SPERO - A PERSONALIZED INTEGRATED E-LEARNING SYSTEM

Tzouveli P., Mylonas P. and Kollias S.
National Technical University of Athens
9, Iroon Polytechneioy, Zographoy, Athens, Greece
{tpar,fmylonas}@image.ntua.gr, stefanos@cs.ntua.gr

ABSTRACT

One of the most important topics in modern Internet-based multimedia e-learning systems oriented towards teachers, students and learners in general, is the treatment of information in a personalized, easy to use and friendly way. In this framework, the automated extraction of learner profiles from such an e-learning system, in order to offer them personalized educational A/V content and e-courses, is a challenging and important problem.
In this paper we present the design and development of an integrated distance learning system. Our approach relies on a suitable adaptation of fundamental e-learning models, such as the IEEE e-learning model, as well as a novel mechanism that creates, updates and uses learners' profiles, extracted directly from user preferences and their usage history. Such preferences are utilized towards the personalization of the multimedia educational content offering and retrieval process.

KEYWORDS

e-learning, distance education, personalized educational multimedia environment, teaching application

1. INTRODUCTION

It is a common fact, that the main goal of current multicultural educational society is to provide knowledge life long, in order to equip students suitably for their future work, via a "school for all" principle. The impact of Information and Communication Technologies (ICT) in such a task has become more and more evident in learning and teaching at all levels of education. E-learning is undoubtfully a revolutionary new way to empower a workforce with the skills and knowledge it needs, providing end-users with the visual and multimedia tools needed.

In this paper, we present the methods followed in the framework of SPERO project [1] for gathering information and estimating the ICT level of teachers and students in all fields of education through a web-based user-friendly interface, as well as a personalized, profile-based scheme to fulfill this purpose. The latter has been designed to enable teachers and students to gracefully increase their ICT knowledge [6] and provide them with credible information and feedback. Additionally, appropriately selected e-courses are provided to the end-users through the easy to use web interface of the system, according to the above dynamically updated profiling information.

SPERO is closely related to e-learning, the so called "Internet-enabled learning". E-learning provides faster learning at reduced costs, increased access to learning information and clear accountability for all participants in the learning process, so, undoubtfully, forms the future approach to the learning process, by eliminating barriers of time, distance and socioeconomic status through simple visual interfaces. Several components of this system include content delivery in multiple audiovisual formats, management of the learning experience in several ways, plus a networked community of learners, teachers, content developers and domain experts. The various media used within the SPERO project for delivering e-learning material include broadcast video, content on demand and virtual classrooms.

The structure of the paper is as follows: In section 2, the overall architecture design of the SPERO system is presented. In section 3, after a short reference to the IEEE e-learning model, we present the system's adaptation procedure and explain the extra features provided by it. In section 4, we consider mostly the e-learning point of view of the system. Furthermore, in section 5, we tackle initially the problem of the learner profile creation, followed by issues concerning the profile extraction, re-evaluation and re-estimation procedures. All of the above are used as the main feedback source for the forthcoming e-course selection and

evaluation process, which is presented in section 6. Finally, in section 7, we present our concluding remarks and future work.

2. ARCHITECTURE DESIGN OF SPERO SYSTEM

The SPERO system forms an integrated e-learning system which is adapted to the learner. The general architecture design is shown in Figure 1. This system consists of three entities, for which an analytical description will follow.

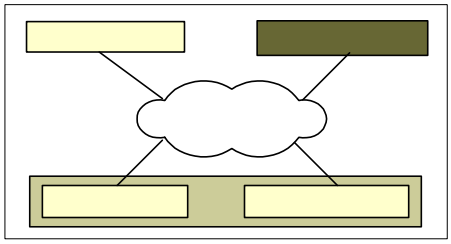

Figure 1: SPERO System Architecture

The first entity is the User or the Group of Users which is comprised of the proposed system's learners. This group supports either single users learning individually, or users learning collaboratively. In addition, different roles of users may exist, like parents, teachers or employers. The Group of Experts is another entity that contains a variety of people like teachers (in general or special education field), experts in e-learning, computer engineers, statistic analysts and psychologists. The group's main goals are the design, illustration and continuous improvement of the SPERO system. The Server System forms the hardware and software of the system that is based on the well-known three-tier architecture, whereas each tier can be implemented on the same machine or on different ones. The three-tier architecture is considered as the most suitable architecture for such large, web-based enterprise applications [8].

The partitioning of the application enables rapid design and development of the system. The modularity makes it easier to make changes to just one tier without affecting the others. Separating the functions into distinct tiers makes it easier to monitor and optimize the performance of each layer. The three tiers reflect the specialized functions of a web-based enterprise application. More specifically:

- The presentation layer delivers the SPERO application to the end users on the Web.
- The business logic layer contains and executes the rules that run the SPERO application.
- The database layer manages the data required by the application.

3. SYSTEM ADAPTATION

The architecture of the SPERO application system is based on the IEEE Reference Model (WG) of the Learning Technology Standards Committee [5]. In brief, this standard contains three different categories as Processes (learner entity, evaluation, coach, and delivery), Stores (learner records, learning resources) and Flows. The Flows are vectors like learning preferences, behaviour, assessment information, performance and preference information, query, catalog info, locator, learning content, multimedia and interaction context. The Processes and Stores entities exchange information and data via the Flows which connect these entities.

Our approach to the e-learning system, which is outlined in the LTSA Draft Standard [5], consists in the personalization of the learner. In this work we attempt to extract learner profiles through the evaluation entity of the IEEE learning system. For this propose, the group of SPERO experts have designed a multilingual e-questionnaire, from which the learner profile can be automatically extracted. In addition, a survey can be carried out according to the answers gathered.

In Figure 2 we present the proposed extension of LTSA IEEE system. Our system contains additionally three new components (Questionnaire's Texts, Learners' Profiles and E-survey) which attempt to lead the learner towards a profile adaptable e-learning system and provide a statistic analysis of the learners' responses.

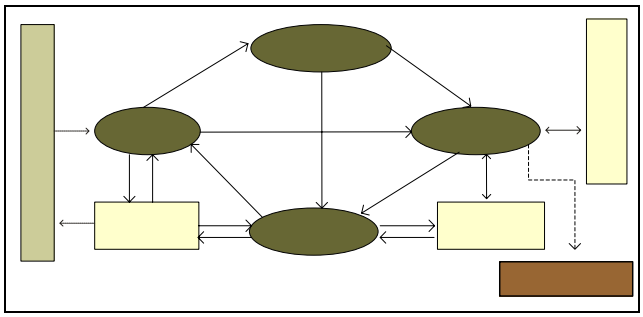

Figure 2: LTSA IEEE learning system components

The Learner process is thought to be the input to the SPERO application system and forms the main entity around which the system is organized. It represents a single learner or a group of learners with different needs. The Delivery, Coach and Evaluation processes belong to the business logic layer of the system architecture, whereas the Learner Resources, Records and Profiles constitute the database layer. The "Questionnaire's Text" is a new component of the Delivery process that transfers the texts of the multilingual questionnaire from the "Learning Resources" to the "Delivery" process. The "Delivery" process receives these texts and then creates and dispatches the multilingual questionnaire presentation to the learner when he/she accesses the SPERO system for the first time.

The role of the “E-survey” component, added to the “Evaluation” process is the presentation of results from statistical analysis [3], which has been conducted on the already stored "Learner Records". The "Delivery" process is responsible for automatically generation of the e-questionnaire. In our approach, the Learner initially receives a questionnaire, defined by the Group of Experts, for which he/she is called to answer. The learning materials which are stored in the "Learner Resources" will be presented to the Learner automatically after his/her profile generation that the evaluation process produces. Consequently, the learner receives a multimedia presentation from the "Delivery" process, which is retrieved from the store "Learning Resources" after having been evaluated. The learner entity's observable behavior is given as input to the "Evaluation" process. The evaluation component produces assessment information which is sent to the "Coach" and creates performance information that is stored in the learner records. Additionally, the learners' records containing the learners' answers to e-questionnaires constitute the data with which an e-survey can be conducted.

The learner profiles, initially defined by the experts, are stored in the profile database. A new learner profile can be created or existing ones can be adapted based on statistics extracted from the e-questionnaire database. New learner profiles or adapted versions of them are stored in the Learner Profiles Database. Whenever an already registered learner accesses the systems its profile is restored from the Learner Profiles Database. Change of the learner's profile is mainly performed based on the progress of its training.

4. SPERO: AN E-LEARNING MULTIMEDIA SYSTEM

The term e-Learning is used to describe a wide range of efforts to provide educational material on the web. These efforts include a diversity of approaches, ranging from static HTML pages with multimedia material to sophisticated interactive educational applications accessible on-line. The SPERO application can be characterized as a flexible interactive educational system being able to adapt to its users preferences. The proposed system has been developed for teachers that work to the main stream or to the special education section in a wide area of the European Union, as well as for their students. The latter forms one of it’s main novelties, as it is designed, implemented and introduces several automated procedures and techniques, in order to fulfil the needs and satisfy the interests of people involved in special education.

This system can adapt to the learners' preferences and languages and can provide educational material for Information and Computer Technologies (ICT). Learners could either be teachers or students. Both of them are in great need of ICT. On the one hand the teachers mainly because their role is continuously evolving and demanding new formation and students because of their need to have distance e-courses in the field of ICT.

The main entity, around which the whole system is structured, is the e-questionnaire [2]. This questionnaire contains three sub-questionnaires. A small part of one of these sub-questionnaires is depicted in Figure 3, while the rest of it is omitted for the sake of space. This subquestionnaire is intended to collect information about general teachers' educational background, as well as their background in ICT.

In addition, information concerning teacher's opinions about pedagogical utilization of ICT and the amount of using ICT in teaching procedure is also extracted by the following sub-questionnaire towards the aim of carrying out a European survey. Another sub-questionnaire, that focuses on the students aims to investigate

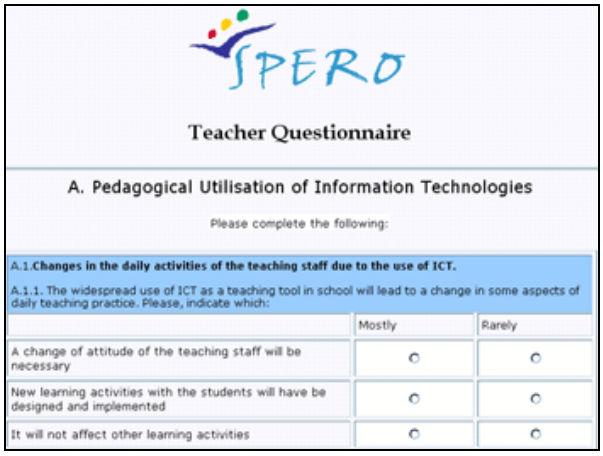

Figure 3: Part of Teachers' Questionnaire

their ICT background, as well their educational interests and needs towards the purpose of selecting from the learning material the appropriate courses for them.

Finally, the evaluation process of the SPERO system extracts statistical analysis results from the learner's responses (Figure 4), which provide useful information about the teachers' opinions regarding new technologies. Moreover, the students' opinions are evaluated, in order to specify their needs and interests.

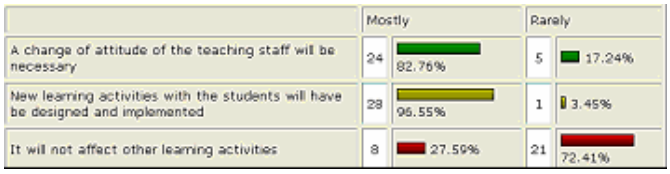

Figure 4: Statistics from the learners' responses

The automatic profile evaluation which takes place as the last step in the proposed system provides extremely useful and fully personalized information regarding SPERO learners and will be analyzed in the next section. Learning resources have been also linked to each profile category that has been defined during the profile evaluation and the group of SPERO experts is responsible to provide a set of e-courses to the group of learners, according to the latter information.

In Figure 5 we show the learners' profiles [4] that the SPERO system has extracted according to learners' responses to the e-questionnaires.

	User ID	Professional Development	Personal ICT Background	Teaching use of ICT
1.	509	beginner	unspecified	beginner
2.	659	beginner	unspecified	beginner
3.	807	expert	expert	beginner
4.	808	unspecified	beginner	beginner
5.	809	expert	expert	advanced
6.	811	expert	expert	advanced
7.	813	expert	expert	advanced
8.	814	expert	expert	advanced

Figure 5: Learners' Profiles

5. LEARNER'S PROFILE CREATION & ADAPTATION

This section of the paper refers to the so called "content personalization" module of the system via the underlying automatically generated creation of SPERO's user profiles. This step is based on the information provided by the users' input data, which are collected from the e-questionnaires. It is a common fact, that several types of personalization exist when dealing with interactive services, applications and content delivery to the end-user [9], where different content can be generated for different individual users or even classes of users. Different approaches exist also in the way the system will deal with the gathering of information and statistical analysis. Unquestionably though, it becomes more and more clearer, that results obtained from any kind of electronic questionnaires, like the ones introduced herein, do not have the desired effectiveness on the imminent e-course selection and evaluation, unless major focus is given to the individual user profiling and the corresponding user behavior. Such functionalities define users' capabilities, as well as their potential future courses choices.

Automated user profile extraction from several sources of knowledge is, in general, a complex and difficult task. It is, however, expected that it could be tackled when dealing with specific application

domains, so herein we focus on the electronic mining of knowledge through the SPERO intelligent system. As already mentioned, in this system, knowledge is collected and stored initially in form of e-questionnaires. Educational material presentation and offering in the form of e-courses and A/V content is at this point based on this type of knowledge. The main target of the mining procedure is the extraction of information from data provided by users as input. It takes into account:

- The nature of useful input data that users may provide through the e-questionnaires.
- The particular context of the input data, when expanded towards content related parts of a specific e-questionnaire.

In our implementation, we present a novel mechanism that creates, updates and finally uses users' profiles during the oncoming e-course selection. The profile model developed is designed in a way that facilitates both the process of using user preferences in profile creation, as well as the process of tracking of these preferences throughout the whole re-evaluation procedure. The overall abstract structure of a user profile is presented in Figure 6.

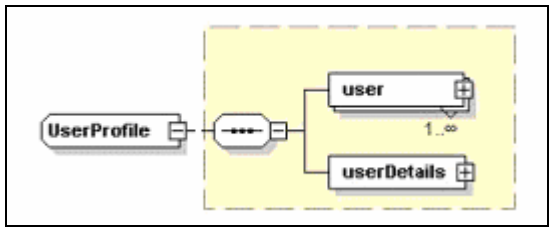

Figure 6. Overall structure of a “UserProfile”

As seen in this figure, the "UserProfile" compound type contains two elements. The first element "user" stores information about the usage history, while the second, "userDetails", stores the user preferences. The sequence of "user" elements states the existence of different SPERO users, filling corresponding e-questionnaires, whereas each one of those corresponds to a different SPERO user. So, one "UserProfile" type contains the user's profile as well as the user's details. The former holds also all the information that is required for personalizing the interaction of SPERO with the user, while the later contains additional authorization-related information, such as the user's level of expertise. Various levels of access rights are defined for each level of expertise, thus classifying users to several classes. These classes describe the actual data (e-courses, A/V content) that each user may access and are represented within the "userDetails" element of the profile.

Furthermore, as the profile extraction process instantiates, all user profiles are stored within a single mapping structure. This step is necessary, in order to ensure proper and balanced classification of the above profile types, according to the underlying information and results into an overall descriptive and representative profiling mechanism. The abstract model of this structure is presented in Figure 7.

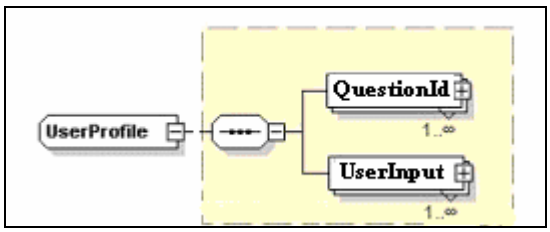

Figure 7. Overall mapping structure of “UserProfile”.

As seen in the figure, the "UserProfile" is mapped against information retrieved either from the structure of e-questionnaire itself, or directly from the input of the users. The first element, "QuestionId" holds all the information required for identifying the underlying e-questionnaire question, as well as its type, aiming at better understanding and fitting of the currently generated profile, while the second element, "UserInput" contains user data related information, such as the user's actual answers in integer or string format. Both, the sequences of "QuestionID" and "UserInput" elements denote the existence of large amount of different SPERO e-questionnaires, questions and users' input data.

Moreover, the main hot points to consider, regarding the specific user attitudes, may be summarized in the following:

- A user may be interested in multiple topics.
- Not all topics that are related to a user's usage history are necessarily of interest to the user.

All of the above issues are tackled using higher level tools and principles. In more detail, our approach performs a fuzzy hierarchical clustering of the user's input, relying on the knowledge stored in the form of e-questionnaires. The essence of clustering the data is to identify homogeneous groups of objects based on the values of their attributes. It is a well-known approach that is related to various scientific and applied fields and has been used in science and in the field of data mining for a long time [11]. In particular, the notion of context has a central role in this process and forms the basis on which the extraction of preferences is built. The common topics of offered materials are used in order to determine which of them are of interest to the user and which exist in the usage history coincidentally. In the case of two documents (multimedia or plain ones), the determination of what is common among them, i.e. their common topics, can be referred to as their common context. The height of their common context is used as a metric that can indicate the degree to which two documents are related and subsequently this can be extended to the case of more than two documents, in order to provide a metric that measures the similarity between several documents. The representation and handling of preferences using fuzzy sets is developed to a greater extent in [10], whereas our novel clustering algorithm is described in details in [12].

Obviously, since a user may have multiple interests, we should not expect all documents of the usage history to be related to the same topics. Quite the contrary, we should expect most documents to be related to just one of the user's preferences. Therefore, in addition to the restriction of allowing the algorithm to be fuzzy, the cardinality of the clusters must be used, as well as a ranking of documents based on their common topics needs to be applied. In this process, documents that are misleading (e.g. documents that the user chose to view once, just to find out that they do not contain anything of interest to him) will probably not be found similar with other documents in the usage history, resulting to only those documents that form the desirable distinct user profile categorization. In order to overcome this obstacle, weights are introduced in order to balance the obtained results from the above user profiling procedure. These weights are initially predefined and stereotyped within the SPERO system specifications from a suitable group of system's experts, categorizing its users into three main categories: headstrong users, intermediate users and unconcerned users. As more and more input from every specific end user is gathered and as more answers from the end user enter the "UserProfile" structure, the above procedure results in a weighted mapping of this end user to a specified profile. This mapping changes in that manner continuously and dynamically every user's profiling, until a final equilibrium profile state is achieved. The combination of the profiling information with those weights, results into an ultimate descriptive and representative profile for each user. This final profile is then used in order to provide him/her with educational content of their interest, as well as with metadata and audiovisual enhancements, such as statistical information and explanatory graphs during transmission of the mainstream content (e.g. during e-courses and educational material browsing).

The ultimate output of this process, after the application of the weights, is the extraction of a "1-1" profile-end user relation for the part of the e-questionnaire taken into consideration. As the final step, all allocated e-questionnaire parts are then taken into consideration, in order to extract one final user profile. Again, suitably selected weights, specific for each part of the e-questionnaire, are used. In that manner, each old, as well as new SPERO end user is classified to a specific profile, which characterizes her/his behaviour, her/his interests, as well as her/his further e-course offerings performed by the underlying SPERO system.

Information obtained from profile extraction is transferred to the learner, as well as used for internal re-evaluation of the above process. In that manner, the process delivering content to the end SPERO user is combined with this evaluation process to achieve the tight coupling necessary for responsive, interactive learning experiences. The core technological target of such a combined process is to blend the achievements in characterizing users with innovative intelligent technologies in order to perform constant adaptation of user profiles to the new input data.

According to user input, all aforementioned profile mappings to end users are re-evaluated and reassigned. In order to tackle such a task, statistical information is gathered and taken into consideration, regarding each question alone, as well as a bunch of questions as a whole, according to the context defined by the profile of individual users or specific user groups; the latter clearly appreciating that such interpretation relies on the context which in turn depends on the specific profile. The SPERO system routines that perform statistical queries and extract relevant statistical information run in parallel to the profiling process and provide useful data from the set of e-questionnaires, as well as from the final content delivered,

regarding tendencies and percentages amongst them. For the sake of completeness, these statistics are also presented to the end users of SPERO, providing first-hand knowledge and feedback.

6. SELECTING APPROPRIATE E-COURSES

The SPERO software is a web-based learning portal, designed and implemented according to well-known learner-friendly solutions and flexible e-learning software application [13]. When teachers or students visit the SPERO site, they are validated against the SPERO learner database and subsequently are called to answer the e-questionnaires in order to automatically establish their user profile. This profile is stored in the profile database as soon as they answer the whole questionnaire. The main menu of the SPERO portal contains links to the following sectors/services:

- Courses Catalogue: It contains the titles, as well as a small textual description of one or more e-learning courses, that learners may take. An intelligent module takes over the selection of e-courses, according to user preferences and profiles, as well as their usage history. A small overview for each e-course is provided, demonstrating its main topics and concepts. A small notion of a selected e-course listing is presented in Figure 8.

Figure 8. Personalized e-course listing sample.

- Announcement Service: This service provides a bulletin board where topics about e-courses or other educational subjects are published. Relative documents, regarding e-courses outlines and requirements are posted herein. Students' and teachers' messages are presented in a threaded view layout.
- Search Service: It provides a search environment to facilitate information and educational materials retrievals from SPERO site, e.g. members, school units, e-lessons, e-books, e-lectures, exercises, "live" educational content broadcasts, etc.
- E-mail Service: SPERO users are able to send and receive e-mails through the SPERO system.
- Upload Files: Learners have their own personal space where they can store their own material to which other learners may or may not have access to. Several levels of authorization access are implemented.
- Help Service: Analytical description of the usage and tasks of SPERO menu choices. It provides information about library links and online resources outside the SPERO system and answers general Frequently Asked Questions.

In order to improve the ICT level of learners, different e-courses are delivered to them according to their ICT level. Indicatively, group of e-courses with increasing difficulty are characteristic by the following topics:

- Group 1: Introduction to Information's Technologies (definition of data, bit, byte, presentation of hardware components, presentation of type of software)
- Group 2: Presentation and usage of operating systems.
- Group 3: Usage of text editor, software for work sheets, software for creation a presentation, educational software
- Group 4: Usage of Internet (explorer in a browser, search machines, sending and receiving e-mails, access to a news group, access to a chat room.

Selected e-courses should correspond to the available SPERO user profiles. So, students who receive an e-course are monitored, regarding their selections, throughout the SPERO system and their behavior is observed. In order to estimate whether the suggested e-course was one that covers their level of knowledge as well as the requirements of the user, some kind of selection's evaluation is required and therefore performed. For the evaluation of e-course selection, the behavior of the user when attending the course is recorded, framed in the appropriate context. This behavior information is extracted and stored in the learner's profile, providing the main source of transparent feedback for the users' future e-courses selections.

7. CONCLUSION

This paper is part of our ongoing work in the field of designing and developing a fully personalized integrated e-learning system, implemented to be used by teachers and students originating from all levels of education. It combines information gathered from appropriately extracted user profiles, preferences and usage history. It extended on previous work [7] on profile extraction and re-evaluation in the direction of automated extraction of user preferences. Such preferences can then be utilized towards the personalization of systems' content deliver process.

Based on an adaptation of the IEEE fundamental e-learning model, SPERO provides faster learning at reduced costs, increased access to learning information and clear accountability for all participants in the learning process, thus indicates an efficient approach to the learning process via simple to use visual interfaces.

A major area of future research for this work is the utilization of a fuzzy relational knowledge representation model in the learners' profile weight estimation process. Our findings so far indicate, that such a combination between semantic and statistical information is possible and will have very interesting results, regarding the personalization of the educational content offered to the end-users.

ACKNOWLEDGEMENT

This work has been partially funded by the "SPERO: TeleInformatics System for Continuous Collection, Processing, Diffusion of Material for Teacher Training in Special Education" European project through the "Leonardo Da Vinci" programme, under "Themes 5: E-Learning".

REFERENCES

[1] SPERO: TeleInformatics System for Continuous Collection, Processing, Diffusion of Material for Teacher Training in Special Education (http://www.image.ntua.gr/spero)

[2] SPERO e-questionnaires: (http://www.image.ntua.gr/questspero)

[3] SPERO statistics: (http://www.image.ntua.gr/statsspero)

[4] SPERO profiles: (http://www.image.ntua.gr/statspero/profiles)

[5] P1484.1 Architecture and Reference Model, IEEE LTSC (http://ltsc.ieee.org/index.html).

[6] Tzouveli, P. et al, 2003, On the Development of User Adapted E-learning Schemes, *Proceedings of NHIE 03*, Santorini, Greece.

[7] Wallace, M. et al, 2004, Automatic extraction of semantic preferences from multimedia documents, *Proceedings of WIAMIS*, Lisboa, Portugal.

[8] Skordalakis, E., 1991, *Introduction to Software Engineering*, Athens.

[9] Correia, N. et al, 2002, Towards an Integrated Personalization Framework: A Taxonomy and Work Proposals, *Proceedings of WPTEPW*, Malaga, Spain.

[10] Wallace, M. et al, 2002, Representation of user preferences and adaptation to context in multimedia content -- based retrieval, *Proceedings of SOFSEM*, Milovy, Czech Republic.

[11] Hirota, K. et al, 1999, Fuzzy computing for data mining, *Proceedings of the IEEE 87:1575–1600*

[12] Mylonas, P. et al, 2004, Using k-nearest neighbor and feature selection as an improvement to hierarchical clustering, *Proceedings of 3rd Hellenic Conference on Artificial Intelligence*, Samos, Greece.

[13] WebCT Campus Edition: Course Management System (http://www.webct.com/)