Intelligent content adaptation in the framework of an integrated e-learning system

Phivos Mylonas
National Technical University of Athens
School of Electrical and Computer Engineering
Department of Computer Science
Image, Video and Multimedia Laboratory,
Tel.: +30 210 7724351
fmylonas@image.ntua.gr

Paraskevi Tzouveli
National Technical University of Athens
School of Electrical and Computer Engineering
Department of Computer Science
Image, Video and Multimedia Laboratory,
Tel.: +30 210 7724352
tpar@image.ntua.gr

Stefanos Kollias
National Technical University of Athens
School of Electrical and Computer Engineering
Department of Computer Science
Image, Video and Multimedia Laboratory,
Tel.: +30 210 7722488
stefanos@cs.ntua.gr

ABSTRACT

It is a common fact that modern e-learning schemes lack educational content representation and user personalization. In this framework, automated extraction of user profiles, to be used in an e-learning content offering system, forms an interesting and important problem. In this paper we present the design and implementation of such a profile-based system, by which content is matched to its environmental context, so that it can be adapted to its user's needs and capabilities. The paper extends previous work on profile extraction and clustering techniques and on integrated e-learning systems. Our approach relies on the fundamental IEEE e-learning model, suitably adapted to reflect the profiling aspects of the system.

Keywords

web-based education, e-learning, personalized education systems, clustering-based profiling methods

1. INTRODUCTION

The nature and structure of our era, dominated by rapid information exchange and instant worldwide communication capabilities, has significant impact on education and the way it evolves. In an epoch where everyone and everything continuously changes, education itself couldn't stay passive and unconcerned: all traditional teaching techniques are revisited and reevaluated and new or sometimes radical ones are introduced. Above all, Internet comes in the foreground, playing a significant role in all fields of education, contributing the most to the educational procedures. As a result, Internet-oriented applications arise in the aid of educational needs, trying to close the gap between traditional educational techniques and the new trends of future, technology-blended education.

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 [5]. E-learning is unquestionably the revolutionary new way to empower a workforce with the skills and knowledge it needs. Towards that goal, during the last years, e-learning systems were developed in the means of static software applications, lacking in educational multimedia environments and personalized capabilities and without any interest given to the real users input and feedback.

In this work which was conducted in the framework of the Leonardo SPERO project [9], we present a novel method for gathering information and estimating the ICT level of learners in all fields of education through a web-based user-friendly interface that takes into consideration personalized, profile-based schemes. The latter has been designed to enable learners to gracefully increase their ICT knowledge and provide them with credible information and feedback, such as suitably selected e-courses and multimedia educational content.

The structure of this paper is as follows: In section 2, the overall architecture design of the SPERO system is presented; its basic corresponding groups and components and its notion of e-questionnaires. Afterwards, a short reference to the IEEE e-learning model is presented and current approach's adaptation is analyzed and the extra features provided by it are explained. In section 3, we tackle initially the problem of the learner profile creation, followed by issues concerning the initial static profile extraction procedures. All of the above are used as the main feedback source for the forthcoming intelligent clustering profiling procedure, which is presented in section 4. In the same section is provided a description of the utilized clustering algorithm, together with experimental results on the clustering-based profiling scheme. Next section 5 describes the general context for this work, dealing with the educational content offering of the system in general and briefly presenting its categorized e-courses. Finally, in section 6, we present our concluding remarks and future work.

Hypertext'05, September 6-9, 2005, Salzburg, Austria.

2. SYSTEM'S ARCHITECTURE

In the effort of designing, implementing and evaluating a novel, integrated e-learning system, the first step to consider is the definition of its basic architecture. The general architecture design of SPERO is shown in Figure 1:

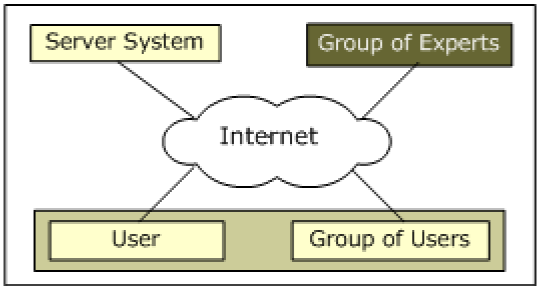

Figure 1. SPERO System Architecture

Three main networked components can be identified which are: a) the group of the *system's users*, b) the group of *system's experts*, who play a key role in the personalization process; this group includes teachers, experts in e-learning, data analysts, psychologists and software engineers and c) the actual *server system*, which includes all hardware and software needed to establish a 2-tier system core [4]. The first group includes every kind of teacher working either for general education or for the Special Education sector. Eventually, expanding the system's architecture, a user could be identified as any learner, whether a student, teacher or employee. The second group's role is crucial in the personalization process of the SPERO system, since it defines the initial set of specifications and limitations of the end-users' profiles, which justifies the variety of people consisting it. The third group includes all hardware and software that enables a web-server to be active, as well as efficient and robust. All SPERO web applications [11][12][13] together with an underlying RDBMS system to support profiling and user information are included in this setup.

SPERO experts have designed and illustrated two groups of e-questionnaires: The first group of e-questionnaires contains questions about school units in order to collect general details about them. The second group contains questions about teachers' Information and Computer Technologies (ICT) background. The questions, which are addressed to the teachers, are intended to collect information about teachers' educational background, as well as their background in the Information and Computer Technologies. In addition, information concerning teachers' opinions about pedagogical utilisation of ICT and the amount of using ICT in teaching procedure is also extracted. 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 group of e-questionnaires is accessible through the SPERO web server [11]. Each questionnaire is divided into several subsections, a portion of which is depicted in Figure 2:

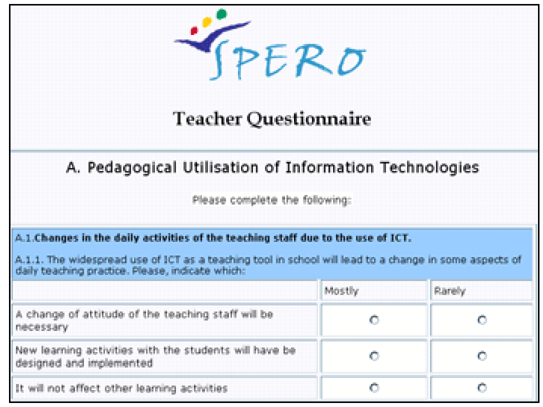

Figure 2. Part of Teachers' Questionnaire

This subquestionnaire is intended to collect information about general teachers' educational background, as well as their background in ICT. SPERO e-questionnaires are developed in the framework of conducting a European survey. Consequently, they are translated in eight European languages and these translations are stored within the SPERO database. Software has been developed for automatic presentation of the e-questionnaires in every one of these eight languages. Moreover, the e-questionnaire is used to estimate the ICT level of individual users, by using the calculated user's profile categorizations, that are automatically extracted in the following by the SPERO software. More than one learning resources (e-courses and educational material) are selected by experts, up to one for each of the distinct collaborative user profiles categories. The set of e-questionnaires is used for ICT level estimation in the framework of the distance-learning architecture that is presented in Figure 3:

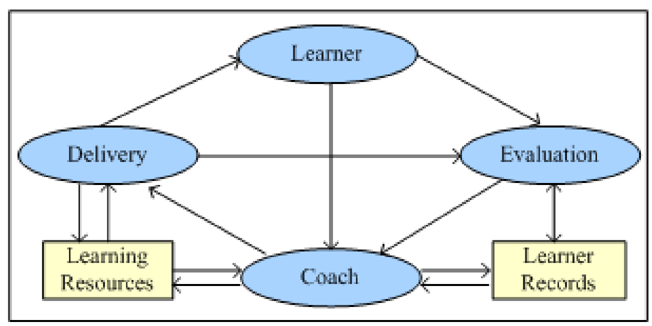

Figure 3. IEEE learning system components

In this work we attempt to extract learner profiles through the evaluation entity of the above architecture, proposed by the IEEE Reference Model (WG) of the Learning Technology Standards Committee [8]. However, in this generic approach to e-learning systems, a system's ability to adapt its operation to the user is not defined; although an evaluation process exists. Aiming at extracting learner profiles through this entity, we are proposing the replacement of the IEEE standard *Evaluation entity*, with the novel *Re-evaluation entity*, which, additionally, is strongly related to two new entities: the *E-survey entity* for gathering statistical information and the *Profile Database entity* dealing with all learners' profiles. A schematic diagram of this procedure is presented in Figure 4, whereas an extremely detailed description of the above concepts can be found in [2].

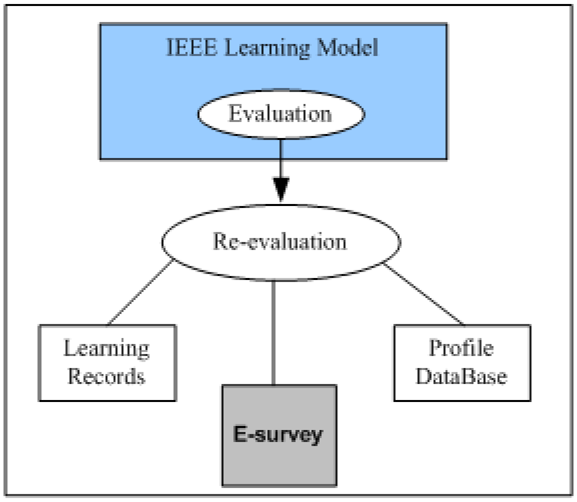

Figure 4. Proposed replacement of evaluation entity

In order to assist the profiling process, the need for this evaluation step is essential; the usage of an appropriate e-questionnaire was considered necessary in order to collect user input data and base profiling information on top of them. For this purpose, the system's experts designed and illustrated such a e-questionnaire [11] which collects information about learners' ICT background, learners' opinions about pedagogical utilization of ICT and the amount of using ICT in teaching procedure. Additionally, software has been developed to allow e-surveys to be conducted based on users' answers. As a result, the core of the system relies on this replaced evaluation entity. The latter forms an independent personalization subsystem, where user profiling information is extracted, according to statistics gathered from the e-questionnaire database and the e-survey. Delivery of educational content is then possible, based on the results of the profiling procedure, providing personalized views to the system's end-users and taking into consideration their particular ICT levels of education and needs.

3. PROFILING INITIALIZATION

At this point, a brief presentation of the system's personalization subsystem is indispensable. The initial profile representation approach of the system consists of a static profiling mechanism. Experts, based on experience and intuition, define a set of three user characterizations, forming a static profiling representation. These initial characterizations are also utilized at a later stage, during the intelligent profiling process, providing a rock-solid point of reference and although they are thought to be static, they are actually generated automatically from the system. Their basis is information provided by the users' input data, obtained from the e-questionnaires [9]. In this case, personalization was needed in order to aid with the ultimate educational content offered by the system; this was successful, based on the electronic mining of knowledge gathered from the system's questionnaires subsystem.

The profiling mechanism creates updates and uses system's user profiles, mapping specific e-questionnaires question triggers, to particular identified patterns. The profile model's design facilitates both the process of using user preferences in profile creation, as well as the process of preference tracking throughout the whole profiling procedure. Furthermore, it is designed in a way that allows for the automated extraction of user profiles, based on these preferences and the users' input history. This model forms an initial static version of the user profile denoted by "UserProfile" and presented in Figure 5:

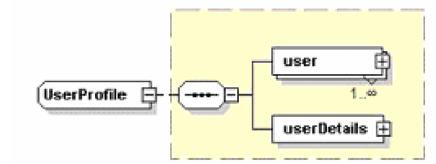

Figure 5. Structure of a "UserProfile"

As seen in the figure, the main abstract structure of "UserProfile" compound type contains two elements. The first, "user", stores information about the user's history, while the second, "userDetails", stores the user preferences. As the initial profiling process instantiates, all user profiles are stored within a single, central mapping structure, whose abstract model is presented in Figure 6. The "UserProfile" is mapped against information retrieved either from the 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. The second element, "UserInput" contains user data related information, such as the user's answers. Both, the sequences of "QuestionId" and "UserInput" elements denote the existence of large amount of different system's e-questionnaires, questions and users' input data.

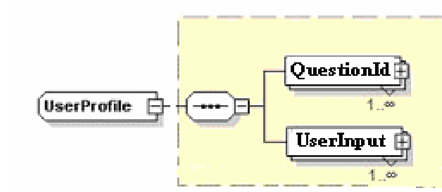

Figure 6. Mapping structure of "UserProfiles"**.**

The core of this methodology is summarized in the following step of weights association, performed according to the following guidelines: Once a user answers a question of the input e-questionnaire, a relevance degree is associated to it and adjusted to her/his specific "QuestionId" element, and thus also propagated to the "UserProfile" element. As more and more answers from the end user enter the "UserProfile" structure, additional relevance degrees are registered to the corresponding "QuestionID" elements. Depending on the particular question, as well as the part of the e-questionnaire that this question belongs, different degrees are propagated. The latter is based on comparison of the provided numerical values with the range of values a-priori associated with the profiles. In order to better understand the underlying mapping structures, examples are presented in Figure 7, Figure 8 and **Figure 9**, derived from the group of experts directly from the e-questionnaires:

B.1.4.Teaching Experience	
Under 5 years	Beginner
5 to 10 years	Advanced
11 to 20 years	Advanced / Expert
Over 20 years	Expert

Figure 7. Static profile mapping example
(1st part of e-questionnaire)

.

B.2.5.Which of the following tasks have you performed at least once ?	
Installation of software	a
Installation of a printer	b
Greation of backup	b
Installation of applications (e.g. MS Office)	b
Problem solving related to software	e
Problem solving related to the printer or Internet card	f
None of the above	g
Other, please indicate:	text

Checked (a…f)>4 => "Expert"
2<Checked (a…f) <4 => "Advanced"
2 < Checked (a…f) OR Checked (g) => "Beginner"

Figure 8. Static profile mapping example
(2nd part of e-questionnaire)

B.3.4.How many hours a week, on average, do you use the Internet or educational software with your students?	
General	if (General or/ and SEN teaching hours)>5 "=> Expert"
SEN	else if 2 < t. h.) <5=>"Advanced" else if (t. h.)<2=>"Beginner"

Figure 9. Static profile mapping example
(3rd part of e-questionnaire)

The e-questionnaire acts as an intermediate towards the information gathering process, and as the mount of the answered questions increases, the more entries are summed up in the "UserProfile" structure. Thus, the overall process results in an aggregated weighted mapping of the end user to the specified profile, which is different for each user's answers and depends on their particular nature. This mapping is temporarily preserved and as the completion of the e-questionnaire is carried out, the above mentioned weighted mappings are aggregated. In that manner, they continuously and dynamically change every user's profiling, until a final equilibrium profile state is achieved. Test bed experimental results within the SPERO project indicate that after an approximately 50% sample of questions has been answered, it is possible for the system to balance to a solid, static, initial user profile with great confidence.

The final output of this process, following the application of the weights, is the extraction of a "1-1" profile–end user relation. In that manner, each end user is classified to an initial, static profile that characterizes his behavior, his interests and his further treatment from the system. This particular profile characterization forms the basis of the following intelligent clustering procedure, which includes the notion of profile extraction and integration within this system. In Figure 10 we present an indicative sample of the end-users' static profiling, extracted by previously analyzed procedure within the SPERO system, according to the users' answers collected by 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 10. Initial static profiling mapping

4. CLUSTERING-BASED PROFILING

The core of the clustering data concept is to identify homogeneous groups of objects based on the values of their attributes. It is in general a difficult problem to tackle and is undoubtfully related to various scientific and applied fields [1]. The problem gets more and more challenging, as input space dimensions become larger and feature scales are different from each other, as is the case in our system. In particular, a consideration of the original set of questions of the e-questionnaires as input space, results into a large number of 176 unique features to be taken into consideration when performing clustering on the user answers. The best way to go in this direction is to use a hierarchical clustering algorithm, which is able to tackle such a large scale of features [3]. Although such a method does not demand the number of clusters as input, still it does not provide a satisfactory framework for extracting meaningful results. This is mainly due to the "curse of dimensionality" that dominates such an approach, as well as the inevitable initial error propagation and complexity along with data set size issues.

In order to increase the robustness and reliability of the whole clustering step of our system, the use of an unsupervised extension to hierarchical clustering in the means of feature selection was evident [3]. Using the results of the application of this clustering to a portion of the system's dataset in question are then refined and extended to the whole dataset. The performance of the proposed methodology is finally compared to the previous step of fixed clustering, using the predefined profile characterizations as a priori label information.

The general structure of such hierarchical clustering algorithms, which forms the structure of SPERO's clustering approach as well, is summarized in the following steps and presented analytically in [3] :

1. Turn each input element into a singleton, i.e. into a cluster of a single element.
2. For each pair of clusters c1, c2 calculate their distance d(c1, c2).
3. Merge the pair of clusters that have the smallest distance.

4. Continue at step 2, until the termination criterion is satisfied. The termination criterion most commonly used is thresholding of the value of the distance.

It is worth noticing, though, that in our case, where the input space dimensions are large, the Euclidean distance is thought to be the best distance measure used [7]. Still, this is not always the case, due to the nature of the individual features; consequently a selection of meaningful features needs to be performed, prior to calculating the distance [6]. Moreover, one feature might be more important than others, while all of the features are useful, each one to its own degree. In this work we tackle weighting of features based on the following principles:

a) we expect elements of a given meaningful set to have random distances from one another according to most features, but we expect them to have small distances according to the features that relate them,

b) we select meaningful features based on the nature of the specific questions of the e-questionnaires. In particular, system experts perform an initial selection of meaningful questions, restricting the input space dimensions and

c) we further perform a second level filtering of the input data, based on the type of the input, leaving out answers - and thus questions - of arbitrary dimensions, such as free text input boxes of the e-questionnaires. Information collected from such answers fails out of the scope of clustering data and identifying user profiling information, being more useful for plain statistical approaches.

More formally, let c_1 and c_2 be two clusters of elements. Let also r_i, $i \in N_F$ be the metric that compares the *i*-th feature, and F the overall count of features (the dimension of the input space). A distance measure between the two clusters, when considering just the *i*-th feature, is given by:

$$f_i(c_1,c_2) = \sqrt[\kappa]{\frac{\sum_{\alpha \in c_1, b \in c_2} r_i(a_i,b_i)^\kappa}{|c_1||c_2|}} \quad \textbf{(1)}$$

where e_i is the *i*-th feature of element e, |c| is the cardinality of cluster c and κ is a constant. The overall distance between c_1 and c_2 is calculated as:

$$d(c_1,c_2) = \sum_{i \in N_F} x_i(c_1,c_2)^\lambda f_i(c_1,c_2) \quad \textbf{(2)}$$

where x_i is the degree to which i, and therefore f_i, is included in the soft selection of features, $i \in N_F$ and λ is a constant. Based on the principle presented above, values of vector x are selected through the minimization of distance d [14].

In the following, we present the proposed algorithm implementation with our system's data set, using the Euclidean distance as the distance measure. The clustering algorithm has been applied to a small portion of the data set, namely a 10% of the overall system's users; it contained 100 elements/users, characterized by 44 meaningful features/questions. Considering a 25% of the features as meaningful ones, proved accurate and efficient in the process. These features correspond to a set of questionnaire questions that are summarized in the following Table 1 and have been considered appropriate for the profiling extraction process. Features are grouped by the corresponding question of the e-questionnaire.

#	*QuestionID*	*Description*
1	125	Are you a teacher dedicated to or working in Special Education Needs?
2, 3, 4	126, 127, 128	Qualification/training in Special Education Needs?
5	130	Teaching Experience
6	133	Do you have a computer at home?
7	134	Do you have access to the Internet from your home?
8	139	How often do you personally use your Internet connection at home?
9, 10, 11, 12, 13	141, 142, 143, 144, 145	For which of the following did you use the computer at least once in the past month?
14, 15, 16, 17, 18, 19, 20	147, 148, 149, 150, 151, 152, 153	Which of the following tasks have you performed at least once, without any help?
21	155	Are there any computers in your work environment?
22, 23	300, 301	How often did you use the computer last week at the school?
24, 25	168, 169	Do you have access to the Internet or educational software in your work environment?
26, 27	174, 175	In your teaching, how many hours a week, on average, do you use the Internet or educational software with your students?
28, 29	176, 177	Do you use the Internet for search and retrieval of information relating to the needs and problems faced by SEN students?
30, 31	182, 183	Do you use the Internet from the school in order to find additional sources of educational material?
32, 33	400, 401	Do you use the Internet to connect with other schools?
34, 35, 36, 37	261, 262, 264, 265	Is your post permanent – temporary?
38, 39, 40, 41, 42, 43	267, 268, 270, 271, 273, 274	Age of your students
44	275	Area served by your school

Table 1. Features selection

The above elements belonged to three fixed profile classes, but this labeling information was not used during clustering; the labels were used, though, for the evaluation of the quality of the clustering procedure, as described in [3], prior to projecting the results to the whole data set. More specifically, each cluster was then assigned to the class that dominated it. Results are shown in Tables 2, 3 and 4, whereas the numbers inside parenthesis separated by commas denote the elements belonging to its one of the three profile classes in each step.

Performing the initial clustering on a mere 10% subset is not only more efficient computationally wise, it is also better in the means of quality and performance, when compared to the approach of applying the hierarchical process to the whole data set. Although clustering over this 10% of the data set resulted in different possible identifiable clusters, optimal results have been obtained for a number of nine clusters, as indicated in the following Tables 2-4, where clustering results are presented for three variations of output clusters (3, 5 and 9):

Clusters	***Elements***	***%***
1st	(2, 6, **9**)	(11.77%, 35.29%, **52.94%**)
2nd	(11, 2, **25**)	(28.95%, 5.26%, **65.79%**)
3rd	(14, 1, **30**)	(31.11%, 2.22%, **66.67%**)

Table 2. 100 users clustering results – 3 clusters

Clusters	***Elements***	***%***
1st	(3, 1, **7**)	(27.27%, 9.09%, **63.64%**)
2nd	(5, 1, **8**)	(35.72%, 7.14%, **57.14%**)
3rd	(5, 1, **13**)	(26.32%, 5.26%, **68.42%**)
4th	(5, 9, **11**)	(20.00%, 36.00%, **44.00%**)
5th	(11, 1, **19**)	(35.48%, 3.23%, **61.29%**)

Table 3. 100 users clustering results - 5 clusters

Clusters	***Elements***	***%***
1st	(1, 1, **4**)	(16.66%, 16.66%, **66.66%**)
2nd	(0, 1, **6**)	(0.00%, 14.28%, **85.71%**)
3rd	(4, 2, **5**)	(36.36%, 18.18%, **45.45%**)
4th	(3, 3, **6**)	(25.00%, 25.00%, **50.00%**)
5th	(4, 2, **5**)	(36.36%, 18.18%, **45.45%**)
6th	(**8**, 4, 5)	(**47.05%**, 23.52%, 29.41%)
7th	(**4**, 1, 4)	(**44.44%**, 11.11%, **44.44%**)
8th	(3, **10**, 6)	(15.78%, **52.63%**, 31.57%)
9th	(1, **4**, 3)	(12.50%, **50.00%**, 37.50%)

Table 4. 100 users clustering results – 9 clusters

As expected, the results of the clustering step demonstrate the clear trend underlying in the system's input data: System users are characterized by intermediate ICT skills and expertise. This observation is extremely evident in the third column of Table 4, which indicates clearly that most users of the system belong to the static, intermediate "Advanced" profile. The first two clusters identified by our algorithm are unambiguously dominated by the third profile class. Additionally, clusters 3, 4 and 5 indicate a clear majority of the third class in their elements as well. Consequently, 5 out of 9 clusters (55.55%) are indicating a clear advantage of the "Advanced" profile class. Moreover, cluster 7 acts as an intermediary between classes "Advanced" and "Expert", as it illustrates a draw in the elements between those two profile classes. Clusters 8 and 9 are dominated by the "Beginner" profile class, whereas cluster 6 forms a solid representative of the "Expert" class. The above clustering approach forms the basis procedure, with the aid of which each SPERO end user is automatically categorized to a specific profile class that characterizes his behavior and his future interests and choices within the system. According to the cluster to which each user belongs, educational content, appropriately selected by the system's experts, is offered to him. Because of flexibility and protection of crucial personal data reasons, the step of user characterization is only provided as an added value characteristic to the users that are willing to use it. Suitable verification procedures ensure that content offering filtering features are only enabled according to each end user's will.

5. CONTENT ADAPTATION & USER TRACKING

The SPERO system software forms an integrated, web-based learning portal, designed and implemented according to well-known learner-friendly solutions and flexible e-learning software applications [10]. When system's users visit the SPERO portal, validation against the system user database is performed. Subsequently, they are called to answer the e-questionnaires in order to automatically establish their user profile. This automatic profile extraction provides the extremely useful and fully personalized information needed. Learning resources have been linked up to each profile category that has been defined during the profile extraction process. The group of system's experts is responsible to provide a set of appropriate selected e-courses to each group of user profiles, indicated by the nine groups identified during the previous clustering step.

SPERO 's content offering [9] contains links to educational content, separated into various sectors and providing services, like: *Courses Catalogue*, *Announcement Service*, *Search Service*, *E-mail Service*, *Upload Files* and *Help Service*. In particular, 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 11.

Figure 11. 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 also designed and implemented. Indicatively, groups of e-courses characterized by increasing difficulty and strong topic relativity are possible, such as the following chain of e-courses: *Introduction to Information's Technologies, Introduction to Operating Systems*, *Presentation and usage of Office and Educational software, Introduction and Usage of Internet.* Each group is characterized by the following aspects:

- **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.

The offered e-courses correspond to the available SPERO user profiles obtained previously. This is ensured by an intelligent user tracking mechanism. This mechanism is based on each user individual session, the starts and stops of which are signaled by the time the user enters and leaves the SPERO portal respectively. Session information, along with validation and user access rights is stored in the “userDetails” part of the static user profile. It provides a robust and reliable method to ensure independency amongst the system’s users and efficiency of the whole user profiling-based content offering. As an illustrative example, consider that students who receive an e-course are tracked throughout the SPERO system and their behavior is observed and tracked internally. The overall procedure is transparent and provides the main source of feedback for the users' future system (e-courses and material) selections.

6. CONCLUSIONS & FUTURE WORK

The current approach to e-learning applications forms an integrated, state-of-the-art system that is able to identify its individual users. It extends work performed on precise, high level personalization algorithms. The system utilizes internally personalization techniques towards profiling extraction, introducing a novel conjunction of static and dynamic profiling mechanisms. 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.

This work is part of our ongoing efforts in the field of designing and implementing an integrated, fully automated e-learning portal system. The main focus is given to the personalization aspects of the system’s user handling and educational content offering. Possible future work includes better selection of the clustering algorithm threshold criteria and possible increase of the static profiles categories. Moreover, the system’s e-questionnaires are susceptible to evaluation and improvements, as well as an increase in the number of participants in the e-surveys is viable. The overall proposed architecture of SPERO could be easily adapted to other e-learning schemes, mainly due to its robustness and entities clarity.

7. ACKNOWLEDGMENTS

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".

8. REFERENCES

[1] Hirota K., Pedrycz W. (1999), Fuzzy computing for data mining, Proceedings of the IEEE 87:1575–1600

[2] Mylonas P., Tzouveli P. and Kollias S., Towards a personalized e-learning scheme for teachers, 4th IEEE International Conference on Advanced Learning Technologies, Joensuu, Finland, August 2004

[3] Mylonas P., Wallace M. and Kollias S., Using k-nearest neighbor and feature selection as an improvement to hierarchical clustering, 3rd Hellenic Conference on Artificial Intelligence, Samos, Greece, May 2004

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

[5] Tzouveli P., Tsapatsoulis N., Kollias S., Michaelidou M., On the Development of User Adapted E-learning Schemes, ICNHIE, Santorini,Greece, August 2003

[6] Wallace M., Stamou G. (2002), Towards a Context Aware Mining of User Interests for Consumption of Multimedia Documents, Proceedings of the IEEE International Conference on Multimedia and Expo (ICME), Lausanne, Switzerland

[7] Yager R.R. (2000), Intelligent control of the hierarchical agglomerative clustering process, IEEE Transactions on Systems, Man and Cybernetics, Part B 30(6): 835–845

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

[9] SPERO: "Leonardo Da Vinci" programme, Themes 5: E-Learning, 2001-2004, (http://www.image.ntua.gr/spero)

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

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

[12] SPERO statistics web application: (http://www.image.ntua.gr/statsspero)

[13] SPERO profiles web application: (http://www.image.ntua.gr/statspero/profiles)

[14] Wallace M. and Kollias S., Soft Attribute Selection for Hierarchical Clustering in High Dimensions, Proceedings of the International Fuzzy Systems Association World Congress(IFSA), Istanbul, Turkey, June-July 2003