Personalization of e-Learning Services using Web Mining and Semantic Web

Sandesh Jain, Dhanander K. Jain, Harihar Bhojak, Ankit Bhilwar, and Mamatha J.

***Abstract*—The e-Learning has become matured learning paradigm with the advent of web based learning and content management tools, and shifted the focus of entire world from instructor centric learning paradigm to learner centric approach. Now for making the learning process more streamlined and standardized, the implementing agencies are emphasizing on moving towards service oriented architectural design approach to create, deploy and manage reusable e-Learning services, thus benefiting education sector. For providing the intelligence to evaluation system and other e-Learning services, various domains like data mining, web mining, semantic web etc. can be utilized intelligently. In this paper, we will describe an approach aiming to achieve personalization in e-Learning services using web mining and semantic web.**

***Index Terms*—Semantic web, web mining, service oriented architecture (SOA), learning object (LO), knowledge space, knowledge domain**

I. Introduction

E-Learning is a general term used to refer to a form of learning in which the instructor and learner are separated by space or time where the gap between the two is bridged through the use of online technologies. With web-based learning, it is possible for the learners to learn from anywhere, anytime, at their pace. Web-based learning brings unprecedented level of accessibility to courses in remote area, courses prohibited by budget constraints, courses updated to recently discovered knowledge, qualified instructors, and instruction at any time. Taking in view the difficulty in maintaining centralized online learning systems for the whole country and geo-graphically distributed administrative domains of educational bodies in the country, a service oriented architecture based learning system is the need of the hour. The strength of this approach is reusability and interoperability. Developing the learning platform based on this approach will involve assembling the required web services. The learning platform can be developed and run on different hardware and software platform as a middleware. Above all the major benefit of this approach is enhancements in quality of e-Learning services on the fly without affecting the underlying architecture of subscribers of services.

II. Personalization in E-learning

The relationship between learner characteristics, tasks assigned and effectiveness is very crucial for defining the environment towards building a personalized online educational system. The target audience are be categorized into four groups viz., Individualist, Result-Oriented, Pragmatic and Avant-Gardes. Based on the main characteristics associated with these groups it is mentioned that the target groups differ in their demands for communication and tutor support as well as group activities and social contacts in an online course.

Personalization in e-learning can be understood as education related technology which is capable of individualizing the interaction between system and learner, based on the personal needs and preferences of learner, and help them to shape their own learning boundaries and to collaborate in terms of thoughts, information and knowledge entities.

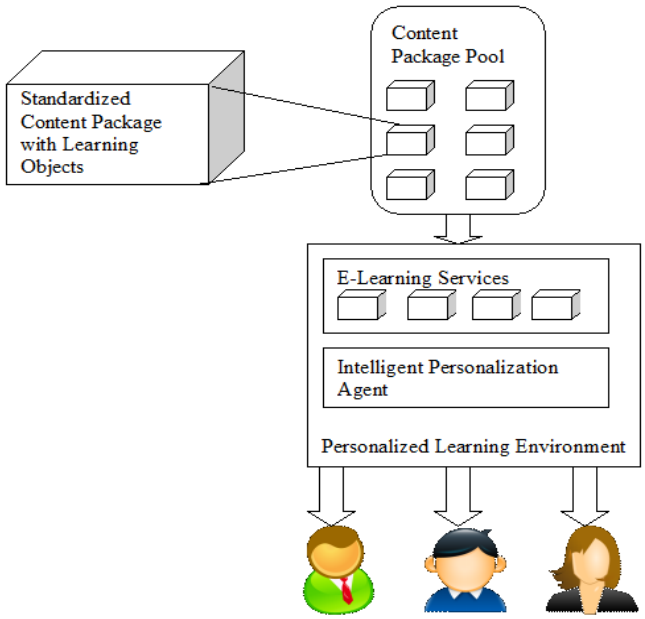

Fig. 1. Conceptual Workflow for Personalization

Personalized learning technological systems (Fig. 1) support learners to set their own learning goals, manage their learning, managing both content and process, communicate with others in the process of learning, and thereby achieve learning goals. These systems may be composed of one or more sub-systems which may in turn built on desktop based application or on web-based services.

III. Personalized E-Learning Services

Some basic and important e-learning services augmented with personalization (Fig. 2) features are as follows:

Manuscript received July 17, 1012; revised September 10, 2012.
Authors are with Center of Development of Advanced Computing, Hyderabad, India (e-mail: hariharb@cdac.in).

A. Personalized Course Management:

- Courses, compliant to e-learning standards, with various forms of web enabled e-content adhering to same learning object, for serving the individual needs and preferences of learners [1].
- Identification and creation of the learner led multidisciplinary courses assisted by instructor for supporting learners to set their own learning goals, manage their learning.
- Intelligent course activity planning based on personalized cognitive patterns.
- Learner preference based self-organizing course content repository.

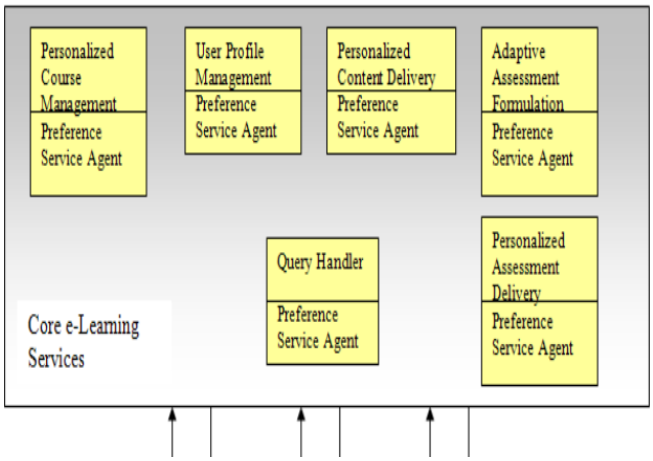

Fig. 2. Personalized e-Learning services

B. User Profile Management:

- Learners' Grade book management for providing automated assistance on specified learning object to the learner, based on individual performance.
- Learners' e-Portfolio management [2], for extracting and interpreting the learners' preferences, based on artifacts, for rendering the e-learning services according to each learner.

C. Personalized Content Delivery:

- Context based device independent content delivery. Eg. Mobile, PC, Palm top, etc. using adaptive hypermedia.
- Multilingual option for delivery of instructions.
- Content delivery of preferred e-content format which is backed with personalized course management service and assisted by user profile management service.

D. Adaptive Assessment Formulation:

- Summative and Formative Assessments.
- Formation of assessment based on the learner's cognitive pattern.
- Building of question repository conforming to standards backed by semantic web, web mining technology and assisted by instructor.

E. Personalized Assessment Delivery:

- e-Learning assessment standard based delivery.
- Personalized interaction types for assessment.
- Tracking and evaluation.

F. Query Handler:

- Supervised automated support for queries implemented using semantic web and web mining technologies.

Technology assisted personalized instructor.

IV. Personalization Using Semantic Web and Web-Mining

Semantic Web [3] is about explicitly declaring the knowledge embedded in many web-based applications, integrating information in an intelligent way, providing semantic based access to the Internet and extracting information from texts. Semantic web would enable web-based applications to interoperate both on the syntactic and semantic level. At its core, Semantic web comprises a set of design principles, collaborative working groups, and a variety of enabling technologies like Resource Description Framework, Web Ontology Language etc. [4], to provide a formal description of concepts, terms, and relationships within a given knowledge domain.

There are a number of important issues related to the Semantic Web. Roughly, they belong to four categories: Semantic Web Languages, Ontologies, Semantic Markup of Web pages and Semantic Web services.

A. Semantic Web Languages

In order to represent information on the Semantic Web and simultaneously make that information both syntactically and semantically interoperable across applications, it is necessary to use specific languages. There are a lot of such languages around and most of them are based on XML (eXtensible Markup Language), XML Schemas, RDF (Resource Definition Framework) and RDF Schemas, all four developed under the auspices of W3C and using XML syntax.

B. Ontologies

Ontology comprises a set of knowledge terms, including the vocabulary, the semantic interconnections and some simple rules of inference and logic for some particular topic. Ontologies applied to the Web are creating the Semantic Web. Ontologies provide the necessary armature around which knowledge bases should be built and set grounds for developing reusable Web-contents, Web-services and applications [5] [6] [7]. Ontologies facilitate knowledge sharing and reuse, i.e. a common understanding of various contents that reaches across people and applications [6].

C. Semantic Markup

Ontologies merely serve to standardize and provide interpretations for Web content, but are not enough to build the Semantic Web. To make Web content machine-understandable, Web pages and documents themselves must contain semantic markup, i.e. annotations which use the terminology that one or more ontologies define and contain pointers to the network of ontologies. Using ontologies as references in marking-up pages and services on the Semantic Web enables knowledge based indexing and retrieval of services by intelligent agents, agent brokers and humans alike.

D. Semantic Web Services

Intelligent, high-level services like information brokers, search agents, information filters, intelligent information integration and knowledge management, are what the users want from the Semantic Web [8]. They are possible only if a number of ontologies populate the Web, enabling semantic interoperation between the agents and the applications on the Semantic Web, i.e. semantic mappings between terms within the data, which requires content analysis [6].

In e-learning domain, semantic web technology can guide and support developers, instructors, and learners to organize, personalize, and publish learning content and even to discover, generate, and compose learning content [5]. A conceptual content development and deployment architecture allows us to distinguish and locate the different applications and to discuss and assess the potential of the underlying technologies.

The knowledge space which gets created due to the interactions of instructors, developers and learners has the implicit relations between knowledge units and behavioral preferences of learners that can be made understood to the machines with the help of semantic web technology [5].

One of the e-learning services which can potentially be enhanced with the use of semantic web is query handler, where an intelligent system can build the semantic relations (ontologies) between different types of queries from various learners and its appropriate answers from instructor, and able to satisfy the upcoming similar queries with refined answers to learners.

The Leap2A specification for e-portfolio portability supported by CETIS and funded by JISC, UK have also taken the semantic web compatibility into account for providing machine understandable collaboration and learning environment.

E. Web Mining

The World Wide Web is huge, widely distributed, global information service center for Information services: news, advertisements, consumer information, financial management, education, government, e-commerce, etc. Web mining refers to the overall process of discovering potentially useful and previously unknown information or knowledge from the Web data. The web mining is at the cross road of research from various research communities, such as database, information retrieval, with in AI and NLP. World Wide Web is a popular and interactive medium to disseminate information today.

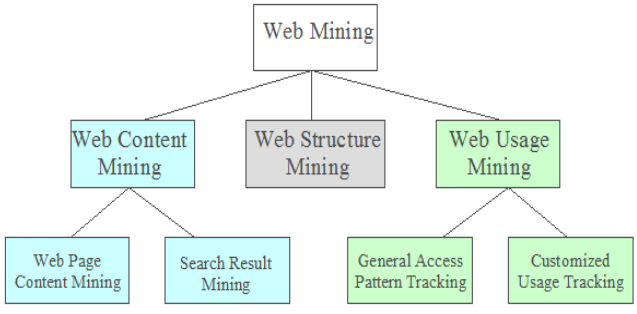

Fig. 3. Categorization of Web Mining

Web content mining (Fig. 3) is the process to discover useful information from text, image, audio or video data in the web. Web content mining sometimes is called web text mining, because the text content is the most widely researched area. The technologies that are normally used in web content mining are NLP (Natural language processing) and IR (Information retrieval).

Web structure mining (Fig. 3) is the process of using graph theory to analyze the node and connection structure of a web site. According to the type of web structural data, web structure mining can be divided into two kinds: extracting patterns from hyperlinks in the web and extracting patterns from document structure.

Web usage mining (Fig. 3), also known as Web log mining, is process of discovering interesting patterns in Web access logs to predict user behavior from interaction with the Web. Two main categories:

- Learning a user profile (personalized)-Web users would be interested in techniques that learn their needs and preferences automatically.
- Learning user navigation patterns (impersonalized)-Information providers would be interested in techniques that improve the effectiveness of their Web site or biasing the users towards the goals of the site.

The usage of web mining for providing personalization in e-learning can be approached as:

- Web Content Mining techniques can be used for the retrieval of relevant content from web to formulate a Learning Object (LO) like topic or chapter, based on learner's preference.
- Web Structure Mining techniques can be used to analysis and establish the linkages in the form of sequencing and navigation between different LO's in order to formulate a instructional unit like Course or Module, based on learner's preference.

Web Usage Mining techniques can be used to track, assist, and guide the learner based on his/her cognitive patterns.

V. Conclusion

In this paper, we have outlined the general principles of a new approach to perform personalization in e-learning platforms by achieving machine understand-ability through semantic web and relying on web mining technology to take care of one of the important steps in content and behavior personalization. Web content mining to standardized educational content metadata, to build content models, and Web usage mining techniques to build learner profiles. We are currently implementing and exploring several techniques and strategies for achieving the personalization in core e-learning services. We are also exploring the possibility of integrating educational preferences in the learner's model such as learning styles, media types, etc. [9] The learner's model to consider, in the future work, should be composed of three main components: learner's profile, learner's knowledge and learner's educational preferences [10]. E-learning systems should be capable of detecting all these components automatically. Constructing the learner models, it lead us for building group models based on collaborative

modeling approach, which in turn enhance the quality of learning object recommendations from an instructional point of view.

REFERENCES

[1] *SCORM Sharable Content Object Reference Model*, Advanced Distributed Learning. (2004). Retrieved January 7, 2009. [Online]. Available: http://www.adlnet.gov/Scorm/ index.aspx

[2] Imsglobal, [Online]. Available: http://www.imsglobal.org/profiles/ index.html

[3] T. Berners-Lee, J. Hendler and O. Lassila, The semantic web, *Scientific American*, 284(5), 34-43, 2001.

[4] [Online]. Available: http://www.w3.org/TR/owl-features/

[5] P. Dolog and M. Sintek, "Personalization in Distributed Learning Environments," *WWW 2004*, ACM 1-58113-912-8/04/0005, New York, USA, May 17-22, 2004.

[6] L. S. Mufioz and J. P. M. de Oliveira, "Applying Semantic Web Technologies to Achieve Personalization and Reuse of Content in Educational Adaptive Hypermedia Systems," *Proceedings of the SWEL Workshop at Adaptive Hypermedia 2004*, pp. 348-353, 2004.

[7] L. Aroyo, D. Dicheva, and A. Cristea, "Ontological support for web courseware authoring," *Proceedings of the Conference on Intelligent Tutoring Systems ITS'02*, 270-280. Berlin: Springer-Verlag 2002.

[8] N. Henze, P. Dolog and W. Nejdl, "Reasoning and Ontologies for Personalized E-Learning in the Semantic Web," *Educational Technology & Society*, 7 (4), 82-97.

[9] R. Boondao, A. J. Hurst and J. I. Sheard, "Understanding Cultural Influences: Principles for Personalized E-Learning Systems," *Proceedings of World Academy of Science, Engineering and Technology* Volume 36, ISSN 2070-3740 December 2008.

[10] D. Sampson, C. Karagiannidis and Kinshuk, "Personalised Learning: Educational, Technological and Standardisation Perspective," *Interactive Educational Multimedia* number 4 pp 24-39, April 2002.

[11] J. Bourdeau and R. Mizoguchi, "Collaborative ontological engineering of instructional design knowledge for an ITS authoring environment," *Proceedings of the 6th Intl. Conference on Intelligent Tutoring Systems*, ITS'2002, 399-409, 2002

Sandesh Jain received his Bachelor's degree from RGPU, Bhopal in Information technology. He is currently working as Sr. Technical Officer and leading the Personalized eLearning Team at Centre for Development of Advanced Computing, Government of India. He also acted as Co-Investigator for several government funded projects. His areas of interest are Web Technologies, eLearning, Cloud computing, Data Mining and Data warehousing.

Dhanander K. Jain received his Bachelor's degree from Thapar Engineering College, Patiala in Electronics and Telecommunications. He has over 30 years of experience in R&D, Training and Project Planning, Coordination, Implementation and currently working as Director at Centre for Development of Advanced Computing at Hyderabad. His areas of interest include Ubiquitous computing, Information Security, Embedded Systems design & VLSI design, Wireless Sensor Networks, eLearning & Education Technologies.

Harihar Bhojak received his Bachelor's degree from RU, Rajasthan (India) in Computer Science. He is currently working as Technical Officer at Centre for Development of Advanced Computing (CDAC), Hyderabad. His areas of interest are Web Technologies, eLearning and Cloud computing.

Ankit Bhilwar received his Bachelor's degree from UPTU, U.P. (India) in Information Technology. He is currently working as Technical Officer at Centre for Development of Advanced Computing (CDAC), Hyderabad. His areas of interest are Web Technologies, eLearning and Cloud computing.

Mamatha J. received her Bachelor's degree from JNTU, Hyderabad (India) in Computer Science. She is currently working as Technical Officer at Centre for Development of Advanced Computing (CDAC), Hyderabad. Her areas of interest are Web Technologies, eLearning and Semantic Web.