators) shows the distribution of avatars in different areas of the simulator, types of communication tools and course rank by attendance within every 2 hours. It is publically available from the administrative area of the simulator (Fig. 2). The “Student’s engagement” (S1) and “Teacher’s activity” dashboards (T1) (Fig. 3) are reporting personal indicators, visible only by each student and teacher by means of heads-up displays (HUDs) “worn” by their avatars, and require a Google account.

The teacher’s activity shows:
- classification of students according to their engagement indicator;
- the less attended activity (course/laboratory);
- the most attended activity (course/laboratory). The student’s engagement report shows the weighted engagement score.

Fig. 2 - Simulator’s activity accessed from an in-world media panel
The initial research hypothesis was also evaluated with the survey method, summarized in Table 1. The responses were given using a 5-point Likert scale, from strongly disagree to strongly agree.

<table>
<thead>
<tr>
<th>Questions</th>
<th>LA objective</th>
<th>No. of Students</th>
<th>No. of Teachers</th>
</tr>
</thead>
<tbody>
<tr>
<td>The provided LA tools are helpful feedback tools</td>
<td>Feedback</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>The provided LA tools helped me monitor and analyze activities</td>
<td>Monitoring and self-evaluation</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>The provided LA tools helped me reflect upon and improve my performance/activities</td>
<td>Reflection</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>The provided LA tools will help receive/provide adapted content</td>
<td>Adaptation</td>
<td>4</td>
<td>2</td>
</tr>
</tbody>
</table>

7 Discussion

A 3D educational simulator is a complex environment, with rich content, several possible usage scenarios and numerous users concurrently connected. Therefore, the evaluation is not a trivial task, usually being based on formal research methods and users surveys. These traditional methods offer a synthetic evaluation and require the time consuming organization of case studies and user experiments. A critical problem is finding a sufficient number of voluntary users among teachers and students willing to take part in a complex study.

For the experimental 3D educational simulator technical solutions were developed with the purpose to compensate these difficulties and implement an
automated method for a more objective analysis and support teaching, learning and assessment in complex 3D MUVEs. The proposed method allows the implementation in a 3D simulator, otherwise a close environment, of quantitative and qualitative Learning Analytics tools, which in time can lead to better understanding teachers’ and students’ preferences and performances, will encourage self-monitoring and reflection, and the development of services based on specific cognitive models.

The repeated experiments with the two focus groups generated a high volume of data. These were collected with different purposes, e.g. logins and logouts, avatar positions, time stamps, object interactions, quiz results, communication and environment options, and require the development of specialized post-processing database services. Some of the collected data are similar to those employed in traditional LMS, while other data are specific to a 3D environment, having both time and space dimensions, complex visual interface and real-time multi-concurrent usage.

The dashboards are powerful synthetic presentation tools which may need experts to understand and correlate the information. For our research purposes we considered useful to design dashboards with more explicit quantitative charts. These can be accessed in the administrative area of the simulator for public information and also from the HUDs attached to avatars, for personal inquiries.

The results form the automated tracking system were similar to those from user surveys. The upper graph from Fig.4 indicates that most of the users visited the LABORATORIES, MEETING ROOMS and the EXHIBIT ROOMS, and less frequent the LOUNGE, the administrative area and the outdoor areas.

Fig. 4 - Simulator’s activity dashboard
The preferred communication media in the virtual environment was the text-chat (the middle graph from Fig. 4). The participation to the Q&A session was relatively poor (the lower graph from Fig. 4).

The majority of the responders of the user survey recognized the benefits of the LA tools for analysis, feedback and self-evaluation, but their role for reflection was not very well perceived, mainly by students, and consequently will need further attention. The responses from the questionnaires indicate that after the experimentation of the virtual campus the students became more aware of the benefits of the educational usage of the 3D simulator.

Conclusions and future work

This study proposes a tracking and monitoring solution complemented with LA tools for evaluation of an experimental 3D virtual campus from two perspectives: a) the functional design of the environment; b) the effectiveness of teaching, learning and assessment in a 3D immersive context with collaborative and mixed reality Learning Objects. Instead of analyzing large log files, which are valuable performance analysis tools, the data are collected using an in-world tracking system integrated with several other data sources (e.g. Moodle LMS) for a common baseline. The analysis of data from the automatic evaluation was confronted with data from surveys from two focus groups, students and teachers, which showed similar positive results regarding the usefulness of the LA tools for better supporting teaching and learning in the 3D campus simulator.

The LA reporting tools employed processed historical data under the form of online dashboards, showing the educational usage of the simulator, students’ and teachers’ personal analytics.

The proposed solution opens the way to complex decision and support systems (e.g. recommendation systems) for activity-driven 3DMUVEs, such as a virtual campus. Different stakeholders can benefit from using a virtual campus with analytical tools. In the future more complex analytical tools can be designed, such as more interactive and real-time solutions for the reporting tools, provision of adaptive in-world services.

Even the results are preliminary and had the role to validate the research method they are encouraging for further investigation of the 3DMUVEs and streamline their integration in the current educational activities.

Acknowledgements

This research was supported by the Sectoral Operational Programme Human Resources Development (SOP HRD), financed from the European Social Fund and the Romanian Government under the contract number
POSDRU/159/1.5/S/137390. The experiments were carried out within the “Time Maps: Real Communities. Virtual Worlds - Experimented Pasts” PNCD II IDEI Project.

REFERENCES


Kickmeier-Rust, M. D., Bull, S., Meissl-Egghart, G. (2014), *Collaborative Language Learning In Immersive Virtual Worlds: Competence-Based Formative Feedback and Open Learner Modeling*, International Journal of Serious Games, Volume I,


Steiner, C. M., Kickmeier-Rust, M. D., Albert, D. (2014), *Learning Analytics and Educational Data Mining: An Overview of Recent Techniques*, Learning Analytics for and in Serious Games, 6.


Virtual University of Edinburgh Grid (VUED) (updated 2016/03/07), URL: http://vue.ed.ac.uk/openvue/ (accessed on 5th February 2015).
