CHAPTER 3
SYSTEM OVERVIEW

3.1 INTRODUCTION

E-learning allows learners individually, to learn, ‘anywhere, anytime’ and offers immediate access to specific information. However, learners have different behaviors, learning styles, attitudes and aptitudes which affect their learning process, and therefore learning environments need to adapt according to these differences, so as to increase the results of the learning process. The main objective of this research work is to discover learning materials and effectively design the adaptive e-learning content, in order to improve the learning progress of the learners. Towards this end, the research work aims to retrieve suitable learning material, dynamically model the learner profile, classify the learners and finally adapt the learning content according to the learner profile, learner behavior, pedagogical aspects and contents of learning material. The major components of the adaptive e-learning system, are, content discovery, dynamic learner profiling, learner classification, learning object optimization, content sequencing, learning modules selection and dynamic course composition. In this research, we have used machine learning techniques, to rank the learning materials, to classify the learners, to optimize the learning objects, to activate the learning path and to compose the dynamic courses. In addition, we have used Information Retrieval (IR) techniques, such as, query generation, focused crawling and ranking mechanisms for e-learning content discovery. Moreover, we consider learning objectives, instructional design strategies and pedagogical aspects while providing learning contents.
Figure 3.1 shows the overview of our research work. Initially we start with the retrieval of learning materials from the web, by using query generation, focused crawling and re-ranking methods, based on domain ontology and topic profile. Topic profile is a profile of subject, which has been specially designed for e-learning. It incorporates details from more than one curriculum with pedagogical information, which results in a hierarchical representation of the topics, sub-topics and terms that should be covered in a topic. Next, we convert these learning materials into learning objects using LO Generator (Sathiyamurthy et al 2012). Subsequently, we optimize and sequence the learning objects, in order to generate customized e-content. Then, we model the dynamic learner profile and classify the learners, in order to create an adaptive e-learning environment. Finally, we provide the facility of dynamic learning modules selection and offer the learning courses based on the learning objectives, learners’ performance, behavior and instructional design strategies.

Figure 3.1  Adaptive E-learning system
3.2 E-LEARNING CONTENT RETRIEVAL

The first step in providing appropriate learning content to the learner, is the collection of learning material according to the topic. In this research, content discovery component is designed to discover suitable learning materials for e-learning. This work uses IR techniques, and technologies specifically designed to navigate the web and collect the educational resources, categorized by topic area. An enormous amount of learning material is needed for an efficient e-learning content management system. The availability of online educational contents is rapidly increasing, which leads to issues, such as, increased complexity in the management and accessibility of educational resources from the web.

This work attempts to tackle these issues by proposing two novel approaches specially designed for retrieval of learning material, one using Ontology and the other using a specially designed subject oriented topic profile. The learning resources which we retrieved from the web not only covered the queried concepts but also covered the associated concepts by using domain ontology. To tackle the issues of key word based content retrieval techniques and tf-idf (term frequency-inverse document frequency) based ranking mechanisms, the ontology based content retrieval scheme is required because it integrates some additional information about relationships from the ontology suited for learning. Thus, we use the concepts of a domain specific ontology, to query the web to obtain seed documents. The domain ontology used in this work is a specially designed computer science ontology based on the ACM classification hierarchy (ACM 2012).

The next method is the topic profile based content discovery, which is specially designed to retrieve learning materials with greater topic coverage. This method constructs a topic profile as a spanning tree for a particular course (subject). Topic profiling is a module which takes collection
of documents (which can be text books, presentations or references) as input. We extract the ‘table of contents’ pages from all the input documents (text books). The table of contents corresponding to all topics in the index is clustered, and we compute the Kullback Leibler (KL) divergence (Kullback & Leibler 1951) between each index’s cluster probabilities. We construct topic profile as an acyclic minimum spanning tree for each course which has the topics, sub-topics, and terms. This tree is the topic profile that needs to be followed in the course (Profile of subject). Based on the topic profile levels and terms, we present an automatic generation of simple and concatenated search engine queries, which are relevant to specific topics, to obtain seed URLs for topic focused crawler.

This component also integrates a re-ranking scheme, using topic profile based Latent Dirichlet Allocation (LDA), to re-order a set of documents, so as to improve the retrieval accuracy and maximize the topic coverage at the top ranks of final results. In general, document ranking methods, depends upon the essential measurements of similarity between documents and query. The ranking of results of typical search engines, are determined by the frequency level of occurrence of the query word in the document, in other words is based purely on similarity between query word and document under consideration and on hyperlink analysis of the web graph. However, in this work from the e-learning viewpoint, the significance to the query depends upon the relevancy of the document to the topic of the query and hence, the retrieval needs to be based on the query, an essence of the topic related to the query and the documents themselves. Therefore, we propose topic profile based LDA for the re-ranking scheme which has the capability of finding the hidden structure of the ‘topics’ in the resulting documents. The topics which are not comprehensive while describing some other important topics are said to be hidden topics. We also state that re-ranking of results could help in minimizing redundancy among the retrieved contents based on the content similarity.
3.3 DYNAMIC LEARNER PROFILING AND CLASSIFICATION

The main challenge in e-learning systems is the heterogeneous and changing needs of each individual in the system. Ignoring these needs, lead to a uniform static profile construction where, “one size fits all” policy may demotivate the learner from achieving his learning goal. Adaptive e-learning is a significant requisite to improve the learning performance of the learners. Providing adaptive learning environment and catering to the changing needs and behavior of the learner, can be achieved by evolving dynamic learner profiles. Hence, to provide the dynamic element of the learning process to the learner, a dynamic learner profile is created and updated actively.

In this work, dynamic learner profiling integrates navigation logs to analyze the learners’ behavior over a period of time. This use of dynamic learner profiles results in improvement in the learning process, by considering both personalization and changing requirements. However, profiling learners’ behavior dynamically is an important part which is missed out in most of the existing e-learning systems (Chiu 2008). In addition, we need to identify their learning capability, based on their performance and knowledge level over a period of time. Therefore, we classify the learners according to their knowledge level and performance. While tracking learners, their overall test performance results can be collected. Based on these results learner’s knowledge level, learners' state and the complexity level of the learning content can be determined, using classification techniques. In this research, machine learning techniques, such as, Bayesian Belief Network (BBN) and Decision Tree (DT) are used to classify the learners. Once the learners get classified, this classification could be updated in the learner profile immediately. Hence, these classification techniques play a major role in learner profile construction.
3.4 LEARNING OBJECTS OPTIMIZATION AND CONTENT SEQUENCING

In adaptive e-learning, there is also a big challenge in the design of learning resources. Learning material varies in level at which the topic is covered, granularity of coverage, pedagogical perspective of coverage and assumptions regarding learner’s previous knowledge. Currently, learning objects are used to develop and deploy the e-learning content in new and interesting ways. Due to heterogeneity in the available learning material, the nature and level of the learning objects, produced from such material, varies. In addition, the presented learning content (in the form of learning objects) should not have redundancy and depending on the learner the content offered should maintain a specific sequence. In order to provide the enriched learning content to learners, and to avoid the redundant learning objects, there is a need to optimize the learning objects, while offering the learning courses.

In LO optimization and sequencing component, we use Genetic Algorithm (GA) with a modified three-parent-crossover and used swap mutation operator in a new way to optimize and sequence the learning content (learning objects). The first step is randomly generating an initial population of topics (set of learning contents with their topic id and pedagogy id), called individuals, which represent different possible solutions to the problem. In our scenario, we define a set of learning objects as one chromosome and each learning object acts as genes of chromosomes. We define the chromosome using a three level topic, pedagogy and learning object representation and associated learner and topic parameters as chromosome metadata which is later used to define the fitness function. Each chromosome acts as pedagogical learning content, such as, concept, definition, introduction, examples, etc. Moreover, the learning content design in adaptive e-learning incorporates the sequencing of learning contents based on the learning
necessity, the learning history information, the curriculum details of the learning objects, and the characteristics of the learning objects and learners. In this work, we consider learning styles and learners’ pedagogy preferences, while sequencing the learning objects.

3.5 DYNAMIC LEARNING CONTENT SELECTION AND COURSE COMPOSITION

Dynamic course composition component is used to compose the adaptive e-learning courses, dynamically based on the learner activities, learning objectives and instructional design strategies. Here, we use machine learning techniques such as Spreading Activation Network (SAN) and Reinforcement Learning (RL) to provide the learning content selection mechanism for learners to choose and obtain their appropriate learning modules at each stage. While the learner learns through the learning content of the course, the system continues utilizing the spreading activation to analyze, pre and post learning conditions and furnish course agenda towards the learning objectives dynamically in an uncertain environment. Here, the SAN consists of learning modules that are interconnected through the conditions and learning module is a representation of the course materials. RL learns the behavior of the users automatically, and provides the course materials dynamically based on the positive and/or negative feedback of the learners.

Though, e-learning has many facilities, it needs a path selection mechanism to select the learning modules, to achieve the course goal. In addition, we use Q-learning, a famous approach to reinforcement learning (Watkins 1992). In Q-learning, the system learns a value function Q, which predicts the value each of its actions in every state. In our scenario, the learning modules are represented as states and links to the states are actions. The learning modules which we used here are learning materials, exercises
and assignments. The system performs the action, observes the new state \( s' \) and receives a reward \( r \). The system uses this information to update the Q function. When an action-value function is learned, the optimal policy can be constructed by simply choosing the action with the highest value in each state.

The forthcoming chapters describe each component of the adaptive e-learning system in detail.