Open Service for Video Learning Analytics

Konstantinos Chorianopoulos[1], Michail N. Giannakos[2, 3], Nikos Chrisochoides[2], Scott Reed[4]
[1] Department of Informatics, Ionian University, Corfu, Greece
[2] Department of Computer Science, Old Dominion University, Norfolk, VA, USA
[3] Department of Computer and Information Science, Norwegian University of Science and Technology, Trondheim, Norway
[4] Lakeland High School, Suffolk, VA, USA
choko@ionio.gr, mgiannakos@acm.org; nikos@cs.odu.edu; scottreed@spsk12.net

***Abstract*—Video learning analytics are not open to education stakeholders, such as researchers and teachers, because online video platforms do not share the interactions of the users with their systems. Nevertheless, video learning analytics are necessary to all researchers and teachers that need to understand and improve the effectiveness of the video lecture pedagogy. In this paper, we present an open video learning analytics service, which is freely accessible online. The video learning analytics service (named SocialSkip) facilitates the analysis of video learning behavior by capturing learners' interactions with the video player (e.g., seek/scrub, play, pause). The service empowers any researcher or teacher to create a custom video-based experiment by selecting: 1) a video lecture from YouTube, 2) quiz questions from Google Drive, and 3) custom video player buttons. The open video analytics system has been validated through dozens of user studies, which produced thousands of video interactions. In this study, we present an indicative example, which highlights the usability and usefulness of the system. In addition to interaction frequencies, the system models the captured data as a learner activity time series. Further research should consider user modeling and personalization in order to dynamically respond to the interactivity of students with video lectures.**

Keywords- User Interactions, Learning Analytics, Video Lecture, Open Learning System, Video Analytics.

I. Introduction

Researchers [1] [2] have recently begun to explore student interactions with video-based learning in order to provide educators with valuable information about students (e.g., Khan Academy, Coursera). However, the capture and analysis of this information is still in an embryotic research stage, mainly because there are few open access tools that facilitate the capturing and analysis of video interactions. The service described herein might contribute to this critical research effort.

With the widespread adoption of online video lecture communities, such as Khan Academy, it has become critical to conduct research to understand how students learn via video lectures. A significant body of related research into the impact of video lectures has been made. However, the majority of previous efforts have been mainly focused on: a *sporadic, or one time* use of video lectures in an educational context and/or the investigation of only a *single factor* like student performance.

Video lectures have given rise to flipped (or inverted) classrooms. This specific type of blended-learning classroom utilizes technology, such as video, to move lectures outside the classroom, thereby giving students and teachers time for active learning in the classroom [3]. At the same time, recent technical and infrastructural developments [3] make the potential of video based learning ripe for exploration. Capturing and sharing learners' interactions in emerging learning technologies can clearly provide scholars and educators with valuable information.

II. Related Work

Video lectures have emerged as one of the premier Open Educational Resource. Many instructors in higher education are implementing video lectures in a variety of ways, such as broadcasting lectures in distance education, delivering recordings of in-class lectures with face-to-face meetings for review purposes [4], and delivering lecture recordings before class to conserve class time and flipping the day for hands-on activities [5]. Other uses include showing videos that demonstrate course topics, and providing supplementary video learning materials for self-study [6]. Researchers have delineated the educational advantages and disadvantages of video lectures [4]. However, previous efforts have been mainly focused on the sporadic use of video lectures and the investigation of a specific feature.

Our **motivation** for this work is based on emerging developments. First, the use of videos for learning has become widely employed in recent years. Video-based technological tools have been developed, and many educational institutions and digital libraries have incorporated video into their instructional materials. Second, despite the growing number and variety of video lectures available, there is limited understanding of the nature and quality of their effectiveness, in terms of how students use and learn from video lectures. Specifically, limited research currently exists regarding guidelines for the use of video lectures and the design of pedagogical systems for their use. For example, it is established that learners benefit from highly structured learning material, but the manual editing of video is not feasible for most learning organizations and instructors.

III. Video Learning Analytics Service

In this section, we present the video analytics service. We used the Google App Engine (GAE) cloud platform and the YouTube Player API to develop the system. The system has several advantages in comparison to stand-alone applications. Users do not need to go through an installation process (http://goo.gl/SJng6O); rather, they just have to visit the link. If there is an updated version, they simply have to refresh the page. In addition, system's architecture is modular and allows re-use of the components.

There are several benefits of the selected architecture (GAE, YouTube, Google accounts); and although, learners can (optionally) use their personal Google account in order to sign in and watch the uploaded videos. The system has been designed to work with no need for any personal data from the user. Each time a user visits the web-video player application, a new log is created. Whenever a button is pressed, an abbreviation of the button's name and the time it occurred are stored.

Practically, every user with a Google account can be a researcher. To do so, one has only to sign in to the service (http://goo.gl/SJng6O), connect the selected video from YouTube, configure the video player buttons or slider, and type the address of the online survey he/she wants to use (see Figure 1, up). By taking these few steps, a respective video experiment URL is created (Figure 1, down).

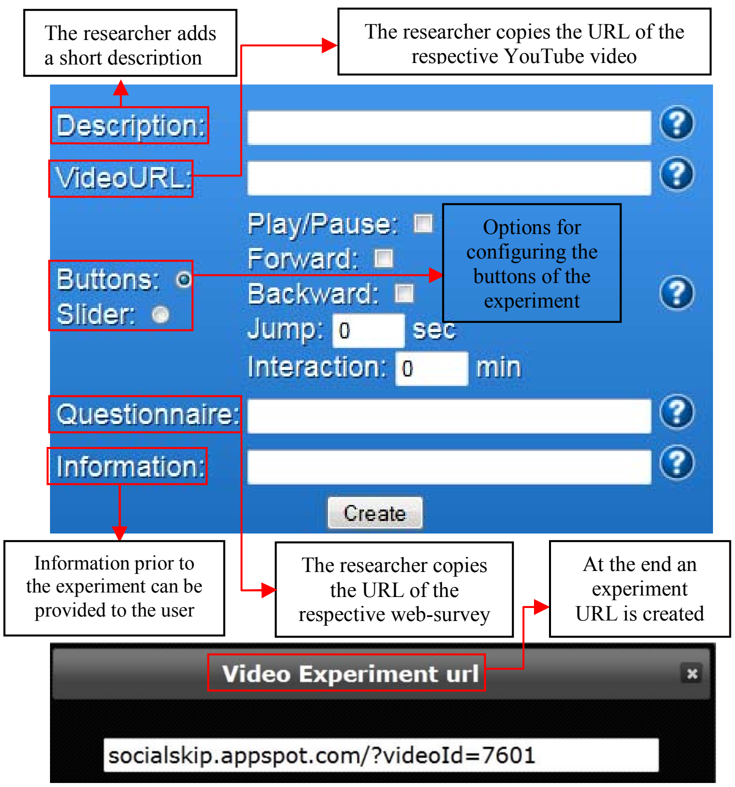

Figure 1. Experiment creation window (up); video experiment url (down)

This URL contains the interface of the video analytics system. The video analytics system (Figure 2) employs the utilities selected by the researcher buttons.

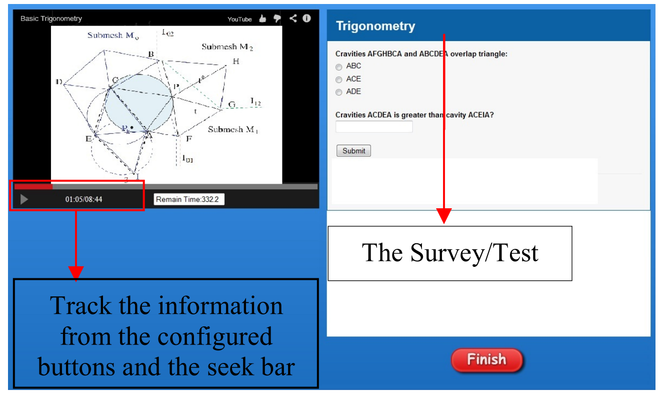

Figure 2. The interface of the system has familiar buttons, seek bar, as well as questionnaire functionality

When the researcher terminates the specific experiment, he can download all the collected data, visualize the activity of each video, configure the experiment, and, thereafter delete it (Figure 3). These options give the researcher the flexibility to test different activities or functionalities on different groups of students, analyze the results, and develop useful conclusions about how students use and learn from video-based learning systems.

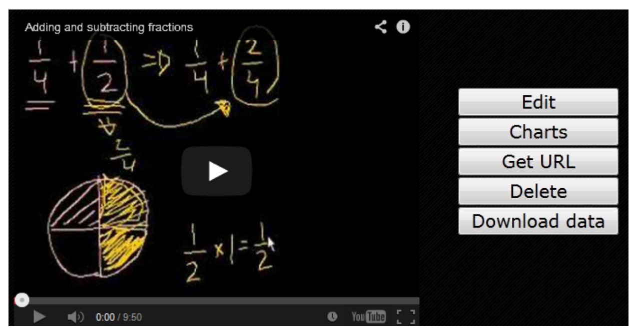

Figure 3. The interface of researcher's configuration/management area

Concerning the visualization capabilities of the system, we have opted to use time series to represent learner activity. A time series is a sequence of data points, measured typically at successive points in time and spaced at uniform time intervals. Time series analysis provides methods for analyzing time series data, in order to extract meaningful statistics and other characteristics of the data [7]. Figure 4 exhibits an example of the visualization of the learner activity via time series technique. We provide two types of time series: The first one (important) depicts only the replay user interactions, which have been found to be the most representative for user activity [8]. The second one (summary) depicts of all user interactions (pause, play, seek), In addition to the preselected visualizations of time series, the user of the system has the option to download the data-set locally in a Comma Separated Values (CSV) format. Then, the researcher might import the CSV file into a visualization program of preference (e.g., R) for more advanced analysis and graphics.

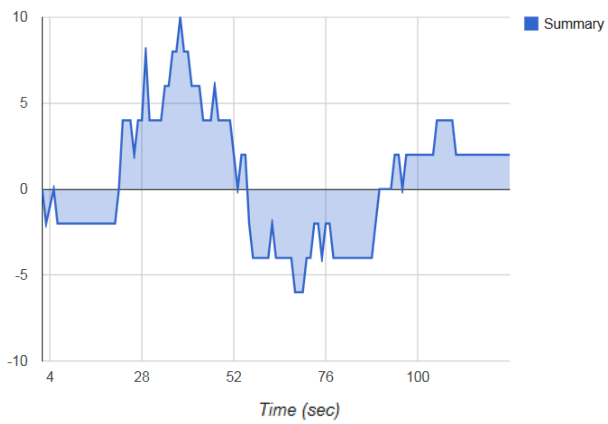

Figure 4. Example Visualization of the Learner Activity

IV. SERVICE VALIDATION

The goal of the system validation was to collect activity data from the learners, as well as to establish a flexible experimental procedure that can be replicated and validated by other researchers. Instead of mining real usage data, we have designed a controlled experiment, because it provides a clean set of data. The experiment took place in a lab with Internet connections, general-purpose computers, and headphones. Twenty-three university students (18-35 years old, 13 females and 10 males) spent approximately ten minutes watching each video (buttons were muted). All students had been attending the Human-Computer Interaction courses at the Department of Informatics at a post- or under-graduate level, and received course credit in the respective courses. Firstly, the students were asked to watch a video lecture, but the video player did not allow any navigation. Next, a time restriction of five minutes was imposed, in order to motivate the users to actively browse through the video and answer the respective questions. Upon determining a basic understanding between the learners' behavior data and the important segments detection, further research could progress to larger scale studies, or even to field studies and the mining of large data-sets.

V. DISCUSSION AND CONCLUSION

As the practice of learners' watching videos on Web-based systems increases, more and more interactions are going to be gathered. Dynamic analysis of this wealth of data will allow us to better understand learner experience. In addition, the combination of richer user profiles and content metadata will provide opportunities for adding value to data obtained from video- based learning.

Although many corporations and academic institutions are making lecture videos and seminars available online, there have been few and scattered research efforts to understand and leverage actual learner experience. In addition, to the best of our knowledge, there are no efforts using diverse data like interactions with the system, students' performance, and attitudes, in order to triangulate them and derive valuable information about how students use and ultimately learn via video systems.

In this paper, we presented a video learning analytics service. The service can be used by anyone who wants to design a custom experiment or a video-based course and to elaborate on student learning. The open-source video learning analytics system is easy to use, may be applied to any viewer, and easily incorporates any video lectures from YouTube and any quiz/survey from Google Drive. In addition, the underlying system is open-source and available for further customization and improvement.

By taking into account learners' interactions and complementary data—such as students' demographic characteristics (gender, ethnicity, English-language skills, prior background knowledge, their success rate in each section, their emotional states, the speed at which they submit their answers, which video lectures seemed to help which students best in which sections, etc.) — new avenues for significant research will open. Captured data can feed powerful algorithms and facilitate personalized feedback. Future work will help to collect diverse data (i.e., success rate, emotional states), which will allow the community to consider the challenges for developing more personalized and effective video learning systems. These advancements will lead us to several earlier interventions (e.g., to prevent drop-out), or to adaptive services and curricula.

ACKNOWLEDGMENT

The authors wish to thank the participants to the study who kindly spent their time and effort. This work is supported by CCF-1139864 NSF grant and the Richard T. Cheng Endowment.

REFERENCES

[1] C. Ullrich, R. Shen, and W. Xie, "Analyzing Student Viewing Patterns in Lecture Videos," Proc. Conf. on Advanced Learning Technologies (ICALT 13), IEEE Press, 2004, pp. 115–117.

[2] J. Kim, P. J. Guo, D. T. Seaton, P. Mitros, K. Z. Gajos, and R. C. Miller, "Understanding in-video dropouts and interaction peaks inonline lecture videos," Proc. of the first ACM conference on Learning @ Scale (L@S '14), ACM Press, 2014, pp. 31-40.

[3] A. Roehl, S.L. Reddy, and G.J. Shannon, "The flipped classroom: An opportunity to engage millennial students through active learning strategies," Journal of Family & Consumer Sciences, vol 105, 2, 2012, pp. 44-49.

[4] J. Brotherton, & G. Abowd, "Lessons learned from eclass: Assessing automated capture and access in the classroom," ACM Transactions on Computer-Human Interaction, vol 11, 2, ACM Press, 2004, pp. 121–155.

[5] J. Day, and J. Foley, "Evaluating web lectures: a case study from HCI," Proc. CHI '06 extended abstracts on Human factors in computing systems (CHI EA '06), ACM Press, 2006, pp. 195-200.

[6] S. Dhonau, and D. McAlpine, "''Streaming'' best practices: Using digital video-teaching segments in the FL/ESL methods course," Foreign Language Annals, vol. 35, 6, 2002, pp. 632–636.

[7] M. Avlonitis and K. Chorianopoulos, "Video Pulses: User-Based Modeling of Interesting Video Segments," Advances in Multimedia, vol. 2014, Article ID 712589, 9 pages, 2014. doi:10.1155/2014/712589

[8] K. Chorianopoulos, K. "Collective intelligence within web video. Human-centric Computing and Information Sciences," 3, 1:10 2013.