

Analysis of Document Summarization and Word Classification in a Smart Environment

Sree Dharinya*

VIT University, Vellore - 632014, Tamil Nadu, India; ssreedharinya@vit.ac.in

Abstract

Background/Objectives: The need for digital learning is rapidly increasing with the effective granularity of learning objects. Retrieval of learning materials from smart environment helps the learner to have a personalized experience using classification algorithms. **Methods/Statistical Analysis:** The digital learning objects are stored in a local repository in the smart architecture. The framework involves mobile agents which play a vital role in predicting the intuitive nature of the learner which helps the learner to get the targeted content. Content retrieval is based on the document summarisation and word classification which is proposed using classification algorithms like content similarity algorithm. Effective personalization is obtained by better accuracy in terms of precision and recall. **Findings:** Document summarization is done using classification algorithm initially based on the type of document retrieved by the learner. The word classification based on the retrieved document is further processed with a pre processing method which is used in the domain ontology. The similarity of the topical terms retrieved helps in classifying the document and presenting the content to the learner for more personalized experience. Further these contents which are stored in a smart framework include mobile agents which help in providing synchronous communication between the users working on similar problems at the same time and gives uninterrupted content to the user based on the query even while the user is moving in different locations. The content similarity algorithm provides better accuracy of retrieved data thereby providing the exact document requested by the learner. The samples of the data classification provide the expected the result of the retrieved data which involve document summarization based on the word classification. The results have proved that the learner has a better intuitive experience with the targeted document wherein the learner gets better personalization. **Applications/Improvements:** More intelligent agents can be involved for learner as well as author to make the system more personalized and provide better collaboration. The classification algorithms along with clustering of documents obtained after summarization can yield better accuracy for large data sets.

Keywords: Agents, Classification, Document Summarization and Framework, Learning Objects

1. Introduction

Enhancing e-learning systems is challenging for most researchers and practitioners. In the past few years much architecture have been framed using learning objects which are related to intelligent agents to improve teaching learning performances. The framework to obtain a smart collaborative design for the system is the main objective¹. The design involves intelligent agents for more personalized, intuitive, global, contextual, and collaborative system² This design is proposed to overcome the dumbness and intuitively handle the habits and desire of individual users.

The structure and work process of the smart system is

to achieve a multimodal collaborative system. The system design also focuses the learner's habits and helps to recall the learner's previous interaction and errors and then present the new content and the process. The system also aims that completion of coursework does not mean that learning is impossible thereafter and hence it involves the lifelong availability factor of data to the learner.

The design of the system also responds to the different learning styles, desired expectation and speed of the individual user using personalization agents³. The learning objects used are expected to functions such as reusability and extension of objects depending on the individual's desire⁴. Various agents are deployed for developing the architecture⁵. Existing systems provide

* Author for correspondence

less support for developing a completely integrated and efficient e-learning environment which meets the needs of individual learners based on the choice of learning. Existing systems require a huge amount of time to develop videos, slides and all other learning materials which can be compiled and stores in the repository⁶.

Agents efficiently hide the difficulty of complex tasks and they perform tasks on behalf of the user to educate them and also help the users to monitor events and procedures. The agents used in designing the system perform these features for its unique behaviour, when learners interact concurrently. The system is to be flexible in retrieving content for a large variety of systems and different queries based on the different grade and difficulty levels of user⁷. Agents with various characteristics for a ubiquitous learning environment in the smart environment exhibit:

- Personalization-Using personalized agents with respect to learners and instructors this factor is responsible for customizing individual learning styles and the speed of the learner⁸.
- Intuitive contextualization focuses on delivering content depending on the learner request and providing the relevant content from deep data for a specific environment.

Collaboration helps learner to interact with the system and also helps to identify a particular topic before the learner spends too much time or is lost or misunderstood a particular topic. Learning objects for intelligent agents have been identified to support learners and to provide better learning outcomes based on their interest. Some of the current approaches which help in effective learning are the IEEE (Learning Object Metadata) LOM-the first model created and now considered to be too simple. The attributes of IEEELOM are extended for various research perspectives. Sharable Content Object Reference Model by Advanced Distributed Learning initiative (SCORM ADL) - This existing model is considered to be too difficult but its implementation is considered to be very consistent. This model proves efficient in terms of learning objects retrieval⁹. Virtual Mentor System -focuses and suggests that better learning outcomes have been identified compared to traditional class room settings. This is used widely for many online courses. Synchronized Multimedia Integration Language (SMIL) was developed by the World Wide Web Consortium (W3C). This standardization was adopted by the World Wide Web Consortium and is an easy-to-learn.

XML-style, allowing easy design and provides

better annotation with respect to metadata extraction. Researchers have given about fuzzy search, instant search and proximity ranking of documents and process of annotating documents based on these search techniques^{10,11}. These methods can be integrated to get better search result and to achieve efficient space and time complexities.

The National Science, Mathematics, Engineering and Technology (SMETE), is constructed to meet learner's and educator's need. This is identified for multidisciplinary learning and provides presentations in education and training process. It also focuses on interactive learning with different resources.

2. System Model

2.1 Generic Architecture

The generalized architecture for an e-learning system is given in Figure 1. This architecture gives the overall and general interaction of learner and author in a learning environment with a centralized content storage management system. The storage of content is based on learning management system architecture. The learning contents are stored in the form of learning objects. The granularity of these learning objects gives more accessibility to the learning contents based on the request of the learner^{12,13}. The storage of contents and retrieval of learning objects is based on the content stored and the query based extraction of data based on the learner's interest. Figure 1 gives the overall architecture of the e-learning environment.

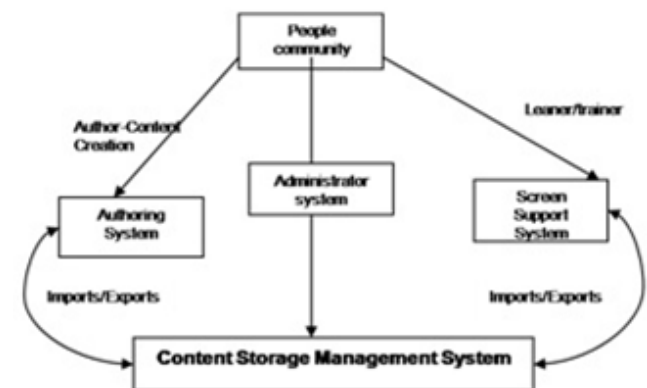


Figure 1. Generic e-learning architecture.

The role of agents in e-learning has proved better personalization. The characteristics of agents involved in e learning are discussed for enhancing the proposed

system model with agent interaction. Intelligent agent is a program which helps to enhance better learning using characteristics such as such as autonomy, better social behaviour, reaction to users and pro active behaviour. Agents work on their own and have a high degree of control over their actions and internal state. This helps in better personalization when learning object retrieval is based on agent interaction. Interaction with learner and system shows the intuitiveness of the user for effective learning¹⁴. Agents perceive their requests and respond in the same time to changes that occur in their environment. This helps in updating the information changes which occur in the learning management system. They also help in communicating with the learner and author side.

Software agents are programs which are defined as an entity and carry out tasks in an organised way as given by the user and also react for the other existing entities involved in the process. They function in a particular agent platform which has other agents and processes^{15,16}. These agents learn from their experiment and they communicate with other agents. Agents play an important role as they are responsible for user friendly interface and provide maximum intuitiveness. The agents to be considered for a general architecture can be listed as the agents for instructors, lesson planning, and resource planning and tutoring. According to the general strategy, a system using intelligent smart agents in a learning environment should abide the following categories of agents, such as diagnostic agent, interacting agent, collaborating agent and social agent^{17,18}.

2.2 A Smart Agent based System

Task Agent: A task agent is involved to do customized tasks like providing information and knowledge by interacting and communicating with other task agents^{19,20}.

Performance Monitoring Agent: The agent is responsible to monitor the work done by the learner. This helps in identifying the intensity of learning of the learner.

Personalization Agent: This agent is identified as a user interface agent between the user and the agent based learning environment²¹. The individual learner can be personalized by a querying process, an indexing process and an evaluation process.

Collaborator Agent: Collaborating agents are used to encourage collaboration between e-learning users and improve the interaction and accuracy of their collaboration^{22,23}. These include suggesting collaboration where appropriate collaboration agents can identify the

tasks of the user in the learning environment provide synchronous communication for the users who are working on similar problems at the same time.

Resource Identifying Agent: The need of the resource identifying agent is to locate the educational resources which are present abundantly in the web and to retrieve them based on the topics, resources and location of the content based on the learners methodology which they support^{24,25}.

Learner Centred Agents: They are responsible for making the learner's interaction with the learning environment to be very effective and problem free. The purpose is to communicate personalization agent and the collaborating agent to enhance learning²⁶. The learner centred agent is responsible for getting feedback from learners about the effectiveness of the targeted learning materials and continuously keeps track of the learning outcomes²⁷.

The learning object repository is responsible for storing learning objects, and metadata which are used for annotating. Annotation tools are used for creating metadata records, searching learning objects, aggregating learning objects into composed learning materials.

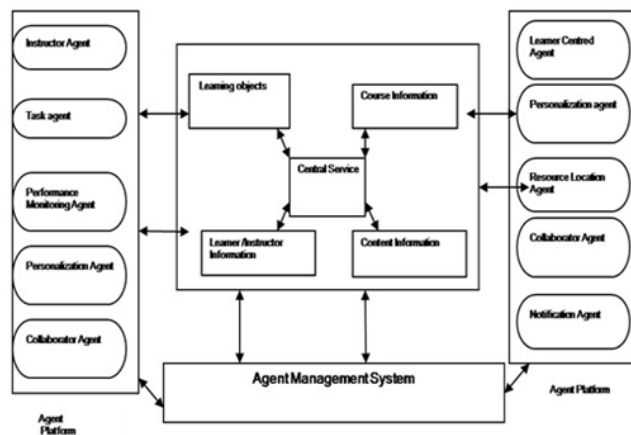


Figure 2. Smart agent based architecture.

2.3 Smart Framework for Mobile Agents for E-Learning

Agent based learning systems are used globally. Smart use of resources of different places is possible. Users do not need to know where the resources are located. The smart framework provides uninterrupted access to the local or remote data which is required to complete laborious calculations²⁸. Recovery protocols, during failure, transactional guarantees such as atomicity, restoration

of operational state after a failure is a few constraints in for e-learning. Services to a smart framework could be massive and this may cause bottlenecks and dynamic learning-agent system. The proposed framework shows the need of mobile agents for retrieving the content from the repository. Figure 3 shows the proposed framework with mobile agents in a grid environment. The proposed framework exhibits the following definitions.

2.3.1 Minimized Communication Costs

Distributed computing interacts between multiple computers through a network. Thereby the cost for communication for learning technology is reduced²⁹.

Asynchronous deployment: After moving migrating to the destination-side computer, a mobile agent does not interact with the source side and if a connection or network is lost on the service providing side the agent works efficiently and reaches the destination.

2.3.2 Direct Access

The agent directly executes on the system in which it is and it can be directly accessed. This is very helpful in detecting the failures and also prevents the installation delays which may be caused due to network problems³⁰.

2.3.3 Dynamic Behaviour

Mobile agents can choose their destination and can be used while it is required. This can be done dynamically and is very useful in smart environments with limited resources.

2.3.4 Easy Access to Distributed Applications

Standalone programs can be easily modified using agents since the mobile agents carry information to different computers.

In Figure 3 the smart frame work with mobile agents is shown. Domain ontology is deployed for storage of content in the smart framework. This helps the user to retrieve the content from the repository even if the user is moving. The mobile agent helps in the continuation of data retrieval on the go from the domain ontology. The contents which are learning objects are store in the repository where the user registers in a regional grid. The intermediate agents interact with the mobile agent to keep the process in continuation. This frame work gives uninterrupted content to the user based on his query even while the user is moving in different locations.

3. Content Search in a Smart Architecture

Courses are not designed in terms of learning objects and storage of learning objects in a repository. The major challenges contribute the methods to decompose complete course into meaningful learning objects, to decide on levels of granularity, to represent such levels, and to show relationships among and between related learning objects at different levels of granularity.

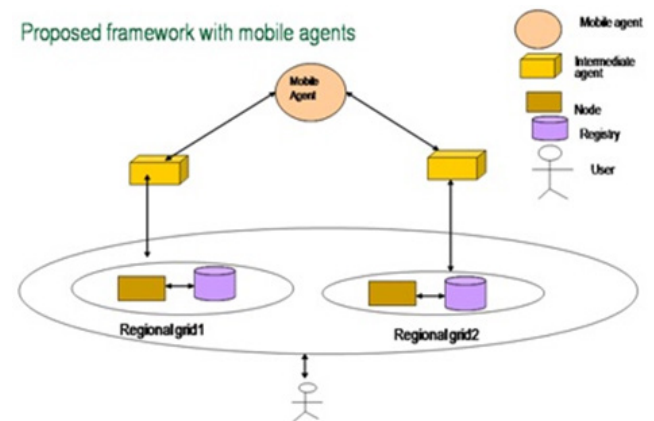


Figure 3. Smart framework with mobile agents.

Redefining the size of the learning object in terms of size and free standing learning objects helps for better granularity. The proof of concept is achieved by obtaining personalization and identifying the difficulties in understanding the levels of granularity and the relationships between them and also to link an adaptive annotation in e-learning. The learning objects can be classified for content retrieval and management, assessing and identifying pedagogical activities which reflect the quality of learning objects.

The constraints in defining the specification of a learning object are identified as the barrier of sharing a measurable learning object and also the lack of ontology. The major difficulty lies in identifying the computational instructional components of the learning objects. The Learning Objects (LO's) are stored in the repository based on domain ontology. The content stored is retrieved based on the content similarity algorithm where in it helps to identify the set of similar learners based on the grade level and the intuitiveness of the learner. The similarity of the contents is identified according to the

learner's choice of query and the similarity is measured with the sample code for the set of learning objects present in the ontology. The content similarity algorithm is used for the repository in the smart environment which shows the search results for a sample set of data. For the learner's choice of topic identify similar learners and initialize their values and get the learners feedback based on the relevant learning object. Then identify the similarity index for each learner and calculate the similarity cosine. A set of similar learners are identified with a similarity index which is greater than the threshold value. The sequence of data classification is showed in Figure 4.

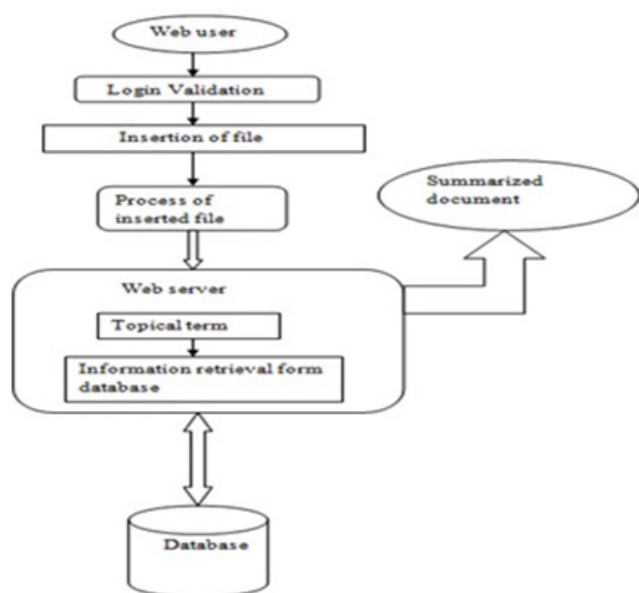


Figure 4. Data classification sequence.

Before sending data from the web user to the web server, the user wants to browse the valid input file to the web server. The web server finds the document to pre-process the data into smaller terms and repeated words are to be calculated and also identify the topical terms related to the given document and matches with database words to find the weight age. The most frequently used words are to be stored into the database tables. Finally the sentence formation is done from the extracted terms. The compressed document is shown to the web user.

3.1 Sample Results of Search Data in the Smart Environment

The classification algorithm in the domain ontology developed using protege has proved more efficient for

a set of data extracted by the user. The similarity of the contents is identified according to the learner's choice of query and the similarity is measured with the sample code for the set of learning objects present in the ontology. The identified results for the different modules with the expected results are shown in Figure 5.

S. No	Module	Expected Input	Expected Output
1.	User Interfaces	IP address	Connection establishment with web server
2.	Peer Login	Username, Password	Login successful
3.	Client Input	Browse File for document summarize	Checks whether file is valid or not
4.	Content Selection and Filtering	Split document terms	Identify terms related to topics
5.	Context based indexing implementation	Topical Terms	Sentence formation
6.	Data presentation	Short sentences	Condensed version of document

Figure 5. Samples of document classification.

For the set of documents the accuracy of the retrieved data is given by precision and recall. Precision and recall are expressed in percentage:

Let A: Number of similar learning objects extracted.

B: Number of similar learning objects not extracted.

C: Number of dissimilar learning objects extracted.

$$\text{Precision} = [A/(A+C)] * 100$$

$$\text{Recall} = [A/(A+B)] * 100$$

4. Conclusion

Learning objects have gained much popularity in recent years, principally because of their reusability. This paper provides a paradigm for providing effective e-learning using agents in a grid environment. Skill based learning environments are used to provide practical skills as well as knowledge development, communication, and problem solving activities. It is vital to provide feedback to the students from these observations based on their actions which helps in assessment and helps researchers to understand the learning process. Using different methods for tagging and using semantic annotation in the smart environments has been investigated for efficient learning. The deployment of this architecture in a smart environment helps better personalization and faster retrieval of search results. The intelligent e-learning agents have the capability of helping authors to design and

schedule courses and obtain learning material location. The mobile agents have the efficiency to select resources and provide solutions. On the whole, agents can be used to provide effective collaboration in the e-learning environment. E-learning mobile agents can be used to provide uninterrupted support to educate learners and also helps authors in different locations.

5. References

1. Lee DH, Shon JG, Kim Y. Design and implementation of OSMD based learning management system for mobile learning. *Indian Journal of Science and Technology*. 2015; 8(1): 154-60.
2. Passier H, Jeuring JT. Ontology based feedback generation in design-orientated e-learning systems. In: Isaias P, Kommers P, McPherson M, editors. *Proceedings of IADIS International Conference on E-Society*; 2004. p. 1-5.
3. Jawahar S, Nirmala K. Quantification of learner characteristics for collaborative agent based e-learning automation. *Indian Journal of Science and Technology*. 2015 Jul; 8(14):1-6. Doi no:10.17485/ijst/2015/v8i14/73119
4. Uddin M, Memon J, Alsaqour R, Shah A, Rozan MZA. Mobile agent based multi-layer security framework for cloud data centers. *Indian Journal of Science and Technology*. 2015 Jun; 8(12):1-10. Doi no:10.17485/ijst/2015/v8i12/52923
5. Murugananthan V, ShivaKumar BL. An adaptive educational data mining technique for mining educational data models in elearning systems. *Indian Journal of Science and Technology*. 2016 Jan; 9(3):1-5. Doi no:10.17485/ijst/2016/v9i3/86392
6. Badri M, Mousavi T, Pour MR, Geravand I, Yeganeh AK. Examine the relationship between use of the ICT and professional development of secondary school teachers in Tabriz. *Indian Journal of Science and Technology*. 2015 Jun; 8(12):1-5. Doi no:10.17485/ijst/2015/v8i12/71737
7. Shu'aibu B, Saud MSB, Bello H, Kamin Y, Buntat Y. Modelling the determinants of ICTs policy formulation in technical and vocational education in Nigerian Institution of Higher Learning. *Indian Journal of Science and Technology*. 2013 Apr; 6(4):1-9. Doi no:10.17485/ijst/2013/v6i4/31853
8. Tella A, Adu EO. Information Communication Technology (ICT) and curriculum development: The challenges for education for sustainable development. *Indian Journal of Science and Technology*. 2009 Mar; 2(3):55-9. Doi no: 10.17485/ijst/2009/v2i3/29416
9. Pandian R, Krishnamoorthy G, Kaliyaperumal K. Rethinking of user studies in academic libraries in the context of paradigm shift due to ICT. *Indian Journal of Science and Technology*. 2009 Oct; 2(10):66-8. Doi no:10.17485/ijst/2009/v2i10/30724
10. Lee SJ, Kim JM, Lee WG. Analysis of elementary students' ICT literacy and their self-evaluation according to their residential environments. *Indian Journal of Science and Technology*. 2015 Jan; 8(S1):81-8. Doi no: 10.17485/ijst/2015/v8iS1/57595
11. Atif Y, Benlamri R, Berri J. Learning objects based framework for self-adapting learning. *Education and Information Technologies*. 2003; 8(4):345-68.
12. De Bra P, Smits D, Stash N. Creating and delivering adaptive courses with AHA. *European Conference on Technology Enhanced Learning (EC-TEL)*; 2006. p. 21-33.
13. Felder RM, Silverman LK. Learning and teaching styles in engineering education. *Engineering Education*. 1988; 78(7):674-81.
14. Felder RM, Spurlin JE. Applications, reliability and validity of the index of learning styles. *International Journal of Engineering Education*. 2005; 21(1):103-12.
15. Limongelli C, Sciarrone F, Temperini M, Vaste G. Adaptive learning with the LS-Plan system: A field evaluation. *IEEE Transactions on Learning Technologies*. 2009; 2(3):203-15.
16. Alumu S, Thiagarajan P. Massive open online courses and E-learning in higher education. *Indian Journal of Science and Technology*. 2016 Feb; 9(6):1-10. Doi no: 10.17485/ijst/2016/v9i6/81170
17. D'Souza MJ, Rodrigues P. Extreme pedagogy: An agile teaching-learning methodology for engineering education. *Indian Journal of Science and Technology*. 2015 May; 8(9):828-33. Doi no:10.17485/ijst/2015/v8i9/53274
18. Kim S. Analysis of an online educational system for early childhood teachers. *Indian Journal of Science and Technology*. 2015 Sep; 8(21):1-5. Doi no:10.17485/ijst/2015/v8i21/79113
19. Zywno MS, Waalen JK. The effect of hypermedia instruction on achievement and attitudes of students with different learning styles. *Proceedings of the Annual ASEE Conference*; 2001. p. 1-9.
20. Sangineto E, Capuano N, Gaeta M, Micarelli A. Adaptive course generation through learning styles representation. *Universal Access in the Information Society*. 2008; 7(1):1-23.
21. Tang T, Hamalainen M. Living lab methods and tools fostering everyday life innovation. *18th International Conference on Engineering, Technology and Innovation (ICE)*; Munich. 2012. p. 1-8.
22. Bui T, Lee J. An agent-based framework for building decision support systems. *Decision Support Systems*. 1999; 25(3):225-37.
23. Jenkins C, Inman D. Server-side automatic metadata generation using qualified Dublin Core and RDF. *International Conference on Digital Libraries: Research and Practice*; Kyoto. 2000. p. 262-9.
24. Jeremic Z, Jovanovic J, Gasevic D. Semantic web technologies for the integration of learning tools and context-aware educational services. *The Semantic Web-ISWC*; USA. 2009. p. 860-75.
25. Kolb DA. *Experiential learning: Experience as the source of learning and development*. USA: Prentice-Hall; 1984.
26. Lau SH, Woods PC. Understanding learner acceptance of

- learning objects: The roles of learning object characteristics and individual differences. *British Journal of Educational Technology*. 2009; 40(6):1059-75.
27. Pai PF, Lyu YJ, Wang YM. Analyzing academic achievement of junior high school students by in improved rough set model. *Computers and Education*. 2010; 54(4):889-900.
 28. Papanikolau KA, Grigoriadou M, Knornilakis H, Magoulas GD. Personalizing the interaction in a web-based educational hypermedia system: The case of INSPIRE. *Journal User Modeling and User-Adapted Interaction*. 2003; 13(3):213-67.
 29. Weber G, Brusilovsky P. ELM-ART: An adaptive versatile system for web-based instruction. *International Journal of Artificial Intelligence in Education*. 2001; 12:351-84.
 30. Zapata-Rivera JD, Greer JE. Inspect able Bayesian student modeling servers in multi-agent tutoring systems. *International Journal of Human-Computer Studies*. 2004; 61(4):535-63.