Evaluation of Social Personalized Adaptive E-Learning Environments: End-User Point of View

Lei Shi, Malik Shahzad Awan, and Alexandra I. Cristea

University of Warwick
Department of Computer Science, University of Warwick, Coventry, CV4 7AL. UK
{lei.shi, malik, acristea}@dcs.warwick.ac.uk

Abstract

The use of adaptations, along with the social affordances of collaboration and networking, carries a great potential for improving e-learning experiences. However, the review of the previous work indicates current e-learning systems have only marginally explored the integration of social features and adaptation techniques. The overall aim of this research, therefore, is to address this gap by evaluating a system developed to foster social personalized adaptive e-learning experiences. We have developed our first prototype system, Topolor, based on the concepts of Adaptive Educational Hypermedia and Social E-Learning. We have also conducted an experimental case study for the evaluation of the prototype system from different perspectives. The results show a considerably high satisfaction of the end users. This paper reports the evaluation results from end user point of view, and generalizes our method to a component-based evaluation framework.

1998 ACM Subject Classification H.3.4 Systems and Software, H.5.1 Multimedia Information Systems, K.3.1 Computer Uses in Education

Keywords and phrases adaptive educational hypermedia, social e-learning, evaluation

Digital Object Identifier 10.4230/OASIcs.ICCSW.2013.103

1 Introduction

Adaptive Educational Hypermedia (AEH) [1] has been designed to offer a personalized learning process. It combines the ideas from Adaptive Hypermedia [2] and Intelligent Tutoring Systems [22], with the aim of producing applications wherein learning contents is adapted to personalized needs. The rapidly emerging social networking apps, offers an opportunity to improve the adaptive e-learning experience by introducing a social dimension. The study on introducing social dimension into adaptive e-learning has recently accumulated a considerable set of theories and techniques, which have been used for building a variety of e-learning systems with both adaptive and social features. As these theories and techniques promote such systems from research labs to real usage, the evaluation becomes more crucial and even more important than proposing new but questionable theories and techniques [5].

This paper aims to 1) present the evaluation of Topolor [18] that was designed based on the concepts of AEH [1] and Social E-Learning [9], and 2) propose a novel component-based evaluation framework that was generalized from our evaluation method. The evaluation was conducted from end user point of view, which reflect critical feedbacks during the real usage. It also intends to minimize the level of biasedness arising from the answers of a user.

In the following, section 2 presents related works, including existing e-learning systems that support adaptation and social interaction, and the existing evaluation methods; section 3 introduces Topolor; section 4 describes the collected data and questionnaire, and reports the evaluation results; section 5 summarizes our evaluation methods to propose a generic evaluation framework; section 6 draws conclusions and outlines future research.

2013 Imperial College Computing Student Workshop (ICCSW'13).
Editors: Andrew V. Jones, Nicholas Ng; pp. 103–110
OpenAccess Series in Informatics
OASICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

2 Related Work

An AEH system aims at producing educational applications in which learning contents is adapted to every learner's personalized needs, such as knowledge level, learning goals and learning styles [3]. To extend an AEH system with social networking services, the support for knowledge networking and community building should be added to the system [8].

Several e-learning systems and services with both adaptive and social features have been developed in the last decade. For instance, Progressor [12] is a visual interface for open student modelling [13]. It provides students with a holistic and easy-to-grasp view on their progress and allows relating it to the progress of other students. QuizGuide [6] is an adaptive e-learning system that helps students in selecting relevant self-assessment quizzes assigned to topics. It uses adaptive annotation to support adaptive navigation, in order to provide personalized guidance to show which topics are more important and which require further study. Knowledge Sea II [4] uses social navigation to help students navigate from lectures to relevant online tutorials in a map style social navigation based on traffic and annotation.

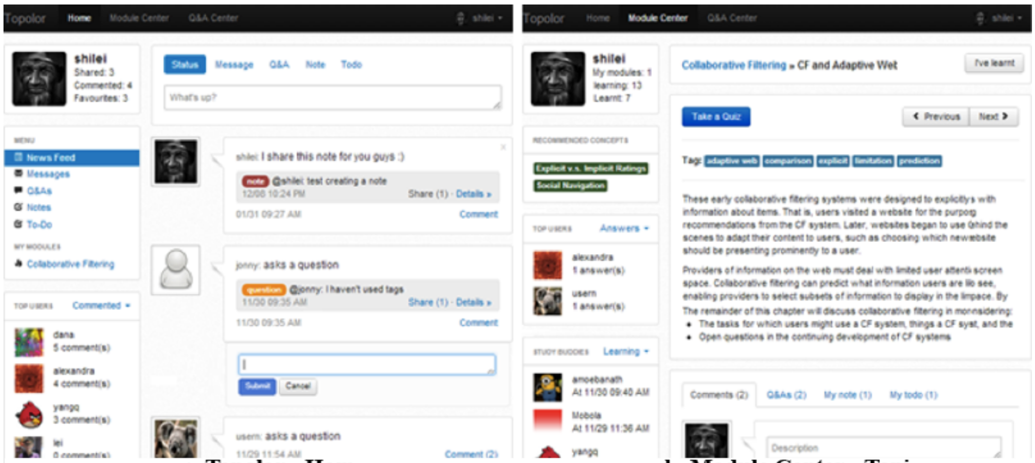
These systems have only focused on a few aspects of adaptive and social facilities in an e-learning environment, thus, requiring the development of more comprehensive systems that support the integration and mutual promotion of these facilities. Further, the flourishing development of social, personalized and adaptive e-learning systems requires more generic approaches to evaluate and compare different systems. Therefore, it is essential to develop new evaluation methods that provide wider coverage in an e-learning environment.

One important issue associated with the development of a social personalized adaptive e-learning system is selecting an effective evaluation method to capture its broader perspective. It's also important that the adopted evaluation method is appropriate and correct [10]. Several evaluation methods have been developed. For example, Chao [7] suggested 5 criteria to such as 1) e-learning material, 2) quality of web learning platform, 3) synchronous learning, 4) self-learning, and 5) learning record. Ozkan [14] proposed a 6-dimension framework 1) system quality, 2) service quality, 3) content quality, 4) learner perspective, 5) instructor attitudes, and 6) supportive issues. These evaluation methods are only able to cover limited perspectives, although they were trying to cover more. Besides, they have no ability to compare, classify or linguistically represent different evaluated e-learning systems.

3 Topolor

Topolor [20] is featured as a social personalized adaptive e-learning system, and has been used as an online learning environment at the University of Warwick. Topolor was developed based on a classical layered architecture (inspired by the Dexter model [11]), extended with a social flavour: a storage layer, a persistence infrastructure for physical entities; and a runtime layer, parsing adaptation strategies to present adaptive and/or adaptable learning materials, providing Web 2.0 tools for social interaction, and monitoring learners' behaviour [15].

Topolor has a Facebook-like appearance (Fig. 1a): the profile avatar and user information, the fixed-position top menu, the left navigation bar, and the information flow wall for social interaction. As shown in Fig. 1b, Topolor supports topic and learning path adaptation, which provides various levels of granularity of learning content adaptation, and personalized order of learning those recommended topics; it supports peer recommendation, which is based on each learner's learning history and previous performance, in order to provide opportunities for them to learning from each other; it also support social interaction, which provides Web 2.0 tools that they are familiar with, so that they can comment on, share a topic, a question, a note, etc. This is a much broader range of adaptation than in regular AEH systems.



a. Topolor - Home b. Module Center - Topic

Figure 1 The screenshots of Topolor: a. Topolor – Home; b. Module Centre – Topic.

4 Evaluation

To evaluate Topolor, an experiment was conducted with the help of 21 students at the University of Warwick. They were learned an online course on "collaborative filtering" using Topolor in a 2-hour intensive learning session. Before the session, a 'to-do list' was handed out to make sure they have a reminder of all functionalities at their disposal. The order of using the functionalities, and if to repeat was up to them. During the session, a logging mechanism kept track of each of their actions. After the session, they were asked to fill in an optional and anonymous questionnaire. 10 out of 21 students returned the questionnaire.

Based on the collected data, Topolor has been evaluated from different perspectives [16][17][21]. We have also investigated learning behaviour patterns of users using machine learning, educational data mining and data visualization techniques [19]. This paper focuses on only the evaluation from the end user's point of view based on the questionnaire analysis.

This method consists of 3 perspectives: *functionality*, *learning perspective* and *system prospect*. The following sections present the analysis process and report evaluation results.

4.1 Functionality

10 functionalities were evaluated from two performance aspects: usefulness and ease of use as below. The score values of the two considered performance aspects of each functionality ranged between 0.85-1.40 and 0.95-1.44 for 'usefulness' and 'ease of use' respectively. Figure 2 presents the classification of them on a Likert Scale from -2 to 2.

Overall-Sub. This represents the overall view of each subsystem.

Status. The status (post) functionality supports learners to publish and share their learning status and comment on each other's status.

Messaging. It aims at helping in making the intra-system communication more efficient.

Q&A. The questioning and answering functionality helps learners learn and manage the queries related to learning topics.

Note. The note management functionality records learners' personal thought related to learning topics. Notes can also be shared to other learners.

To-do. The to-do management functionality aims at helping learners arrange their own learning plans and remind them to finish their tasks.

Module. The module management functionality includes activities such as arranging topics within a module, reviewing topics learnt, accessing to a recommended topic.

Topic. The represents the functionalities that supports topic-learning activities such as accessing the previous and next learning topic according to the recommended learning path, discussing with others who are learning the same topic, commenting on the topic.

Testing. This includes taking quizzes for a learning topic and taking tests for a module (a set of organized learning topics). It also assesses the process of reviewing quizzes/tests, and getting access to the learning topics related to the questions in a quiz/test.

Statistics. This represents the statistics about the numbers of, e.g., how many topics a learner has learnt, how many questions a learner has asked/answered, how many status (post) a learner has commented on and shared, and so on.

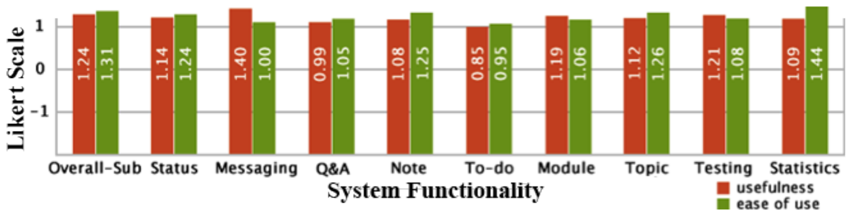


Figure 2 The result from the System Functionality Questionnaire.

4.2 Learning Perspective

The learning perspective was tested by the "Overall Topolor System – Five Scales for Believes" questionnaire, which include 7 questions shown as below. The individual score values for each of the questions are shown in Figure 3. **Q1.** I believe Topolor helps me learn more topics; **Q2.** I believe Topolor helps me learn more profoundly (deeply); **Q3.** I believe Topolor increases my learning outcome; **Q4.** I believe compare to other e-learning system, Topolor is easy to use; **Q5.** I believe compare to other e-learning system, Topolor is useful; **Q6.** I believe that the interaction with Topolor is easy to learn; **Q7.** I believe that the interaction with Topolor is easy to remember how to use.

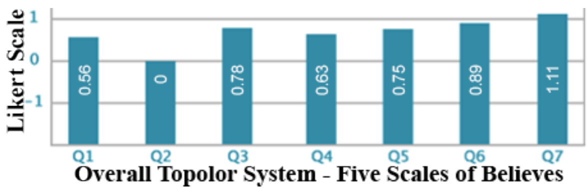


Figure 3 The result from the Overall Topolor System – Five Scales for Believes Questionnaire.

4.3 System Prospect

Two parameters were considered for evaluating the system prospects: 1) identification of the most useful features; and 2) improvements suggested by users. These two parameters were evaluated using a questionnaire with open-ended questions, in order to allow users to present a more practical and broader feedbacks on the features that were either helpful in learning or necessary to improve to be more useful features. From the questionnaire, four 'Best' features (Table 1) and four 'Need to Improve' features (Table 2) were identified respectively.

Table 1 The reported "Best Features".

Best Features	Frequency
Ask & answer questions	3
Filter learning content	3
Review quizzes	2
Discuss with others	2
Identified features: **4**	**10**

Table 2 Suggestions on improvements.

Improvements	Frequency
Share questions	4
To-do management	3
Interaction with contents	2
Searching tools	2
identified features: **4**	**11**

5 Evaluation Framework

We propose a component-based evaluation framework to generalize our methods that and overcome the shortcomings of existing methods. The proposed framework has 4 components. The first 3 evaluate an e-learning system from different perspectives. The 4th component combines the evaluation results of the first 3, summarizes and provides system classification and linguistic representation. The score values for the first 3 components are obtained from a questionnaire that system users fill in, after using the system for a given period of time.

5.1 Functionality

This component aims at using a Likert Scale to evaluate system functionalities from different performance aspects such as accessibility, effectiveness, operability, reliability, scalability and usability. We associate a weight, w with each considered performance aspect, which represents its significance. The score value of this component, C_{FUNC}, is calculated using Eq. 1a for representing the overall system value against the considered system functionalities. The score value is calculated by taking the weighted sum of the considered performance aspects, $SubSys_{(aspectID,subSysID)}$, and the associated weight, $w_{(aspectID,subSysID)}$, where *aspectID* represents a considered performance aspect; *subSysID* represents a considered sub-systems; m represents the number of the considered sub-systems. The generalized description of $SubSys_{(aspect,subSys)}$ for each considered system functionality represented as $F_{(aspectID,funcID,subSysID)}$, could be calculated using Eq. 1b, where *funcID* represents a considered functionality; $w_{(aspectID,funcID)}$ represents the corresponding associated weight; n represents the number of the considered system functionalities within the sub-system. $F_{(aspectID,funcID,subSysID)}$ could be calculated using Eq. 1c, where $q_{(i,j)}$ represents the j$^{\text{th}}$ question related to the considered system functionality, $F_{(aspectID,funcID,subSysID)}$, answered by the i$^{\text{th}}$ respondent; w_j represents the corresponding associated weight of this question; k represents the number of the questions related to the considered performance aspect of a considered system functionality; a represents the total number of respondents. The term *1/a* is used to minimize the level of biasedness arising from the answers of a respondent.

$$C_{FUNC} = \sum_{subSysID=1}^{m} SubSys_{(aspectID,subSysID)} \times w_{(aspectID,subSysID)} \tag{1a}$$

$$SubSys_{(aspectID,subSysID)} = \sum_{funcID=1}^{n} F_{(aspectID,funcID,subSysID)} \times w_{(aspectID,funcID)} \tag{1b}$$

$$F_{(aspectID,funcID,subSysID)} = \frac{1}{a} \times \sum_{i=1}^{a} \sum_{j=1}^{k} q_{(i,j)} \times w_j \tag{1c}$$

5.2 Learning Perspective

This component evaluates the impact that the learning system had on the overall learning experience as perceived by learners. It also evaluates the user-system interaction considered as helpful in learning as perceived by learners based on a Likert Scale against the defined criteria. The score value for this component, C_{LEARN} could be calculated using Eq. 2, where $q_{(j,j)}$ is the $\mathrm{j^{th}}$ question answered by the $\mathrm{i^{th}}$ respondent; w_j represents the corresponding associated weight of this question; k represents the number of the questions related to the considered performance aspect of a functionality; a represents the total number of respondents.

$$C_{LEARN} = \frac{1}{a} \times \sum_{i=1}^{a} \sum_{j=1}^{k} q_{(i,j)} \times w_j \tag{2}$$

5.3 System Prospect

This component identifies the learner characterization for the system features using simple questions grouped into different categories such as 1) identifying "the best feature(s)" of an e-learning system, 2) identifying the features that require further improvements. The score value for this component, 'C_{PROS}' could be calculated using Eq. 3a, where Freq_j could either represent the number of features identified as "the best feature(s)" by a learner or those requiring improvements; w_j represents the associated significance of the feature category, i.e. the most useful and those requiring improvements; k represents the distinct feature categories in the evaluation. Freq_j could be calculated by counting the frequency of an identified feature, $\mathit{FeatureCount}_i$, as grouped by the i^{th} respondent, and a represents the total number of respondents, as shown in Eq. 3b. A value ranging from -2 to 2 represents the frequency of a reported feature. For 0 or 1 reported feature, $\mathit{FeatureCount}_i$ has an assigned value -2; Similarly, for 2, 3 and 4 reported features, the $\mathit{FeatureCount}_i$ has an assigned value of -1, 0 and 1 respectively. For 5 or more features, the assigned value is 2.

$$C_{PROS} = \frac{1}{k} \times \sum_{j=1}^{k} \mathit{Freq}_j \times w_j \tag{3a}$$

$$\mathit{Freq}_j = \frac{1}{a} \times \sum_{i=1}^{a} \mathit{FeatureCount}_i \tag{3b}$$

5.4 Overall System Classification and Linguistic Representation

This component provides an overall system classification and linguistic representation for the evaluated system. Its score value could be calculated using Eq. 4, where C_{FUNC}, C_{LEARN} and C_{PROS} have been calculated using Eq. 1a, Eq. 2 and Eq. 3a respectively, and then rounded off the mean score value to the nearest integer value on a Likert Scale from -2 to 2 (-2: *very bad*; -1: *bad*; 0: *medium*; 1: *good*; 2: *very good*).

$$\mathit{System} = \left\| \tfrac{1}{3} \times (C_{SYSFUNC} + C_{LEARN} + C_{PROS}) \right\| \tag{4}$$

The measurements in this component mainly involved counting the frequency of the system features reported by the learners against the two parameters considered in this case study. To maintain consistency in the value scale with the other two framework components, namely, system functionality and learning perspective, we used the same scale, from -2 to 2, for assigning a score to each reported feature, and the score value is kept constant at 2 when the number of reported features exceeds 5. Likewise, for a non-reported feature or suggestion, a default score of -2 is added for maintaining consistency of results on the defined scale.

6 Conclusions and Future Work

In this paper, we have reported the evaluation of Topolor, a social personalized adaptive e-learning environment, and generalized the evaluation method to a component-based evaluation framework that provides broader and objective perspectives of a system evaluation report. The framework consists of 4 components, namely, *functionality*, *learning perspective*, *system prospect* and *system classification*. These components could be used for evaluating the effectiveness of a social personalized adaptive e-learning system from different perspectives with low level of biasedness. The evaluation involves the calculation of a score value using a component-specific mathematical model for each components. These score values are further used for calculating an overall rounded mean score value for the system. This rounded mean score value is then mapped to a Likert Scale for classifying the system into five (5) different categories, namely, 1) *very bad*, 2) *bad*, 3) *medium*, 4) *good* and 5) *very good*.

We further plan to broaden the data set for the evaluation of Topolor, so that, along with human experts, we could deploy advanced artificial intelligence techniques, e.g., artificial neural networks, case-based reasoning, machine learning, etc., for automatically learning the associated weights of framework components, system functionalities and question items during the system classification process. We further plan to extend the evaluation process by using fuzzy rule-based reasoning and consider the linguistic representation for classifying a social personalized adaptive e-learning system to present a more empirical view of system performance. Our future plans also include introducing mechanisms for estimating cost and effort associated with the measurement using our evaluation framework and perform a cost-and-benefit analysis for evaluating social personalized adaptive e-learning systems.

We claim that the proposed methodology is transferable to other social personalized adaptive e-learning systems and would enable researchers to: 1) gain a global view of assessments on an e-learning system; 2) evaluate each sub-system and functionality from different perspectives; 3) obtain learners perception of learning experience; 4) identify the best system features from learners' point of view and gain suggestions on the system improvement; 5) quantify the evaluation using component-specific mathematical models; and 6) classify the system on a Likert Scale for linguistic representation.

References

1 Peter Brusilovsky. Adaptive educational hypermedia. *International PEG Conference*, strony 8–12, 2001.

2 Peter Brusilovsky. Adaptive hypermedia. *User modeling and user-adapted interaction*, 11(1-2):87–110, 2001.

3 Peter Brusilovsky. Adaptive educational hypermedia: From generation to generation. *Proceedings of 4th Hellenic Conference on Information and Communication Technologies in Education, Athens, Greece*, strony 19–33, 2004.

4 Peter Brusilovsky, Girish Chavan, Rosta Farzan. Social adaptive navigation support for open corpus electronic textbooks. *Adaptive Hypermedia and Adaptive Web-Based Systems*, strony 24–33. Springer, 2004.

5 Peter Brusilovsky, Charalampos Karagiannidis, Demetrios Sampson. Layered evaluation of adaptive learning systems. *International Journal of Continuing Engineering Education and Life Long Learning*, 14(4):402–421, 2004.

6 Peter Brusilovsky, Sergey Sosnovsky, Olena Shcherbinina. Quizguide: Increasing the educational value of individualized self-assessment quizzes with adaptive navigation support. *World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education*, wolumen 2004, strony 1806–1813, 2004.

7 Ru-Jen Chao, Yueh-Hsiang Chen. Evaluation of the criteria and effectiveness of distance e-learning with consistent fuzzy preference relations. *Expert Systems with Applications*, 36(7):10657–10662, 2009.

8 Mohamed Amine Chatti, Matthias Jarke, Dirk Frosch-Wilke. The future of e-learning: a shift to knowledge networking and social software. *International journal of knowledge and learning*, 3(4):404–420, 2007.

9 Stephen Downes. Feature: E-learning 2.0. *Elearn magazine*, 2005(10):1, 2005.

10 Cristina Gena, Stephan Weibelzahl. Usability engineering for the adaptive web. *The adaptive web*, strony 720–762. Springer-Verlag, 2007.

11 Frank Halasz, Mayer Schwartz, Kaj Grønbæk, Randall H Trigg. The dexter hypertext reference model. *Communications of the ACM*, 37(2):30–39, 1994.

12 I-Han Hsiao, Julio Guerra, Denis Parra, Fedor Bakalov, Birgitta König-Ries, Peter Brusilovsky. Comparative social visualization for personalized e-learning. *Proceedings of the International Working Conference on Advanced Visual Interfaces*, 303–307. ACM, 2012.

13 Antonija Mitrovic, Brent Martin. Evaluating the effect of open student models on self-assessment. *Journal of Artificial Intelligence in Education*, 17(2):121–144, 2007.

14 Sevgi Ozkan, Refika Koseler. Multi-dimensional students' evaluation of e-learning systems in the higher education context: An empirical investigation. *Computers & Education*, 53(4):1285–1296, 2009.

15 Lei Shi, Dana Al Qudah, Alexandra I. Cristea. Designing social personalized adaptive e-learning. *The 18th Annual Conference on Innovation and Technology in Computer Science Education (ITiCSE 2013)*, 2013.

16 Lei Shi, Malik Shahzad K Awan, Alexandra I Cristea. Evaluating system functionality in social personalized adaptive e-learning systems. *Scaling up Learning for Sustained Impact*, strony 633–634. Springer, 2013.

17 Lei Shi, Dana Al Qudah, Alexandra I. Cristea. Social e-learning in topolor: a case study. *IADIS International Conference e-Learning*, 2013.

18 Lei Shi, Dana Al Qudah, Alaa Qaffas, Alexandra I Cristea. Topolor: a social personalized adaptive e-learning system. *User Modeling, Adaptation, and Personalization*, strony 338–340. Springer, 2013.

19 Lei Shi, Alexandra I. Cristea, Malik Shahzad Awan, Craig Stewart, Maurice Hendrix. Towards understanding learning behavior patterns in social adaptive personalized e-learning systems. *The 19th Americas Conference on Information Systems (AMCIS 2013)*, strony 708–711. Springer-Verlag, Berlin, Heidelber, 2013.

20 Lei Shi, George Gkotsis, Karen Stepanyan, Dana Al Qudah, Alexandra I. Cristea. Social personalized adaptive e-learning environment - topolor: Implementation and evaluation. *The 16th International Conference on Artificial Intelligence in Education (AIED 2013)*, strony 708–711. Springer-Verlag Berlin Heidelberg, 2013.

21 Lei Shi, Karen Stepanyan, Dana Al Qudah, Alexandra I. Cristea. Evaluation of social interaction features in topolor - a social personalized adaptive e-learning system. *The 13th IEEE International Conference on Advanced Learning Technologies (ICALT 2013)*, 2013.

22 Derek Sleeman, John Seely Brown. Intelligent tutoring systems. 1982.