History-Aware User Awareness in Web Lectures

Markus Ketterl
Center for Information Management and Virtual Teaching
University of Osnabrück, Heger-Tor-Wall 12
49074 Osnabrück
Email: markus.ketterl@uni-osnabrueck.de
Germany

Robert Mertens
Fraunhofer IAIS, Schloß Birlinghoven, 53754 Sankt Augustin
Email: robert.mertens@iais.fraunhofer.de
Germany

Johannes Emden
Center for Information Management and Virtual Teaching
University of Osnabrück, Heger-Tor-Wall 12
49074 Osnabrück
Email: johannes.emden@uni-osnabrueck.de
Germany

Oliver Vornberger
Dept. of Computer Science, University of Osnabrück, Albrechtstrasse 28,
49069 Osnabrück
Email: oliver@uni-osnabrueck.de
Germany

Abstract

User awareness has become a popular feature in many social web applications. In classic text-based web-systems, user awareness features show how many users are online in a web application or how many users are accessing the same web page. When time based media like web lectures are concerned this approach comes to its limits since time-based media are inherently different from classic text-based media.

The main difference in these two types of media is that text-based media can easily be skimmed at a glance while video- or audio-objects have to be fully replayed when users want to grasp their content.

This paper presents an approach that employs time-based usage statistics for all users in a web lecture system in order to provide users with a means to communicate with other users who are online and have watched those parts of a media object that the user is interested in.

The work is implemented in the context of the Opencast Matterhorn project – an open source based project for producing, managing and distributing academic video content.

1. Introduction

In many portals, user awareness is designed to tell users which of their friends are online. While this information is useful in a leisure time scenario, it is only of limited use in an e-learning scenario or, for that matter, in any scenario in which users work with the content. In a learning related scenario, user awareness can be used to indicate whether other users are working with the same content or possess knowledge about that content. The user can then direct questions about that content directly to other users that are both online and familiar with the content.

In classical text and picture-based media, detecting users that work with the same piece of content and are currently online is relatively easy. Access statistics from a web server can be used to determine which object (like a specific chunk of text) a user is currently accessing. If these objects are reasonably sized, which is the case in most web content, a user accessing it can easily tell the contents of that object. However, web lectures and other time-based media differ from classical text and picture based media in that a user can hardly tell the content of those passages of the media-object he has not already replayed [12]. Hence, the fact that other users are online and retrieving the same media-object does not deliver any cue as to whether these other users have actually visited the parts of the media object that are relevant the user.

The approach presented in this paper solves this problem by matching viewing statistics of the current user with viewing statistics of all users that have accessed the respective media-object. The user interface provides a feature to specify arbitrary parts within a web lecture so that users can ask questions about specific passages in a media object.

The remainder of this paper is structured as follows: Section 2 discusses related work in the field of user awareness in text based media as well as current developments bringing social web approaches to web lectures and other time based media. Section 3 explains in which context the work presented in this paper is being used. Section 4 describes the users' perspective on the system. The technical background of the application is explained in sections 5 and 6. A conclusive discussion of the approach is given in section 7.

2. Related Work

Making Web activity more apparent to users is a major part of today's web experience. In literature one can find different types of user awareness. Liechti [11] differentiates besides others four main categories, which are not mutually exclusive: group, workspace, contextual, and peripheral. Dourish and Bellotti [2] proposed one of the first definitions for awareness: "Awareness is an understanding of the activities of others, which provides a context for your own activity". Group awareness conveys information about the state and activity within a team. Workspace awareness facilitates the coordination among users and allows communication and collaboration. Contextual awareness adapts the idea of presenting information to the user situation. Peripheral awareness denotes the way in which a system presents information about effortless monitoring of activity and also the level of detail. Web 2.0 and social network application inherent many of this ideas and combine them in an attractive way. Early ideas for a social web (and social navigation) have been presented by [14] or [1]. Nowadays social network applications allow users inside a network to get information about what is happening on the site or in the application (e.g. who checked my profile, left a comment). Examples for awareness of web activity and description can be found in [5, 17]. Group awareness is an important issue in eLearning applications. Lecturers as well as students need to know the activities, background knowledge to support learning effectively. Ad-hoc communication and collaboration among user can supports online learning or foster coordination work. Establishing and maintaining awareness has been reported to be difficult without appropriate supporting tools [6]. Fundamental privacy techniques and tradeoffs for awareness systems have been presented in [7]. In [8] an awareness component model with a notification mechanism has been presented that enables group awareness within web-based learning components. Researchers have reported a significant effectiveness of applying media technologies, such as chatting tools, shared workspace tools, video and audio tools, visualization representation tools, email, and notification tools in deliver awareness through groupware system [19]. Examples for the visualization of social data that conveys information about the online world and its participants can be found in Lifeline [15], Conversion Map [16], Netscan [18] or the Personal Map project [3]. Social browsers to support awareness and social interaction can be found in work from [10]. The authors explain the use of a people browser to visualize the members of a community and also the human interaction over time. In web lectures, user awareness has – to the author's knowledge - only been implemented in the approach presented in [4]. In that approach, three user awareness features are implemented: a "who is online"-list that offers information on other user who are online, a social scrubber that shows thumbnail pictures of other users who are currently watching the same web lecture on the web lectures timeline and a cumulative footprint feature that shows how many accesses a passage of the video has received.

The "who is online"-list is not different from classical user awareness features from many web portals; it simple shows the users who are currently logged into the system. The social scrubber is an interesting social feature that is tailored to the use in time-based media. It does, however, face the problem that it only shows who is currently watching the same web lecture. It is thus not history aware and loses valuable information. The social footprints help the user to identify what parts of a web lecture have been watched most by other students but they have no relation to other users who are currently online in the system as the footprints are a purely cumulative feature [13].

3. Context of this work

Web lectures have become a reliable learning companion for students at many places. Applications and research in the field of lecture recording have grown exponentially across the world. For the most part, these systems and technologies were originally introduced as research projects and evolved to meet local institutional or academic needs.

To counter this trend, an alternative concept, Opencast[1], was introduced by UC Berkeley to explore an Open Source alternative to the production, distribution, management of and engagement with audiovisual content. Opencast recognized the numerous academic efforts emerging in isolation, and has created a landscape on which institutions can combine efforts and increase innovation around one project [9]. The work that is being presented in this paper is part of the development activity in Opencast.

Most of us would not consider watching a lecture recording as "interactive learning", especially if it is a replay of ex-cathedra teaching. Lecture recordings alone, even before adding advanced learning features, can be more engaging by their very nature. With a system that can capture more than the information delivered in a lecture, a video has the potential to engage the student on a more personal level than other kinds of e-learning content.

It allows the learner to make stronger connections to the lecturer as an actual person, as well as the other learners in the classroom, and benefit from the non-verbal and visual cues of the experience. Opencast Matterhorn engage applications go beyond a pure lecture video experience.

The applications do support fine-grained segment based navigation (based on video OCR), in-video search and multi-stream replay. Application accessibility has been a high priority throughout the development that allows full assistive technology support on many platforms. Figure 1 depicts a multi-stream engage application that presents the lecturer and a content stream in a synchronized way. In addition further content navigation elements have been implemented.

[1] http://www.opencastproject.org/

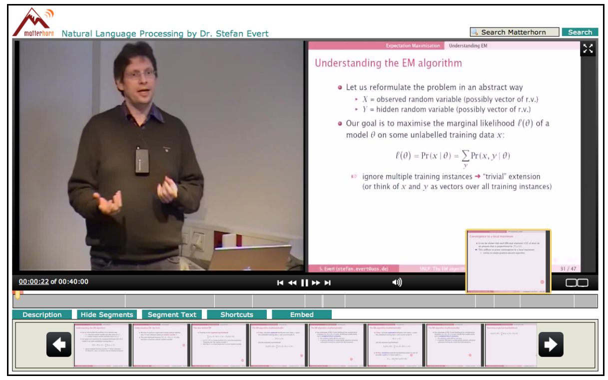

Figure 1: Opencast Matterhorn engage application

The history aware user awareness extensions are being implemented in the context of these applications and are currently in a prototype status. The ideas behind the user awareness features are being described in the next sections.

4. Users Prospect

The context-aware user-awareness feature described in this paper is designed for a special use case. When a user watches a web lecture and has a question about a specific passage in that lecture, he or she can find other users who are online and have already watched that very passage and are thus likely to be capable of answering that question. The overall workflow is shown in figure 2.

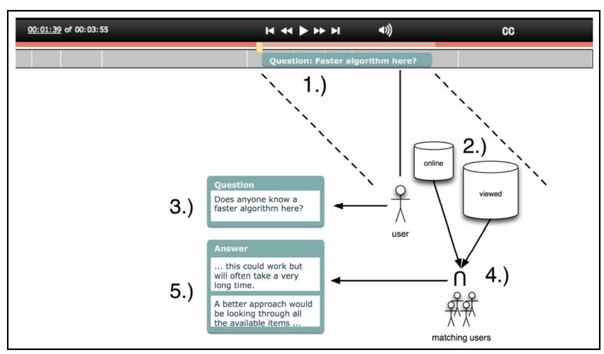

Figure 2: Workflow from user's perspective

First the user marks a passage in the web lecture he or she is currently viewing. This is done by selecting the passage's starting and ending points on the timeline (1). After the user has selected a passage, the system checks if any users who have watched that passage already are online (2). If no such users can be found, the system informs the user. If such users can be found, a message field opens in the interface. The user then types a message – preferably a question – about the passage (3). When the user sends this message, the system forwards it to all users who are online and have watched the passage and are online (4). These users can then start chatting with the user who posted the message (5).

The message is posted to all online users who have watched the passage in question. However, the user who has posted the question is not aware of who these users might be unless they answer his or her question. This way, the other users' privacy is protected.

In cases where privacy protection is not an issue, other mechanism can be imagined.

For instance, a list with all users that are online and have watched the passage in question could be shown after the user has marked the passage. This list could show the users' names and profile picture. In the current version of the system, such a feature is, however, not implemented.

5. Capturing Usage Data

When a user watches a video, so called footprints are generated automatically. These footprints store what passages of the video had been watched and can be used by the individual user to visually indicate what passages he or she has watched. This information can be useful to find new information or to find a passage the user remembers vaguely. The footprints are also evaluated in a central Data Warehouse. For this evaluation, each footprint is linked to an anonymised but unique avatar. This way, user behavior can be analyzed and correlated for different videos and even different lectures without revealing the user's identity. The system also stores metadata associated with the user. These metadata can vary depending on the video-players context. The video-player can be embedded in different contexts (open website, portal, learn-management-system with authentication). In some contexts only a limited amount of information on the user is available. In case of a social network, a complete user profile exists. Each of these contexts poses its own demands on the structure and granularity of the metadata associated with a footprint. The remainder of this article describes four prototypical classes of contexts.

The first class of contexts is a nearly context-free environment like a website without any text and just the video player embedded. In this case, only general information like the client's IP-address, time and date and possibly webpages linking to the site can be obtained. In this class of metadata for footprints (class 1), geo-coding could be used to analyze geographical disparities in learning behavior. This approach is, however, not elaborated on in this paper.

A second class of contexts (class II) is webpages with machine-readable content. This is the case whenever the video-player is embedded in a webpage that contains text. In this case, the metadata of class 1 can be enriched with information on the webpage's content. For this purpose, keywords can be extracted from the text and title of the surrounding webpage.

Class III contexts are given for web lectures that are accessible via a learn-management-system (LMS) with authentication only. In an LMS, users are linked to courses and a simple user profile that gives information about the user's status field of study. Also, users are identifiable since they had to log in using a personal password or similar authentication method. Stand-alone web lecture systems that require a user to log in also fall in this category.

The last class of contexts (class IV) can be found in web lectures that are embedded in social network applications [4]. Facebook or applications that support OpenSocial are a good example for social network applications. Social network applications usually come with user profiles that can store information about a user's professional life, private life and study activities. The amount of information stored is, however, highly dependent on how much information the user leaves and can also be a privacy issue. Additionally, a social graph exists, that captures relations between users (friend, study mate, co-worker, etc.). Users can also be members of a group. Members of a group share a status, an activity or a field interest. Hence, usage data from gathered web lectures in this context come with a wealth of information.

The approach described in this paper is realized by sending messages to users that have already watched the passage in question. Hence, these users have to be identifiable. This is only given for class III and IV contexts.

6. Finding Users

In the Data-Warehouse, all users that have already watched the passage in question are retrieved. This is done by analyzing all footprints stored for the web lecture in question.

A key challenge is finding those users who are suited best to answer the user's question. All users who have watched the passage marked by the user are can be sorted according to a list of criteria. These criteria could be:

- **How often has the user watched the passage?** This criterion is a double sided one. One the one hand, a user should be quite familiar with a passage he or she has watched more than once (for instance to prepare for an exam). On the other hand, repeated viewing could indicate that the user had problems understanding the passage's content. In the current implementation, the criterion is regarded positively, i.e. more views increase the ranking.

- **Has the user watched other videos from the same lecture?**
  Having watched other videos from the lecture give a user a broader knowledge about the topic. These users are thus ranked higher.
- **Is the user online?**
  This is a knock-out criterion. If the user is not online, he or she cannot answer the question online. Hence users who are not online are omitted.
- **Has the user written comments about this passage or a nearby one?**
  A user who has written comments is very likely to have good knowledge about the passage.
- **Has the user already answered questions about this passage or a nearby one?**
  A user who has already answered questions about the passage is also very likely to have good knowledge about the passage.
- **Has the user watched the whole passage?**
  In the current implementation, messages are only sent to users who have watched the whole passages. Percentages of coverage could also be used as a criterion.

In the current implementation, only the criteria "is the user online?" and "has the user watched the whole passage?" are taken into account. A message is send to only to those users who have watched the whole passage in question and are online.

7. Conclusion

Users privacy seems to be more important than ever before. Social networks and social application tempt to collect all kinds of data – oftentimes it is not transparent to the users what kind of data is collected and how it will be used in the future. User feedback from interviews and comments indicate that in particular student users do prefer a strict separation between a fun or leisure usage of these networks and the tools and applications they use for learning at universities.

We believe that social technology like the social awareness features presented in this paper help to build better tools for learners to engage with the material. Group and workspace awareness (see related work) as well as privacy are important trying to build sustainable tools for the learners. Standards like Open Social help to develop applications as independent as possible and allow control about the data that is being recorded, analyzed and published. Opencast Matterhorn is an ideal platform for building academic audio and video applications for learners. The work presented in this paper is one of the first steps to bring the social web also to this platform.

8. References

[1] Dieberger, A. "Supporting social navigation on the World Wide Web", International Journal of Human-Computer Studies, Volume 46, Issue 6, Pages 805-825, ISSN 1071-5819, June 1997.

[2] Dourish, P. and V. Bellotti. "Awareness and Coordination in Shared Workspaces", In proceedings of the ACM Conference on Computer supported cooperative work (CSCW'92), Toronto, Ontario. ACM Press., 1992.

[3] Farnham, S. et al., "Personal Map: Automatically Modeling the User's Online Social Network," Human–Computer Interaction (INTERACT 03), IOS Press, pp. 567-574, 2003.

[4] Fox, P., Emden, J., Neubauer, N. and Vornberger, O. "Integrating Lecture Recordings with Social Networks", Proceedings of the 2009 First International Conference on Advances in Multimedia. IEEE Computer Society, pp. 18-22, 2009.

[5] Gellersen, H. and Schmidt, A. "Look who's visiting: supporting visitor awareness in the web", Int. J. Hum.-Comput. Stud. 56, 1 (Jan. 2002), pp. 25-46, 2002.

[6] Gutwin, C. and Greenberg, S. "A Descriptive Framework of Workspace Awareness for Real-Time Groupware", Computer Supported Cooperative Work (CSCW), Volume 11, Numbers 3-4, pp. 411-446, 2002.

[7] Hudson, S.E. and I. Smith. "Techniques for Addressing Fundamental Privacy and Disruption Tradeoffs in Awareness Support Systems", in proceedings of the ACM Conference on Computer supported cooperative work (CSCW'96). 1996, Cambridge, MA. ACM Press.

[8] Hu, B., Kuhlenkamp, A., Reinema, R., Supporting Group Awareness in Web Based Learning Environments, pp. 525-536, 2002.

[9] Ketterl, M., Schulte, O., Hochman, A. "Opencast Matterhorn: A community-driven Open Source Software project for producing, managing, and distributing academic video", International Journal of Interactive Technology and Smart Education, Emerald Group Publishing Limited, in print, 2010.

[10] Lee, A., Girgensohn, A., Zhang, J. "Browsers to Support Awareness and Social Interaction," IEEE Computer Graphics and Applications, vol. 24, no. 5, pp. 66-75, Sep./Oct. 2004.

[11] Liechti, O. and Sumi, Y. 2002. “Editorial: Awareness and the WWW”, International Journal of Human-Computer Studies, 56, Academic Press: London, pp. 1-5, 2002.

[12] Mertens, R. Schneider, H., Müller, O. and Vornberger, O. (2004). Hypermedia Navigation Concepts for Lecture Recordings World Conference on E-Learning in Corporate, Government, Healthcare & Higher Education, Washington DC, USA, Nov. 2004, pp. 2480-2487

[13] Mertens, R., Farzan, R. & Brusilovsky, P. (2006). Social Navigation in Web Lectures. ACM Hypertext 2006, Odense, Danmark, August 2006. pp. 41-44.

[14] Palfreyman, K. and Rodden, T. 1996. A protocol for user awareness on the World Wide Web. In Proceedings of the 1996 ACM Conference on Computer Supported Cooperative Work (Boston, Massachusetts, United States, November 16 - 20, 1996). M. S. Ackerman, Ed. CSCW '96. ACM, New York, NY, 130-139.

[15] Plaisant, C. “LifeLines: Visualizing Personal Histories,” Proc. SIGCHI Conf. Human Factors in Com- puting Systems (CHI 96), ACM Press, 1996, pp. 221-227.

[16] Sack, W. “Conversation Map: A Content-Based Usenet Newsgroup Browser,” Proc. 5th Int’l Conf. Intelligent User Interfaces, ACM Press, 2000, pp. 233-240.

[17] Schmidt, A. and Gellersen, H. 2001. Visitor awareness in the web. In Proceedings of the 10th international Conference on World Wide Web (Hong Kong, Hong Kong, May 01 - 05, 2001). WWW '01. ACM, New York, NY, 745-753.

[18] Smith, M. A., and Fiore, A. T. “Visualization Components for Persistent Conversations,” Proc. SIGCHI Conf. Human Factors in Computing Systems (CHI 2001), ACM Press, 2001, pp. 136-143.

[19] Tsai, I., Nuankhieo, P., Wu, C. & Chang, Y. (2005). Supports of Group Awareness for Synchronous and Asynchronous Collaboration. In G. Richards (Ed.), Proceedings of World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education 2005 (pp. 2449-2454). Chesapeake, VA: AACE.