The Value of QoE-Based Adaptation Approach in Educational Hypermedia: Empirical Evaluation

Cristina Hava Muntean and Jennifer McManis

Performance Engineering Laboratory, School of Electronic Engineering
Dublin City University, Glasnevin, Dublin 9, Ireland
{havac, mcmanisj}@eeng.dcu.ie

Abstract. This paper reports the results of a comparison-based empirical study on the applicability of the end-user Quality of Experience-based content adaptation mechanism in adaptive educational hypermedia. The focus of the paper will be the experiment itself: the initial settings, testing scenarios and the results. We will show that for low bit rate connections the QoE-based adaptation decreases study session time, information processing time per page and the number of re-visits to a page, it maintains similar learning outcomes while also improving the user quality of experience and satisfaction with the system. Finally we will comment on the results and interpret them.

1 Introduction

New communication technologies can enable Web users to access personalised information "anytime, anywhere". However, the network environments used for accessing the information may have widely varying performance characteristics (e.g bandwidth, level of congestion, mobility support, cost of transmission, etc). It is unrealistic to expect that the quality of delivery of content can be maintained at the same level in this variable environment. Rather an effort must be made to fit the content served to the current delivery conditions, thus ensuring high Quality of Experience (QoE) to the users. Currently, the adaptive hypermedia research places very little emphasis on end-user QoE and performance. However, it should be noted that some Adaptive Hypermedia Systems (AHS) have taken into consideration some performance features such as device capabilities (GUIDE [1], MP^3 [2]), the type of the access, communication protocol (SHAAD [3]), state of the network (SHAAD [3]) etc. in order to improve the end-user QoE. However, these account for only a limited range of factors affecting QoE. In order to respond to this problem we have conducted a more complete analysis of the key factors that affect QoE. We also proposed a QoE adaptation layer that extends the adaptation functionally of the AHS and ensures high level of QoE when users access personalised material via various connectivity network environments [4].

This paper presents an empirical-analysis of the experimental study conducted applying QoE layer in adaptive educational hypermedia to ascertain its effects on the learning process. The analysis compares learner outcome, learning performance, usability and visual quality for an AHA! [5] courseware application with and without application of the QoE adaptation layer.

V. Wade, H. Ashman, and B. Smyth (Eds.): AH 2006, LNCS 4018, pp. 121 – 130, 2006.

2 End-User QoE-Based Adaptation Overview

Most of the AHS proposed in the literature follow the abstract representation of the adaptation process illustrated in the AHAM [6] model. They construct and maintain a Domain Model (DM), a User Model (UM) and an Adaptation Model (AM). Starting from this simple abstract representation we extended the adaptation functionality with a novel content-based adaptation mechanism (called QoE Layer) that improves the end-user Quality of Experience (QoE) [7] when the Web browsing process takes place over heterogeneous and changeable network environments. While the AHS's main role of delivering personalised content is not altered, its functionality and performance is improved and thus the user satisfaction with the service provided.

To better understand the framework for our experiment, a brief presentation of the QoE layer architecture is presented next. More details were provided in other papers [4, 7, 8]. The QoE layer includes the following components: Perceived Performance Model (PPM), Performance Monitor (PM) and Adaptation Algorithm (AA). PM monitors different performance metrics (e.g. download time, round-trip time, throughput, user tolerance for delay) and user behaviour-related actions (e.g. abort request) in real-time during user navigation and delivers them to the PPM. The PPM models this information using a stereotype-based technique and probability and distribution theory [8], in order to learn about the user's operational environment characteristics, about changes in the network connection and the consequences of these changes on the user's quality of experience. Based on this information the PPM suggests optimal Web content characteristics (e.g. the number of embedded objects in the Web page, the dimension of the based-Web page without components and the total dimension of the embedded components) that will provide a satisfactory QoE. The Adaptation Algorithm uses PPM's content-related suggestions and applies various transformations [4] (e.g. modifications in the properties of the embedded images and/or elimination of some of them and placing a link to the image) to the personalised Web page (designed by the AM based on a user profile).

In this context, the adaptation process of an AHS allows for both *user-based* and *QoE-based* adaptations. *User-based* adaptation selects the fragments of information from the DM for inclusion in a user-tailored performance-orientated document, based on the user profile (UM). *QoE-based* adaptation is applied when the delivery of the personalised document over a given connectivity environment would not provide a satisfactory end-user QoE. In this case, the AA adjusts the characteristics of the personalised document.

In conclusion, the main goal of a QoE-aware AHS is to provide personalised material that suits both user's individual characteristics (e.g. goals, knowledge) and the delivery environment in order to ensure high QoE.

3 Context of the Study

As already mentioned, QoE layer aims at enhancing the adaptation functionality in order to improve the overall user experience with AHS when the user's operational environment has widely varying performance characteristics (e.g. bandwidth, delay, level of congestion, mobility support) that may affect the QoE. Therefore, the evaluation of the

QoE layer involved its deployment on the open-source AHA! system, creating QoEAHA. QoEAHA was applied in the educational area as an adaptive courseware application in order to test user perception of content delivery and its effect on learning outcome. The empirical study presented in this paper aims at a comparison-based evaluation of the QoEAHA in a home-like operational environment.

The subjects involved in the experiment were comprised of sixty-two post-graduate students from Faculty of Engineering and Computing at Dublin City University, randomly divided in two groups. They were required to perform two sets of task-based scenarios. Learning outcome, learning performance, system's usability and effectiveness, correlation analysis as well as assessment of the visual quality of the content and user satisfaction were analysed when using QoEAHA and the AHA! system. The impact of the end-user QoE on the student learning performance was also investigated.

4 Design of the Empirical Evaluation

The experiment took place in the Performance Engineering Laboratory, Dublin City University. Two sets of task-based scenarios ("interactive study" and "search for information") were developed and carried out. The first group of subjects used QoEAHA, whereas the second group used AHA!. The subjects were not aware of the type of the system they were using. The educational material consisted of an adaptive tutorial on the AHA! v.2.0 [5] application. None of the students had accessed the material prior the tests. No time limitation was imposed on the execution of tasks.

For both groups the same network conditions were emulated between the subjects' computers and the two systems. NISTNET, a network emulator that allows for the emulation of various network conditions characterised by bandwidth, delay, and loss was used to create low bit rate home-like Internet environments with network speed between 28 kbps and 128 kbps.

4.1 Scenario 1: Interactive Study

Scenario One covered an interactive study session on chapter one from the adaptive tutorial over various connection types. At the start of the study session the subjects were given a short explanation of both system usage and their required duties such as:

- *complete an on-line pre-test questionnaire* [9] with six questions related to the learning topic that aims at assessing subjects' prior knowledge on studied domain.
- *log onto the system* and proceed to browse and *study the material.* Back and forward actions through the studied material were permitted.
- at the end of the study period, *complete a post-test questionnaire* [9] consisting of fifteen questions that tests recollection of facts, terms and concepts as well as knowledge level after completing a study session.
- *Answer an usability questionnaire* [9] consisting of ten questions categorised into navigation, accessibility, presentation, performance and subjective feedback.

In order to fully assess the subjects learning outcome, both pre- and post-tests were such devised that consisted of a combination of four different types of test-items, commonly used in the education: “Yes-No”, “Forced-Choice”, “Multi-Choice” and “Gap-Filling” test items [10]. Each test-item type has different degree of difficulty and therefore a different weight in the final score is assigned for a correct answer. The final scores were normalised in the 0-10 range. Learning outcome, learning performance and system usability were assessed and forty-two students took part in this scenario.

4.2 Scenario 2: Search for Information

This scenario involves visual quality assessment that makes use of an evaluation criteria often used in the educational area: time taken to search for a term described in a page. Since the QoE adaptation mechanism involves applying various modifications on the properties of the embedded images, this scenario has an important role. The goal is to assess whether the quality of images is good enough to perform the required task.

The subjects were asked to look up two different terms and to answer two questions related to these terms. The terms were located on two pages of the tutorial, more precisely, in the embedded images. As these pages include the largest number of embedded images, the highest image quality degradation was caused when using QoEAHA in the tested conditions.

Objective and subjective visual quality assessments were performed involving twenty students. Time taken to complete the task and questionnaire-based evaluation techniques were used. The subjective visual quality assessment on a five-point quality scale (1-“bad”, 2-“poor”, 3-“fair”, 4-“good”, 5-“excellent”) ascertained the impact of QoE-based content adaptation on user experience.

5 Results

5.1 Assessment of the Learning Outcome

Learning outcome provided by an e-learning technology relates to the degree of knowledge accumulation by a person after studying certain material. It is analysed in form of course grades, pre/post-test scores or standardised test scores. In our experiment learning outcome was analysed in term of pre-test/post-test scores of the two groups after a study session. Pre-test scores (AHA_{mean}= 0.35, $QoEAHA_{mean}$= 0.30) showed that both groups of students had the same prior knowledge on the studied domain.

The mean scores for the post-test were 7.05 for the subjects that used QoEAHA and 6.70 for the AHA! group. A two-sample T-Test analysis on these values does not indicate a significant difference in the marks received by the two groups of subjects. (α =0.05, t = -0.79, t-critical =1.68, p(t) = 0.21).

In conclusion, results indicate no significant difference in the learning outcome of the two groups, regardless of the characteristics of the operational environment. Thus, the degradation of image quality does not adversely affect student’s learning outcome and QoEAHA offers similar learning capabilities to the AHA! system.

5.2 Impact of QoE-Based Adaptation on Learning Performance

Learning performance is another important barometer used to assess the utility and value of an e-learning system. In general the term refers to how fast a required task (e.g. studying, searching for information or memorising information displayed on the screen) is performed. The most common metric used in the educational area for measuring learning performance is *Study Session Time (SST)* [11, 12]. Apart of this, other metrics such as *Information Processing Time per page (IPT/page)* and *Number of Accesses to a Page (NoAccesses/page) performed by a person* were measured in our study. These metrics were analysed and compared for both group of subjects.

In the following, results for a 56 kbps emulated network environment are presented. Tests involving various low rate network environments in the range of 28 – 128 kbps provided similar outcomes as the 56 kbps case. These tests were for the first scenario.

5.2.1 Study Session Time (SST)

SST is measured from the moment when the subject logs into the system and proceeds to study, until s/he starts answering the post-test questions.

An analysis of the SST for the two groups, shows that on average students that used QoEAHA (SST_{Avg} = 17.77 min) performed better than the ones that used AHA! (SST_{Avg} = 21.23 min), leading to 16.27% improvement. This fact was confirmed with 99% confidence level by the T-Test analysis. The very large majority of the students that used QoEAHA (71.43 %) performed the study task in up to 20 min. with a large number of students (42.87 %) requiring between 15 min and 20 min of study time. When the AHA! system was used, only 42.85 % of the students succeeded in finishing the task in 20 min. The majority of them (71.42 %) required up to 25 min. with the largest number of students (28.57 %) in the interval 20-25 min (Fig. 1).

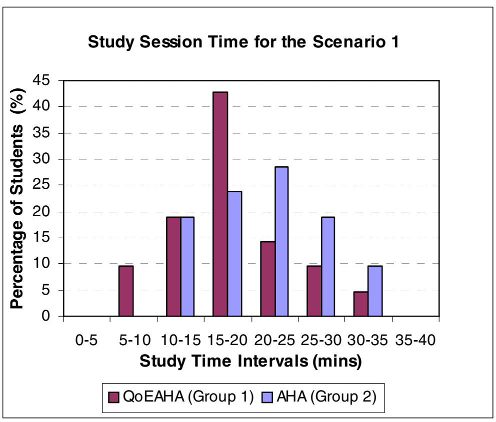

Fig. 1. Distribution of the Study Session Time for the two groups of subjects when performing Scenario 1 over a 56 kbps connection speed

5.2.2 Information Processing Time per Page (IPT/page)

IPT represents the time taken by a student to read and assimilate the information displayed on a Web page. It is measured from the moment when the Web page is delivered and displayed on the computer screen until the user sends a request for another page. The pages were not loaded in a progressive way. The results indicate that on average a lower time per page (IPT = 4.31 min) was spent by a student to process the information delivered through the QoEAHA system, in comparison to the case when AHA! was used (IPT = 4.95 min).

5.2.3 Number of Accesses to a Page (NoAccesses/page)

NoAccesses/page provides an indication of the quality of the learning process. Any re-visit to a page may indicate that the student was not able to recall the information provided in that page and thus the learning process was of poor quality. On average, subjects from the QoEAHA group performed a smaller number of re-visits ($NoAccesses/page_{Avg}$ = 1.40) than those from the AHA! group ($NoAccesses/page_{Avg}$ = 1.73). An unpaired two-tail T-Test with unequal variance assumed, confirmed with 92% confidence that there is a significant difference in the number of visits to a page measured for the two group of subjects.

How the content delivery performance of the two systems has affected *NoAccesses/page* was investigated by analysing the variability of the test-samples. The results for the *NoAccesses/page* showed that the variance for Group 1 (QoEAHA) (σ^2 =0.35) was lower than the variance for Group 2 (AHA!) (σ^2 =0.75). A F-Test analysis confirms with a 95% confidence level that Group 1 and Group 2 do not have the same variance and the difference between the two groups' variances is statistically significant. Therefore, one can conclude that Group 2 has a higher dispersion than Group 1. This indicates that a larger number of students (an average of 55 %) that used the AHA! system required more than one access per page for learning process. At the same time, the majority of students (an average of 65 %) that used the QoEAHA performed only one access per page showing that in general these students succeeded to focus better on the studied material.

Summarising the results, students from the QoEAHA group had shorter *Study Session Times* than those that used the AHA! system. A 16.27 % improvement of SST was obtained with the QoEAHA system. Since the download time per page provided by QoEAHA does not exceed the user tolerance for delay threshold [4] (12-15 sec is considered satisfactory for the Web users by the research community [13]), the students were constantly focused on required task. On average, an improvement of 26.5% [4] on the download time per page was obtained in comparison to the AHA system. Therefore, the QoE-aware system has ensured a smooth learning process. This observation is confirmed when assessing *NoAccesses/page* (on average19% decrease with QoEAHA) and *IPT/page* (on average13% decrease provided with QoEAHA).

5.3 Usability Evaluation

The goal of the usability evaluation was to measure and compare the usability, effectiveness and performance of the two systems. The methodology used for the experiment involves an online questionnaire. Questionnaires are the most widely used method in the education area with other alternatives being interviews, heuristic evaluations, subjects' observation through adequate equipment [8].

Scenario 1 required the subjects to complete an online usability questionnaire consisting of ten questions related to key usability aspects and performance issues. The answers were given on the Likert five–point scale: 1-poor, 2-fair, 3-average, 4-good, 5-excellent. The questions were categorised into: navigation, presentation, subjective feedback, accessibility and user perceived performance. The accessibility and user perceived performance questions assessed the end-user QoE. Four questions of the survey related to this and assessed user opinion about overall information delivery speed (Q6), download time in the context of browsing experience (Q7), user satisfaction in relation to the perceived performance (Q9) and whether the slow access to the content has inhibited them or not (Q5).

The results for 56 kbps connection case are briefly presented. More details on usability assessment were presented in [7]. Fig.2 shows that the QoEAHA system provided better end-user satisfaction (above the "good" level for all questions), than the AHA! system that scored just above the "average" level. This good performance was obtained in spite of the subjects using slow connection (56 kbps) during the study session and not being explicitly informed of this. Overall, the mean value of QoE usability assessment was 4.22 for QoEAHA and 3.58 for AHA!. This leads to a QoE improvement of 17.8 % for the QoEAHA system.

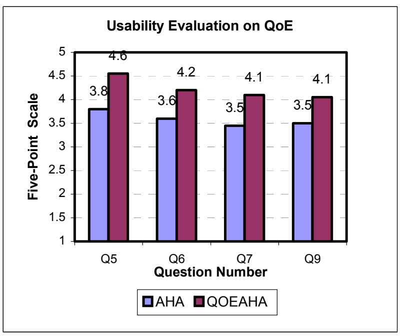

Fig. 2. Usability evaluation marks on questions that assessed the end-user QoE

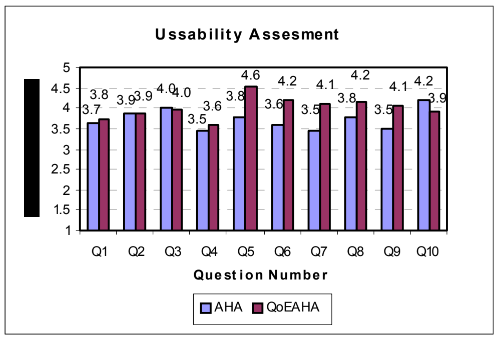

Fig. 3. Comparative presentation of the answers from the usability questionnaire

Finally, an overall assessment when all ten questions (Fig. 3) were considered of equal importance shows that the students considered the QoEAHA (mean value=4.01) significantly more usable than AHA! (mean value=3.73). These results were also confirmed by an unpaired two-tailed T-Test ($t=2.44$, $p<0.03$) with a 97 % degree of confidence. An increase of 7.5 % in the overall QoEAHA usability was achieved due to the higher scores obtained in the questions related to end-user QoE.

5.4 Visual Quality Assessment

The scenario two involved both objective and subjective visual quality assessments. The goal was to investigate the effect of the quality degradation performed in a controlled manner by the QoE Layer on learning performance and student satisfaction.

Objective visual quality assessment involved searching for two terms and answering two questions related to the target information. Time taken to find each term was measured for both QoEAHA and AHA! groups. Results [14] showed that the measured times was similar for both groups. One aspect worth mentioning is all the subjects have successfully completed the task and answered the questions correctly.

The target information was presented in the embedded images of two pages that have the biggest content size. Thus, QoEAHA imposed the highest image quality degradation for these pages. For the worst operational environment case (28 kbps) QoEAHA applied 57 % reduction for Page 1 introducing the first term and 18 % for Page 2 (second term). The subjective-based visual quality assessment investigated through a questionnaire shows that regardless of the high content reduction the average quality grade for the AHA system was 3.9, very close to "good" perceptual level, and only 4.4% lower than the average quality grade (4.3) for AHA! [9]. This suggests that the cost of image quality reduction is not significant as far as user-perceived quality is concerned while at the same time yielding significant improvements in download time and learning performance.

5.5 Correlation Analysis

5.5.1 Correlation Between Learning Outcome and Judgment on Usability

One aspect worth examining is whether or not there is any correlation between students learning outcome (post-test marks) and their judgment on system usability (questionnaire scores). Therefore, we examined if students that performed well in the post-test evaluation, thought that the e-learning system (AHA!/ QoEAHA) was more usable, while students with lower scores expressed bad opinions.

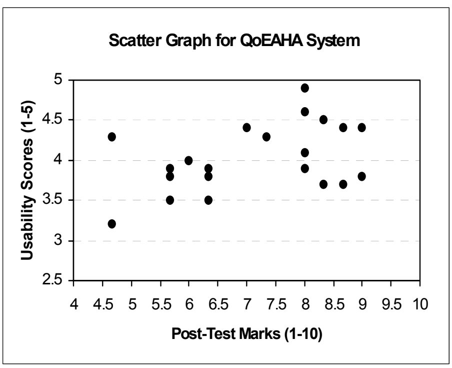

Fig. 4. Correlation analysis between learning outcome and judgment on usability for QoEAHA

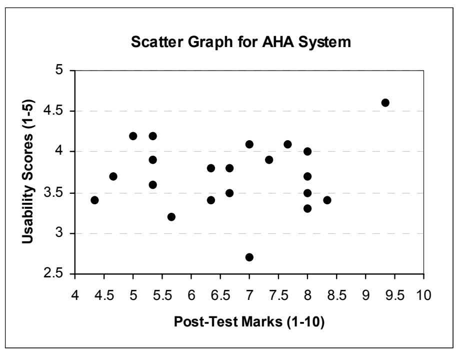

Fig. 5. Correlation analysis between learning outcome and judgment on usability for AHA!

Spearman correlation coefficient (r_s) that indicates the degree of correlation between two sets of data was measured for both systems (QoEAHA: r_s=0.21, AHA!: r_s=0.03). As both values are low we can conclude that there is no strong correlation between learning performance and perception of usability. An alternative solution for examining the correlation is to represent graphically the values in a scatter plot. Fig.4 and Fig. 5 confirm again that there is no correlation.

Summarizing, no correlation has been found between the students learning outcome and their judgment on the system usability. The student opinion expressed in the usability questionnaire was not influenced by his/her mark from the post-test evaluation.

5.5.2 Correlation Analysis Between Study Session Time and Learning Outcome

This analysis aims at investigating whether the time spent studying was correlated to marks in the Post-Test evaluation. The Spearman coefficient for both AHA and QoEAHA systems was computed (QoEAHA: r_s=0.27, AHA!: r_s=0.004). As both coefficients are low value, we can conclude that there is no strong correlation between the two data sets. In consequence a low mark received by a student was not due to a short period of time allocated for the study.

6 Reflections on the Experiment Results

The aim of the experiment presented here was to investigate the usability and effectiveness of novel QoE layer enhancement for AHS, in the educational area. The choice to use AHA! as the foundation for building a QoE-aware AHS (QoEAHA), was based on issues such as availability (open-source), trust (tested), etc. However the QoE layer could be easily embedded in any AHS that respects the principles of the AHAM [5].

The QoEAHA evaluation involved a comparison with the AHA! system in home-like low-bite rate operational environments. Different educational-based evaluation techniques such as learner outcome analysis, learning performance assessment, usability survey, visual quality assessment, and correlation analysis were used in order to assess QoEAHA in comparison to AHA!. Over 60 students from Dublin City University took part in the experiment that involved two scenarios. The first scenario covered a learning session whereas the second one involved a search for information task.

Learning outcome analysis has shown that both groups of students (AHA and QoEAHA) received similar marks on the final evaluation test. Therefore, QoE layer did not affect the learning outcome and QoEAHA offers similar learning capabilities to the classic AHA! system.

Learning performance improvements in terms of Study Session Time (16.27 % decrease), Information Procession Time per page (13% decrease) and smaller number of revisits to a page were obtained with the QoEAHA system. It is noteworthy that most of the QoEAHA group students (71.43 %) finished the study in up to 20 min. while only 42.85 % of the AHA! group students finished in the same period of time.

Results on visual quality assessment confirmed that the controlled degradation of the quality of the content performed by the QoE layer did not affect the functionality of the e-learning system. Both groups of students succeeded to complete the required task in similar period of time.

The system usability investigation performed using an online questionnaire showed that students thought the QoEAHA system provided high end-user QoE. Questions related to the QoE were marked on average over the “good” level (between 4.10 and 4.55 points) while the AHA! system scored between 3.45 and 3.8. Questions related to the other aspects of the system (e.g. navigation, presentation) achieved similar marks for both systems demonstrating that the QoE layer did not affect them. Finally an overall usability assessment shows that students considered QoEAHA significantly more usable than AHA! with QoEAHA achieving a 7.5 % increase in usability due to the high marks awarded on QoE related questions.

Acknowledgement

The support of Enterprise Ireland Commercialisation Fund is gratefully acknowledged.

References

1. Smyth, B., Cotter, P.: Personalized Adaptive Navigation for Mobile Portals. 15th European Conference on Artificial Intelligence, Lyons, France (2002)
2. Cheverst, K., Mitchell, K., Davies, N.: The Role of Adaptive Hypermedia in a Context-Aware Tourist Guide. Communications of the ACM, Vol. 45 (5), (2002) 47-51
3. Merida, M., Fabregat, R., Matzo, J. L.: SHAAD: Adaptable, Adaptive and Dynamic Hypermedia System for Content Delivery. Conference on Adaptive Hypermedia and Adaptive Web-based Systems (AH2002), Malaga, Spain (2002)
4. Muntean, C., H, McManis, J.: Fine Grained Content-based Adaptation Mechanism for Providing High End-user Quality of Experience with Adaptive Hypermedia Systems. 15th Int. World Wide Web Conf. Hypermedia and Multimedia Track, Scotland (2006)
5. AHA! Project http://aha.win.tue.nl/
6. De Bra, P., Houben, G., Wu, H., AHAM: A Dexter-based Reference Model for Adaptive Hypermedia. ACM HYPERTEXT'98 Conference, Germany (1999) 147-156
7. Muntean, H. C., McManis, J.: End-User Quality of Experience Layer for Adaptive Hypermedia Systems. 3rd Int. Conf. on Adaptive Hypermedia and Web-based Systems, Workshop on Individual Differences in Adaptive Hypermedia, Eindhoven, The Netherlands (2004)
8. Muntean, C., H., McManis, J.: A QoS-aware Adaptive Web-based System. IEEE International Conference on Communications (ICC04), France (2004)
9. QoEAHA Evaluation Forms http://www.eeng.dcu.ie/~havac/QoEAHAEvalForms/
10. Muntean, C. H., McManis, J.: Adaptive E-learning Systems: Evaluation Issues. Transactions on Automatic Control and Computer Science Journal, Vol. 49 (63) 2004, 193-198
11. Weber, G., Brusilovsky, P.: ELM-ART: An Adaptive Versatile System for Web based Instruction. Artificial Intelligence in Education, Special Issue on Adaptive and Intelligent Web-based Educational Systems, Vol. 12, No. 4 (2001) 351-384
12. Ng, M. H., Hall, W., Maier, P., Armstrong, R.: The Application and Evaluation of Adaptive Hypermedia Techniques in Web-Based Medical Education. Association for Learning Technology Journal Vol. 10, No. 3 (2002) 19-40
13. Sevcik, P. J.: Understanding How Users View Application Performance. Business Communications Review, Vol. 32, No. 7 (2002) 8-9
14. Muntean, H. C., McManis, J.: QoSAHA: A Performance Oriented Learning System. AACE ED-MEDIA2004 , Lugano, Switzerland (2004)