Proceedings of the IASTED International Conference
WEB-BASED EDUCATION
February 16-18, 2004, Innsbruck, Austria

ON THE DEVELOPMENT OF ADAPTIVE WEB-BASED LEARNING COMMUNITIES

Maria Rigou
Research Academic Computer Technology Institute
Internet & Multimedia Technologies Reasearch Unit
61 Riga Feraiou str., GR-262 21 Patras
Hellas
rigou@cti.gr

Abstract

Online learning communities may greatly benefit from incorporating adaptive features which take advantage of the knowledge and experiences of community members and use it to better serve each individual depending on personal preferences, goals and needs, as well as the history of activity in the community. This paper investigates the incorporation of adaptive features in online learning communities and focuses on deploying web mining techniques for this purpose. It presents a pilot system that experiments with the application of a number of adaptation forms and concludes with identifying some open issues and concerns in the domain of applying adaptiveness to web environments that host learning communities.

Key Words
Learning communities, adaptive, personalization, web mining, recommendations

1. Introduction

Learning is by nature a process closely connected to sociability and in the majority of cases traditional learning implies the formation and operation of a community. Scientific observation during the last years has indicated that learning on the web in many cases is also accompanied and promoted by the creation and maintenance of online communities.

Particularly in the case of online activities that are by nature remote and impersonal, the notion of setting up communities of users is of vital importance on the path leading to successful learning. Research provides evidence that *"strong feelings of community may not only increase persistence in courses but may also increase the commitment to group goals, cooperation among members, satisfaction with group efforts, and motivation to learn"* [1]. Communities share common goals, needs and problems and can promote solutions and progress if one gains insight into their "accumulated" knowledge. Thus, if strong sense of community is related to increased persistence as well as to increased learning, then the sense of community becomes a foundation upon which to design and facilitate online teaching. In real life most communities are formed through geographical proximity, but online communities are mostly formed around a shared interest or need, and are a powerful tool for building trust and relationships, knowledge acquisition and exchange.

Defining online communities is not a trivial task. A search in the related bibliography (in both the sociology and the IT domains) results in a variety of definitions with different focus and prerequisites as to what constitutes an online community. Probably the best known definition of online communities comes from Howard Rheingold [2] who described them as "*cultural aggregations that emerge when enough people bump into each other often enough in cyberspace*" (p. 57). Schmid [3] pursues a more agent-based approach (that does not solely take into account real people), according to which, communities are put together through agents – these can be human or software – which are linked by a common language and set of values and pursue common interests. These agents are tied together through a medium in which their roles interact with each other accordingly. Another approach from the IT domain comes form Preece [4] and according to it, online communities consist of :

- *People*, who interact socially as they strive to satisfy their own needs or perform special roles, such as leading or moderating.
- A *shared purpose*, such as an interest, need, information exchange, or service that provides a reason for the community.
- *Policies*, in the form of tacit assumptions, rituals, protocols, rules, and laws that guide people's interactions.
- *Computer systems*, to support and mediate social interaction and facilitate a sense of togetherness.

Core attributes of an online community (in the sense that communities with more such attributes are clearer examples of communities than those that have fewer) comprise [5]:

- A shared goal, interest, need or activity
- Repeated, active participation, with intense interactions and strong emotional ties between participants
- Access to shared resources with policies to determine access
- Reciprocity of information, support and services between members
- Shared context (social conventions, language, protocols)

The different types of online communities can be broken down by the purpose, and shared characteristics of their members and can be categorized as:

- Communities of *practice*, where individuals share the same profession,
- Communities of *circumstance*, where individuals share a personal situation,
- Communities of *purpose*, where individuals share a common objective or purpose, and
- Communities of *interest*, where individuals share an interest.

In some cases a community may fall into more than one definition, and over time a community may develop sub-communities formed around special interest groups.

Learning or educational communities are typically categorized as communities of purpose, with the purpose being learning. In the context of learning, the introduction of online communities has proved to be a quite promising concept, allowing the improvement of both the quality of online courses and the attractiveness of web-based learning environments. According to Reinmann-Rothmeier et al. [7] a learning community is a community in which people are joined together by a mutual interest to intensively examine a particular theme, and in so doing are able to learn together, exchange existing knowledge and jointly work on aspects of problem solving. Ideally, within the context of a learning community, knowledge and meaning are actively constructed, and the community enhances the acquisition of knowledge and understanding, and satisfies the learning needs of its members. Moreover, communities can counteract the isolation of the independent learner (and the associated dropout quota) [8]. Members of a learning community may be students, lecturers, tutors, researchers, practitioners and domain experts.

Designing and implementing an online environment for supporting a community requires much more that merely providing for the communication and resource sharing capabilities. Online community designers are people who must combine "… the world of technology and the world of people, and try to bring the two together" [6]. In attempting to set up a successful online community many things can go wrong and the road from assuring all technical prerequisites to having people participating and keeping the community alive, is long and winding. This paper argues that by monitoring the behavior of community members, their expertise, skills, opinions and/or preferences and requirements and by applying certain adaptation mechanisms, the experience and effectiveness of learning online can be drastically improved. The rest of the paper is structured as follows: section 2 introduces the adaptiveness dimension and describes the process of applying it using web mining for delivering personalization. The third section presents a pilot system for supporting learning communities. Section 4 discusses some open issues and concerns and the last section concludes.

2. The Adaptiveness Dimension

Today's web-based learning environments, apart from ensuring high quality content, correct and efficient structuring, as well as support for the tasks of all user profiles participating in the learning process, have drastically evolved and incorporated methods and techniques from other domains and application areas (such as data mining, web content, structure and usage mining, user modeling and profiling, artificial intelligence and agent technologies, and knowledge discovery). More recently, techniques that were initially developed for the e-commerce domain, in support of activities such as personalization, cross-selling, up-selling, and recommendations (based on the underlying technology of clustering, similarity indexing, association rules mining, collaborative or content-based filtering, and more) are transferred and applied to e-learning applications.

These techniques aim to tailor and deliver to the user an instance (or a "view") of the e-learning environment that best suits his personal needs, preferences and objectives, or the view that best implements the teaching strategy decided by the tutor for the specific student cluster. This approach has been dictated by the fact that just like in real life, in online communities user tasks are different and users themselves are different. To this end, researchers develop systems that are able to adapt themselves by observing, recording and analyzing user activity (*adaptive* systems), or can be explicitly "tuned" by the user (*adaptable* or *customizable* systems) [9].

Research activity in the e-learning domain and more specifically the ways of applying adaptive (or personalized) features in web-based learning environments has been intense. Several years after Brusilovsky's work of 1996 [10] on methods and techniques of adaptive hypermedia, it is widely accepted that adaptive systems adapt to user data (goals/tasks, knowledge, preferences, interests, etc.), usage data (data about the user interaction that cannot be resolved to user characteristics) and/or environment data (covering all aspects of the user environment that are not related to the users themselves).

Based on the same source and a later version of the initial document [11], adaptive systems may produce as output:

- adaptive *presentation* (text or multimedia adaptations);
- adaptive *navigation support* (link hiding, sorting, annotation, direct guidance, or hypertext map annotation);
- adaptive *link generation*: discovery of new links and addition to the rest, link generation for similarity-based navigation, or dynamic recommendation of relevant links;
- adaptation of *modality* (in the sense that apart from using text in order to communicate content other media types may be used).

Figure 1 illustrates a typical personalization process based on web usage mining (adapted from [12]).

At a research level, certain systems have focused on specific aspects and theoretical issues deriving from the area of adaptive web applications and that of teaching and learning strategies; we indicatively refer to some of the most representative ones. On the topic of personalizing web-based learning InterBook [13] focuses on adaptive navigation support in e-learning systems and more specifically on link annotation techniques, while AHA! (Adaptive Hypermedia Architecture) uses link hiding [14]. NetCoach [15] derived from ELM-ART, which was one of the first adaptive web-based educational systems [16], and is a system designed to enable authors to develop adaptive learning courses without programming knowledge. WebPersonalizer [17] is a more general-purpose system used to provide a list of recommended hypertext links to a user while browsing through a website. OOHDM (Object-Oriented Hypermedia Design Method) is a methodology for designing personalized web applications and managing personalized views [18].

3. An adaptive system for supporting online learning communities

With the purpose of examining the effect of adaptation mechanisms for delivering personalized views to members of an online learning community, a pilot system has been implemented (details on the design, development and functionalities of the system can be found in [19].

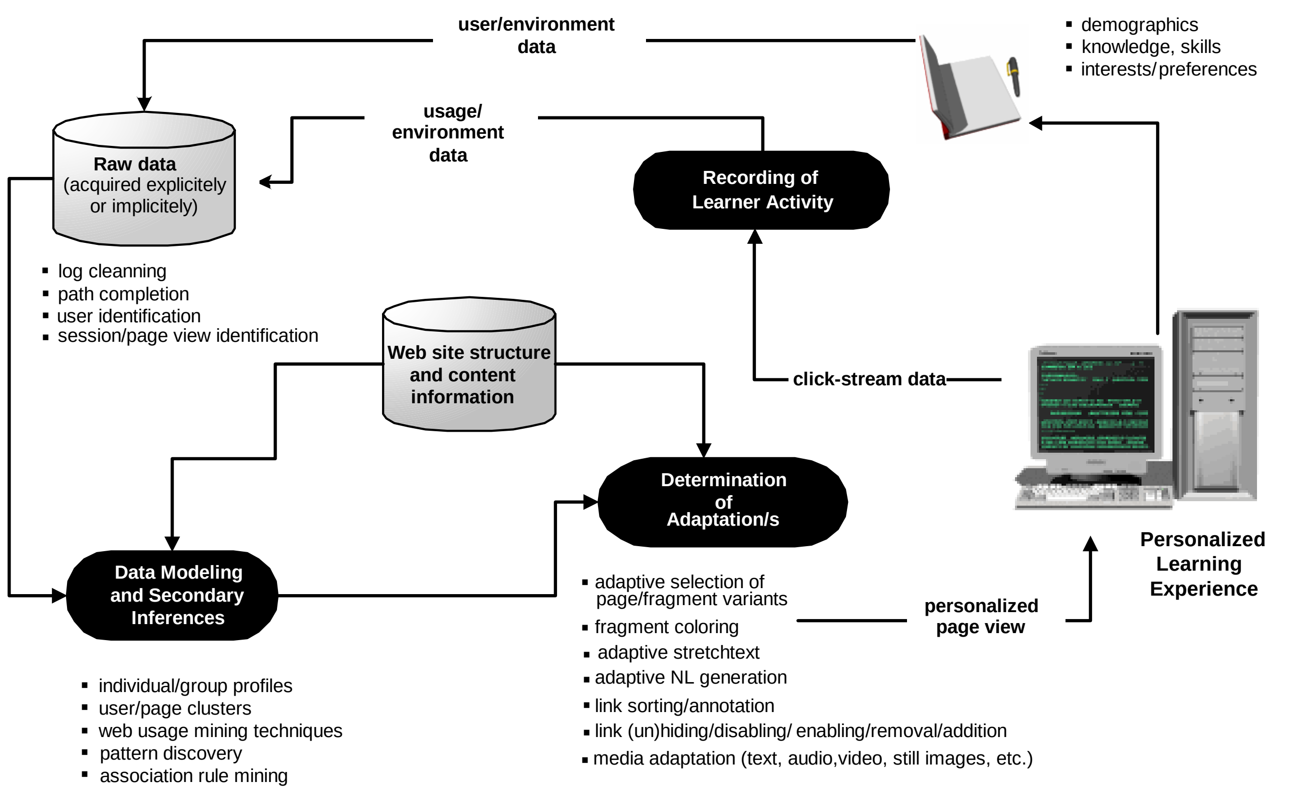

Figure 1. The Process of Personalizing the Learning Experience using Web Usage Mining.

The system extracts community knowledge and experience from the recorded personal learning history of the community learners and combines it with the domain expertise and didactic experience of the community teachers, as well as the activity level of each learner, thus resulting in the construction and delivery of personalized system views constructed dynamically and delivered to each community member. The system experiments with various adaptation forms in order to deliver personalized content and cope better with diverse user profiles, preferences, goals and needs. This section focuses on selected issues regarding the implementation of the adaptive features and the resulting personalized views aiming at upgrading the overall user experience, allocating more power to users and humanizing the feeling of belonging to an electronic community. Figure 2 captures various parts of the user interface elements of the system and the way adaptations are delivered.

Determining system views based on the user profile (administrator, tutor or student).

In the pilot system, the options available in each profile view vary: while administrators are offered the complete set of options and functionalities (in the form of hyperlinks that lead to forms), tutors have a comparatively smaller set of options (since they do not need access to account management options, neither to personal account data). Students on the other hand, can access an even smaller part of options, since they should not be able to interfere in composing new course material or altering the existing modules, neither in determining the underlying recommendation mechanisms. Views are determined using simple filtering and implemented using link hiding (e.g. neither students nor tutors see the hyperlink *"See pending New Account requests"*, available to administrators).

Visual representation of student activity level.

This approach was taken on the assumption that the system should foster a rewarding method for those active participants of the learning process, allowing the positive distinction of certain students by displaying a number of stars beside their nickname. More specifically, for each student in the system a record is maintained for storing profile and usage data. Material coverage $C_{i,j}$ of skill *i* by user *j* complies with customary coverage definitions:

$$C_{i,j} = \frac{|R_i \cap R_{i,j}|}{|R_i|}$$

where R_i is the number or all available topics in skill *i* (e.g. all topics in the MS Word skill) and $R_{i,j}$ is the number of topics in skill *i* marked as read in the personal progress of student *j*. The maximum number of stars (corresponding to available scaling levels along with the actual function that allocates students a certain number of stars) can be determined by administrators or tutors. Usage data collected using cookies and server log analysis are currently used to calculate and deliver the adequate number of stars characterizing the current user, but can also feature as a quite descriptive source of data for assessing the overall user activity on the part of the tutor. Administrators can tune the calculations to match any didactic scenario of student "rewarding" through visual clues in his/her representation in the system communicational areas. Scenarios may be based on any –weighed- combination of metrics such as time connected, material coverage, number of messages posted to the Forum, number of questions submitted, etc.

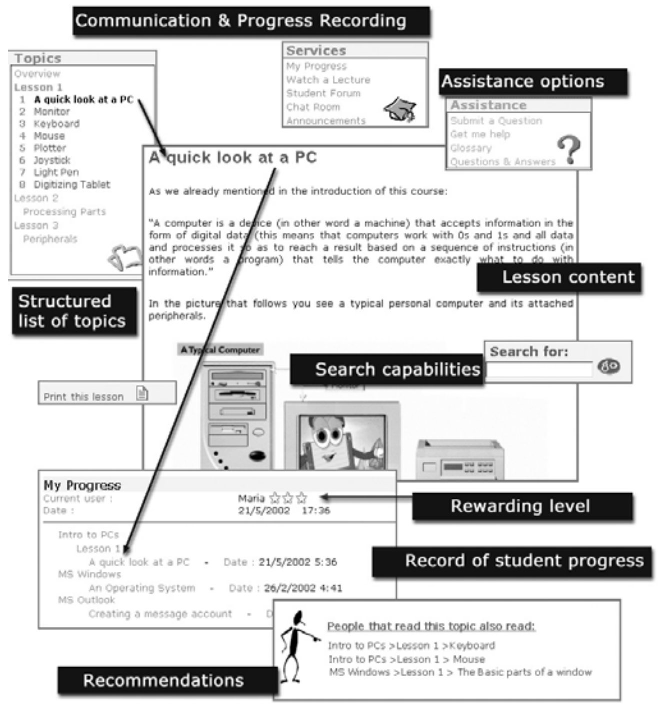

Figure 2. The Pilot System View for Students

Using the recorded activity of the student community to suggest further reading.

The pilot system provides a progress monitoring mechanism that is kept manual (i.e. students are required to explicitly mark the topics they have already studied). The reason for adapting this approach and not resorting to automatic recording of topics pages visited by each user is that topic recommendations produced using student activity should not be based on all topics a student happened to come by while browsing through the available material but only those he/she took the time to study. This way we can hope for more qualitative recommendations. The production of recommendations for further reading is based on association rules mining: topics marked as read in the progress of students that have also read the current topic are recommended under the *"People that read this topic also read…"* section). Association rules ([20], [21]) are used to capture the relationships among topics based on co-occurrence patterns observed in the personal progress during successive student sessions.

For association rules of the form "$A \rightarrow B$", where A and B are sets of topics (A is the set of topics in the current student's progress, and B is the set of candidate topics to be recommended to the student), support is defined as:

$$\sup port(A \rightarrow B) = \frac{\#\_progress\mathrm{Re}cords\_containig\_both\_A\_and\_B}{total\_\#\_of\_progress\mathrm{Re}cords}$$

The *support* of an association rule refers to the percentage of the progress records (in our case) for which the rule is true.

A certainty measure for association rules of the same form is *confidence*. Given a set of recorded studied topics A (in each student's progress), confidence is defined as:

$$confidence(A \rightarrow B) = \frac{\#\_progress\mathrm{Re}cords\_containig\_both\_A\_and\_B}{\#\_progress\mathrm{Re}cords\_containing\_A}$$

For example, the association rule:

$$\{topic123, topic34\} \rightarrow \{topic15\}$$
$$[support=0.02,\ confidence=0.68]$$

conveys the relationship that students who read topic123 and topic34 also tend (with a confidence of 68%) to read topic15. The *support* value represents the fact that the set $\{topic123, topic34, topic15\}$ is observed in 2% of student sessions recorded in their respective personal progress. Association rules mining typically identifies URI references recorded in server logs on a per-session basis and requires log analysis in order to derive sessions/transactions and then references to URIs of interest, but in our case the personal progress provides a more secure (since we indeed want to recommend topics actually studied by other students and not just accessed) and less demanding option (in terms of required processing). Recommendations returned to the user depend on the minimum support and confidence values set by administrators, as well as the preferences set in the current user account data (a user may change the maximum default number of recommended topics or even disable recommendations, at all).

Using the tutors' expertise to suggest further reading. A second set of recommendations is assembled and placed under "*Your tutors suggest that you also study…*" and contains topics recommended based on associations defined by tutors. The tutor that uploaded a new topic creates context links towards topics that relate to the concepts and terms encountered in the new topic. These connections are then used by the system to set up the recommendation list for the students that will study this topic. Again, students can determine whether there will be a recommendation list on his/her page and how many topics will be on that list. Naturally, both types of recommendations exclude from their list the topics already marked as read by the current user.

4. Concerns and Open Issues

Speed – web users have a certain low tolerance to delays as concerns the time required from when they click on a link or type a URL to the time the requested page appears (or starts appearing) in their browser window. Web based learning environments and especially learning communities (which impose further demands on speed due to their communicational 'ingredients') should not suffer from big delays that may lead to abandonment. The underlying mechanisms, computations, recorded data and mining tasks that are part and parcel of a typical adaptation/personalization process should not jeopardize low response times. This means either that part of the process is executed offline or that special algorithms, structures and configurations assure fast online operation.

Effectiveness – Adaptive approaches to web applications impose demanding space and computational requirements. Thus one has to wonder if the specific application needs such mechanisms and to what degree: which parts should be dynamically adapted and what types of data should be recorded; how these data are to be processed and what kind of techniques should be applied on them in order to produce adaptations; and after all, is it really working? The effectiveness of adaptations and their accuracy (did the system succeed in serving user needs and preferences, or were the recommendations interesting for the user?) are part of a bigger issue researchers refer to as adaptation evaluation, and remains a 'gray' area.

Loss of control - The problem of loss of control is observed in situations where the user is not in control of when and what change occurs and it is referenced in literature as a usability degrading factor. Personalization with all the automated adaptations it 'triggers' transparently, is a blessing only if the user is allowed to control what is adapted automatically and how. This way, locus of control remains at the user side, where it should be.

User intrusion – It may be caused by either explicit profiling (when the user is asked a number of questions so that this information can be used in the corresponding profile) or by the wrong way of delivering personalization features (for instance if recommendations for further reading appear in a new window that is placed on top of everything else and stays there until the user closes it). Interface designers should ensure that the system treats users and their attention with respect (especially when the system is a learning environment where attention should not be distracted).

Privacy – As an alternative to explicit (and intrusive) profiling or in combination with it, adaptive systems try to collect as much data as possible from users, usually without user's initiative and sometimes without their awareness. They are striving to identify the user, record the user's online behavior in as much detail as possible and extract needs and preferences in a way the user cannot notice, understand or control. This situation brings up the invaded privacy hazard. A number of protocols and standards are already in place for protecting the user's right to privacy with the big majority of them is designed for the e-commerce domain (P3P, OPS, CPEX, PIDL and more).

5. Conclusion

This paper investigated online learning communities and the way they promote online learning. Yet another factor that may greatly upgrade the online learning experience and diminish the high drop-out rates is the deployment of web mining for tailoring the learning experience to each individual user (or group of users) and enforcing the sense of belonging to a community. Web usage mining (combined in some cases with structure and content information as well) can discover and take advantage of the accumulated community knowledge recorded in the history and traces of everyday practice of community members, in the form of web server logs. A pilot system that experiments with the application of a set of adaptation (or personalization) techniques incorporated in a web-based learning environment has been described. Finally, the author discussed a number of potential restrictions and concerns regarding the use of web mining and personalization in general, in the domain of online learning and learning communities.

References

[1] A. P. Rovai, Sense of community, perceived cognitive learning and persistence in asynchronous learning networks, *Internet and Higher Education*, *5*(4), 2002, 319-332.

[2] H. Rheingold, A slice of life in my virtual community. In L.M. Harasin (Ed.), *Global networks: Computers and international communication*. Cambridge, MA: MIT Press, 1994, 57-80.

[3] M. P. Schmidt, *Knowledge Communities. Mit virtuellen Wissensmärkten das Wissen in Unternehmen effektiv nutzen (Using knowledge in a business effectively with virtual knowledge markets)* (München: Addison-Wesley, 2000).

[4] J. Preece, *Online Communities. Designing Usability, Supporting Sociability* (Chichester, New York et. al.: John Wiley, 2000).

[5] S. Whittaker, E. Isaacs, & V. O'Day, Widening the Net:Workshop, Report on the Theory and Practice of Physical and Network Communities, *SIGCHI Bulletin*, *29*(3), ACM Press.

[6] M. Kapor, A Software Design Manifesto, in T. Winograd (Ed.), *Bringing Design to Software*. (Reading, MA: Addison-Wesley, 1996).

[7] G. Reinmann-Rothmeier, H. Mandl, & M. Prenzel, *Computerunterstützte Lernumgebungen. Planung, Gestaltung und Bewertung (Computer-supported learning environments. Planning, formation and assessment)* (München: Wiley-VCH, 2000).

[8] S. Seufert, Design and Management of Online Learning Communities, *Proc. of the European Academy of Management (EURAM) 2002*, Stockholm.

[9] G. Manber, A. Patel, & J. Robinson, Experience with personalization on Yahoo!, *Communications of the ACM 43*(8), 2000, 35-39.

[10] P. Brusilovsky, Methods and Techniques of Adaptive Hypermedia. *User Modeling and User Adapted Interaction 6*, 2/3, 1996, 87-129.

[11] P. Brusilovsky, Adaptive Hypermedia. *User Modeling and User-Adapted Interaction 11*, 2001, 87-110.

[12] P. Markellou, M. Rigou, & S. Sirmakessis, Mining for Web Personalization, in A. Scime (Eds.), *Web Mining: Applications and Techniques* (Idea Group Publishing Inc., January 2004, in press).

[13] P. Brusilovsky, J. Eklund, & E. Schwarz, Web-based education for all: A tool for developing adaptive courseware. *Computer Networks and ISDN Systems 30*(1–7), 1998, 291–300.

[14] P. De Bra, & L. Calvi, AHA! An open adaptive hypermedia architecture. *The New Review of Hypermedia* and Multimedia 4, 1998, 115–139.

[15] G. Weber, H.-C. Kuhl, & S. Weibelzahl, Developing Adaptive Internet Based Courses with the Authoring System NetCoach, in: P. De Bra P. and Brusilovsky (Eds.) *Proc. of the Third Workshop on Adaptive Hypertext and Hypermedia*, Berlin: Springer, 2001.

[16] G. Weber, & M. Specht, User modeling and adaptive navigation support in WWW-based tutoring systems, in A. Jameson & C. Tasso (Eds.), *User Modeling: Proceedings of the Sixth International Conference, UM97*, 1997, 289–300.

[17] B. Mobasher, R. Cooley, & J. Srivastava, Automatic Personalization based on Web Usage Mining. *Communications of the ACM, 43*(8), 2000, 142-150.

[18] G. Rossi, D. Schwabe, & R.M. Guimarães, Designing Personalized Web Applications, *Proc. of the Tenth International World Wide Web Conference*, Hong Kong, 2001, 275-284.

[19] S. Sirmakessis, M. Rigou, M. Katsis, E. Rigou, & A. Tsakalidis, Learncom: Supporting Communities in the Learning Process, in *Proc. of the Fourth International Conference on Information Communication Technologies in Education*, July 3-5, 2003, Samos island, Greece, 211-217.

[20] R. Agrawal, & R. Srikant, Fast Algorithms for mining association rules, *Proc. of the 20th VLDB conference,* Santiago, Chile, 1994, 487-499.

[21] J. Han, & M. Kamber, *Data Mining Concepts and Techniques* (US: Morgan Kaufmann Publishers, 2001).