Behavioral patterns in hypermedia systems: a short study of e-commerce vs. e-learning practices

A. Stefani, B. Vassiliadis, M. Xenos

Computer Science, Hellenic Open University
Greece
{stefani,bb,xenos}@eap.gr

1. Introduction

Web based systems are extremely popular to both end users and developers thanks to their ease of use and cost effectiveness respectively. Two of the most popular applications of web based systems nowadays are e-learning and e-commerce. Despite their differences, both types of applications are facing similar challenges: they rely on a "pull" model of information flow, they are hypermedia based, they use similar techniques for adaptation and they benefit from semantic technologies (Brusilovsky and Nejdl, 2004). The underlying business models also share the same basic principle: users access digital resources from a distance without the physical presence of a teacher or a seller. The above mentioned similarities suggest that, at least, some user behavioral patterns are similar to both applications.

It is at least intriguing to notice that many published works relating to e-learning adaptation focus on machine learning techniques for log analysis or static user profiles with common characteristics. Despite the large number of related publications, very limited claims of success of such traditional adaptation techniques have been reported under real life e-learning situations (Xenos et al., 2002). Only within the past few years, researchers, originating mainly from the educational domain point out that little has been done for context-aware adaptation (Stash et al., 2004; Wolf, 2003; Spiliopoulou et al., 2002). In the case of e-learning, context-aware parameters that seem to have been ignored to date, stem from behavioral and pedagogical theories. From our point of view, a simple but yet important parameter still eludes many e-learning adaptive hypermedia efforts: systems and techniques should be designed by both engineers and educationalists in order for research to bear fruits in real situations.

Although adaptability of web based systems has been the focus of a wide range of recent research efforts, these efforts are mainly focused on e-commerce. Most existing adaptation techniques for web-based systems are based on log analysis, user modeling or pre-determined rules (Brusilovsky and Maybury, 2002). E-learning adaptation has many

common characteristics with e-commerce applications and many differences as well. This means that advances in e-commerce adaptation may be partially used for e-learning adaptation (Dolog et al., 2004; Gams and Reich, 2004; Kim et al., 2000). Adaptive Web-based e-learning systems are a class of adaptive Web systems, an attempt to replace the classic "one size fits all" approach to hypermedia.

In this work, we discuss some of the major similarities and differences between e-learning and e-commerce systems (and particularly B2C systems) that affect adaptation and system usability. First, a presentation of the state of the art in adaptation techniques will be presented. These techniques, focused more on the e-commerce field, are analyzed in the light of e-learning adaptation requirements. Special attention will be given to adaptation that relies on user modeling since a great deal of research effort has been placed on this area. We also suggest that strategies used for user modeling comprise the major difference between the two types of applications. Our methodology will then examine the different adaptation needs of e-learning and e-commerce in terms of diversity, context and background. Architectural issues are also of importance since different techniques impose different architectural styles for the same category of systems. Theoretical comparison suggests that crude e-commerce adaptation techniques are not very well suited for closed and formal e-learning systems. However, they are quite useful and cost-effective in open and informal systems. As formal systems we define applications that are used by organizations with well structured procedures such as Open or Virtual Universities. Informal learning systems are often free internet applications targeted to wide user populations. Finally, we conclude that machine learning techniques used for adaptation should take into account behavioral and educational theories for distance learning in order to have a serious impact.

2. Adaptive hypermedia systems: state of the art

Adaptation in hypermedia systems followed initially two distinct approaches that later converged: adaptation based on rules and on algorithmic methods.

Rule based adaptation was dominating the field before algorithmic methods became a trend. A few years ago, most existing systems realized adaptation through the use of pre-determined rules, which assigned adaptation constituents to interaction situations in a rather "arbitrary" way. These rules are usually hard-coded in the user interface, and cannot be easily modified or reused across different applications. Additionally, there are no explicit representation of the goals underlying adaptation processes, and, in this sense, the latter cannot be taken explicitly into account in the

adaptation process. Decision-theoretic frameworks for run-time adaptation, are mostly utility-based decision making techniques in the context of the standard reference model for intelligent hypermedia presentation systems. Recent trends examine the efficiency of hybrid rule-based/ algorithmic methods for adaptation (Mendes et al., 2003).

A very popular, algorithmic approach for discovering user behavior in hypermedia systems is a special area of data mining: web mining (Han and Kamber, 2001). In turn, Web mining is generally focused on content, structure and usage mining (Chakrabarti, 2002). Web mining relies on standard data mining algorithmic approaches which face two major problems in this particular application domain: weak relations between users and the hypermedia system, and complicated behaviors (Wang and Meinel, 2004). The first problem is created by the fact that users often do not have clear goals or the system does not have enough information about them to adapt to their preferences (as in the case of first time or infrequent users). The second one is caused by the diversity of the user population.

Content mining analyses the content of hypermedia documents using information retrieval techniques. Usage mining initiated largely by approaches like the one of (Srivastava et al., 2000), gained greater attention and presented significant results in the next few years (Pierrakos et al., 2003; Eirinaki and Vazirgiannis, 2003; Srikant and Yang, 2001). A special area of web usage mining is collaborative (social) filtering (Kim et al., 2004) which refers to the categorization users to groups according to their preferences. The most popular methods are Markov models, association rules, sequential patterns, most forward access patterns (Chen et al., 1998), tree structure access patterns (Zaki, 2002), clustering and hybrid models.

User modelling is a popular adaptation method and already counts several years of application to a variety of systems. A user model maintains an explicit and dynamic representation of the user. It represents the system's understanding of a user and it consists both of known facts about the user (such as personal information) and inferred beliefs based on previous interactions. A typical model classifies user-related information into four main categories (Vassiliadis et al., 2003):

- Personal information (e.g. age, sex, preferences)
- Information about how the user interacted with the offered services (e.g. path used)
- Information about services the user has used
- Explanation of the result of specific service actions (e.g. unsuccessful buying attempts)

E-learning was between the first category of applications, and probably the first among the on-line ones, that used user models for adaptation. Early approaches used machine learning techniques initially developed for the information system domain to support adaptation. The advent of the

internet and the spectacular increase of the user population, both in number and diversity, posed the need for new approaches. New methods focus on the way in which learners acquire, store, process and share knowledge rather than forcing them to follow a stereotype, predetermined behavior (Tsiriga and Virvou, 2004; Schewe et al., 2005).

Learner (student) modeling has been extensively researched because there is already a strong background on pedagogical theories (De Bra, 2002; De Bra et al., 1999). The most interesting approaches include the overlay, the perturbation, the analytic, the synthetic and the mixed model.

Static buyer profiles were the primary source of knowledge in early implementations of adaptive B2C applications. A striking difference with student modeling is that comparatively little progress has been made in the development of user modeling components for e-commerce systems (Strachan et al., 2000).

Besides traditional approaches that produce user models based on log file analysis or questionnaires, works worth mentioning are those applied to live help systems for e-commerce web sites (like the one of Aberg et al., 2001) and recommender mechanisms.

Although we will not discuss in detail these approaches, it is obvious that there is no underlying theory in constructing buyer models nor any standardization efforts. Although machine learning does offer solutions to some problems of buyer modeling, it cannot be considered as a panacea with the argument that there are no other approaches/theories to be used.

4. Adaptation requirements: e-learning vs. e-commerce

A fundamental question in this comparison study is 'what should be adaptable and how?' (Rumetshofer and Wöβ, 2003). In this section, we explore requirements that affect the design of adaptation mechanisms.

The answer to the question of what can be adapted is mainly of a technological nature, while the one referring to how, a pedagogical - technological one. Content, layout and navigation are the three types of adaptation common in most adaptive hypermedia systems. Adaptation strategies should also be context specific. In the case of e-commerce, they rely on behavioral, social, marketing and other theories. In the case of e-learning, they rely on pedagogical and learning theories.

User behavior differs significantly. According to (Moe, 2001), B2C user behavior falls under four general categories: directed buying, search/deliberation, hedonic browsing and knowledge-building. The first two are goal-oriented, while the other two resemble explorative search behavior. On-line learning behavior, on the other hand, may follow the social/collaborative, contextualized or experiential model. There is an obvious difference in complexity but also an interesting point: knowledge

building is also used in e-commerce: navigation and hyperlinked data aid collaborative knowledge building which is becoming available for all kinds of adaptive hypermedia systems (Wu et al., 2003).

It is obvious that, although both e-learning and e-commerce are largely hypermedia based and use the same internet protocols and technologies to work, user goals differ significantly. Starting from e-learning applications, user goals are, ideally, to reach a set of predefined educational objectives, to learn. These objectives are the same for both the designers and the users of such system. E-learning is closely connected to educational, pedagogical and behavioral theories. Most current applications follow the information transfer paradigm where information is passed from the system to the user. Advanced learning models, the future trends in e-learning, anticipate knowledge construction and sharing and most importantly collaboration through the formation and management of virtual learning communities. Most formal platforms support the on-line presence of tutors (either real persons or avatars). Since e-learning uses on-line resources in its core mechanism, requirements such as efficient browsing and searching mechanisms are necessary. But how can context be described without the proper metadata? Semantic enriched, context-aware hypermedia may be the solution.

In the case of e-commerce and especially B2C (Business to Consumer) applications, goals are somewhat different for designers and users. Designers follow marketing strategies in order to sell as many products as possible. This includes making the user experience as seamless as possible, recommender systems and automatic offers. User goals are the same as every buyer goals: locate the appropriate product as simply as possible and access as much and relevant information as possible. It is obvious that designers and users, in this context, do not share the same goals in so far as an underlying theory is concerned. Marketing theories are followed by designers for increasing profits, not by users to directly gain benefit. In B2C, information is transferred from the application to the user and there is no direct possibility of knowledge building as in the case of e-learning. There is, however, the support of knowledge sharing through the use of off-line collaboration (mainly forums where buyers exchange opinions). Another difference is the absence of on-line guidance in the strict and formal form of tutoring. Off-line consultation with experts, or on-line support is significant but not as crucial as in the case of e-learning. On-line communities are present in the form of communities of practice and share significant information. They are however, in most cases, informal and communicate mostly by off-line tools.

New trends in pedagogy concentrate more on constructivism (Duffy and Jonassen, 1992), the building of knowledge by way of social interaction and collaboration on-line. Although constructivism has been identified by many researchers as one of the most appropriate methods for learning,

science has not yet comprehended and analyzed the mental processes of human knowledge building, collaboration, sharing, evolution and reuse. Thus, a learner does not behave exactly as a buyer does. Another important parameter is the identification of how users perceive and process information and how they complete tasks. It seems that the "one-size fits all" approach has proved to be relatively successful in e-commerce, which is not the case in e-learning. The number of different learning styles is large, and each of them is largely affected by numerous context parameters known only to the teaching staff. So, traditional adaptation mechanisms are somewhat superficial. One interesting proposal by (Stash et al., 2004) is to let the teaching staff configure adaptation by choosing the appropriate learning style that best describes the educational context. This approach has merit in theory, but it is difficult to be accomplished in practice since such software may become too difficult to use. Flexibility is the key to the success of this approach.

A popular misconception is that adaptation in both e-learning and e-commerce is governed by the same principles (Li and Zaiane, 2004). Techniques such as log analysis with the use of machine learning techniques, or general purpose user profiling will do the job in both cases. Although the above mentioned strategies are quire successful in the e-commerce paradigm, this is not always the case in e-learning. Most efforts fail to take under consideration pedagogical models and educational goals. E-learning adaptation in formal systems is more about sequencing learning material and workflow of learning processes. E-commerce recommendation/adaptation mechanisms simply will not do because they rely on common beliefs (preferences) which are often misconceptions and possibly not quite useful pedagogically. For example, the logical structure of a course is not taken into account when links or documents are recommended by an adaptation mechanism. This way, a pedagogically false sequence may be initiated by common user mistakes. E-learning is a procedure that is guided by formal theories, with well defined goals and methods in order to impose some kind of pedagogical control over the learning process.

These are some of the differences that drive the diversity of user behavior in e-learning and e-commerce.

5. Conclusions

User behaviour is diverse in e-learning and e-commerce hypermedia applications. Furthermore, different research approaches have flourished in these domains: in the former user modeling and in the latter machine learning. The main problem in current implementations is that these techniques are used in a straightforward way without any tailoring. E-

learning adaptation uses machine learning techniques mostly used in e-commerce resulting in poor efficiency. E-commerce has benefited from student modeling approaches but missing underlying theories produce static user models.

In this work, we reviewed the state of the art in discovering user behavior in both hypermedia contexts and briefly discussed differences and similarities. We argued that data mining techniques used in e-commerce should be used as a basis for e-learning. In fact, they should be combined with pedagogical approaches and theories.

Although the field of hypermedia adaptation is huge we believe that we made a small contribution to the endeavor of real adaptation in hypermedia. Our main argument, although simple but overlooked, is that adaptation should be context-specific and thus be in accordance with the underlying theory. Future work includes a more thorough investigation of social, cultural and economic factors that impose diversity in adaptive hypermedia systems and some recommendations particularly for e-learning and e-commerce applications. Furthermore, to investigate how new architectural models such as service oriented computing affect traditional adaptation in hypermedia.

Acknowledgements

This work was supported by the EPEAEKII Iraklitos programme (code: 88738) at HOU.

References

Aberg J, Shahmehri N, Maciuszek D (2001) User modelling for live help systems: initial results. Proceedings of the 3rd ACM conference on Electronic Commerce, pp 194 – 197

Brusilovsky P,. Maybury MT (2002). From adaptive hypermedia to adaptive Web. Communications of the ACM, 45(5): 31-33

Brusilovsky P, Nejdl W (2004). Adaptive hypermedia and adaptive web. In:Singh M, (ed) Practical Handbook of Internet Computing. Chapman & Hall/CRC

Chakrabarti S, (2002). Mining the Web: Analysis of Hypertext and Semi Structured Data. Morgan Kaufmann

Chen MS, Jong JS, Yu PS (1998). IEEE Transactions on Knowledge and Data Engineering, 10(2):209 – 221

De Bra P (2002). Adaptive educational hypermedia on the web. Communications of the ACM, 45(5):60 – 61

De Bra P, Brusilovsky P, Houben GJ (1999). Adaptive hypermedia: from systems to framework. ACM Computing Surveys (CSUR), 31(4), Article No. 12

Dolog P, Henze N, Nejdl W, Sintek M, (2004). Personalization in Distributed e-Learning Environments. In Proceedings of 13th World Wide Web Conference, pp 170-179

Duffy TM, Jonassen DH (1992). Constructivism and the Technology of Instruction: A Conversation. Lawrence Erlbaum Associates.

Eirinaki M, Vazirgiannis M (2003). Web mining for web personalization. ACM Transactions on Internet Technology (TOIT), 3(1):1-27

Gams E, Reich S (2004). An analysis of the applicability of user trails in web applications. 2004 Web Engineering Workshop

Han J, Kamber M (2001). Data Mining: Concepts and Techniques. Morgan Kaufmann, San Francisco

Kim DH, Atluri V, Bieber M, Adam N, Yesha Y, (2004). Web personalization: A clickstream-based collaborative filtering personalization model: towards a better performance. Proceedings of the 6th annual ACM international workshop on Web information and data management, pp 88 - 95

Li L, Zaiane OR (2004) Combining Usage, Content and Structure Data to Improve Web Site Recommendation. 5th International Conference on Electronic Commerce and Web Technologies (EC-Web 04), Springer Verlag LNCS 3182, pp 305-315

Mendes E, Mosley N, Counsell S (2003). Do adaptation rules improve web cost estimation?. Proceedings of the fourteenth ACM conference on Hypertext and hypermedia, Nottingham, UK, pp 173-183

Moe W (2001) Buying, searching, or browsing: Differentiating between online shoppers using in-store navigational clickstream. Journal of Consumer Psychology, 13(1&2):29-40

Pierrakos D, Paliouras G, Papatheodorou Ch, Spyropoulos CD (2003). Web Usage Mining as a Tool for Personalization: A Survey. User Modeling and User-Adapted Interaction, 13(4):311-372

Rumetshofer H, Wöβ W (2003). An approach for adaptable learning systems with respect to psychological aspects. 2003 ACM Symposium on Applied Computing (SAC 2003), pp 558-563

Schewe KD, Thalheim B, Binemann-Zdanowicz A, Kaschek R, Kuss T, Tschiedel B (2005). A Conceptual View of Web-Based E-Learning Systems. Education and Information Technologies, 10(1&2):83 – 110

Spiliopoulou M, Pohle C, Teltzrow M (2002). Modelling Web Site Usage with Sequences of Goal-Oriented Tasks. Multikonferenz Wirtschaftsinformatik, in: E-Commerce - Netze, Märkte, Technologien, Physica-Verlag, Heidelberg

Srikant R, Yang Y (2001). Mining Web Logs to Improve Web Site Organization. Proc. Of the WWW01 Conference, pp 430-437

Srivastava J, Cooley R, Deshpande M, Tan PN (2000), Web usage mining: discovery and applications of usage patterns from Web data. ACM SIGKDD Explorations Newsletter, 1(2):12-23

Stash N, Cristea A, De Bra P (2004). Authoring of Learning Styles in Adaptive Hypermedia: Problems and Solutions. WWW 2004, May 17–22, 2004, New York, New York, USA, pp 114, 123

Strachan L, Anderson J, Sneesby M, Evans M (2000). Minimalist User Modelling in a Complex Commercial Software System. User Modeling and User-Adapted Interaction, 10(2-3):109 – 146

Tsiriga V, Virvou M (2004). A Framework for the Initialization of Student Models in Web-based Intelligent Tutoring Systems. User Modeling and User-Adapted Interaction, 14(4):289 – 316

Vassiliadis B, Makris C, Tsakalidis A, Bogonikolos N (2003). User Modelling for Adapting and Adaptable Information Retrieval. Journal of Applied System Studies, 4(1).

Wang L, Meinel C (2004). Behaviour Recovery and Complicated Pattern Definition in Web Usage Mining. In: Koch N, Fraternali P, Wirsing M (eds.) LNCS 3140. Springer-Verlag Berlin Heidelberg, pp 531–543

Wolf C (2003). iWeaver: Towards 'Learning Style'-based e-Learning in Computer Science Education. Australasian Computing Education Conference (ACE2003), Adelaide, Proceedings of the fifth Australasian computing education conference on Computing education 2003 - Volume 20, pp 273 - 279

Wu H, Gordon MD, DeMaagd K, Bos N (2003). Link analysis for collaborative knowledge building. Proceedings of the fourteenth ACM conference on Hypertext and hypermedia, Nottingham, UK, pp 216 – 217

Xenos M, Pierrakeas C, Pintelas P (2002). Survey on Student Dropout Rates and Dropout Causes Concerning the Students in the Course of Informatics of the Hellenic Open University. Computers & Education, 39(4):361-377

Zaki MJ (2002). Efficiently Mining Frequent Trees in a Forest. Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining, Edmonton, Alberta, Canada, pp 71 – 80