

Recommender System Based on Web Usage Mining for Personalized E-learning Platforms

Sunil

Section of Computer Engineering
Department of University Polytechnic
Faculty of Engineering & Technology
Jamia Millia Islamia, New Delhi, India
E-mail: sunil@jmi.ac.in

Prof. M. N. Doja

Department of Computer Engineering
Faculty of Engineering & Technology
Jamia Millia Islamia
New Delhi, India
E-mail: mdoja@jmi.ac.in

Abstract: With the rapid increase of information and development of internet the learners prefer online learning. The most important media for the collection, sharing and distribution of information is World Wide Web now days. In a short extent of time E-learning platforms have observed a tremendous growth worldwide. A large quantity of data which is being produced by the E-learning platforms during the transaction is much unpredictable but contains hidden and useful information. From this hidden information learners' usage patterns are find so that the learners can be benefited and provided with an efficient platform and personalized services, this makes a necessity for personalized systems. Web usage mining for E-learning platforms can mine the server logs and database to find the learners usage patterns, to provide the learners with more personalized services. In an E-learning environment a recommender system is a software agent that tries to intelligently recommend actions to a learner based on the actions of the previous learners. This recommendation could be an online activity such as online learning material or could be E-learning web resources. In this regard this paper discusses about the improved mining strategies and builds a model of E-learning platform with learning resources recommendation based on web usage mining which in turn is helpful for learners and provides them an efficient platform and personalized services.

Keywords: Web usage mining, E-learning, Data mining, Recommender system

I. INTRODUCTION

The internet now a day has become an information core that is accessible in a persistent manner. It is an estimable platform for delivering and acquiring information for learners. E-learning systems employ users in a process of learning by providing then essential valuable information on-time and in very effective way. The fast developments in the communication and information technologies have fuelled the pervasive use of internet, which has changed the learning behaviours of the learners.

For a life long term E-learning is undeniably an innovative way to provide education, as compared to the traditional face-to-face style of teaching and learning [1]. Currently more and more learners are getting benefit from the various E-learning platforms. As the learners have a variety of interests for the information collection through E-learning platforms, hence they cannot be treated in an identical way as the learners may have different levels of proficiency. E-learning aims to deliver individualized and dynamic learning content in real time, aiding the expansion of communities of knowledge [2, 5]. E-learning offers new potential in learning. Thus learner can get instant feedback on solution to their problems and learning paths can be individualised. E-learning is a substitute concept to the traditional tutoring system with the application of more and better computer techniques in computer. It is of immense importance to provide a personalized system which can

suggest and adapt to the interests of learners. In recent years most of the E-learning platforms have not been personalized [3].

However, even nowadays, personalized systems are still mainly restrained to research laboratories, and most of the existing E-learning platforms are still delivering the same educational resources in the same way to learners with different profiles. In general, to facilitate personalization, existing systems used one or more types of knowledge. The well-known features of E-learning platforms are the learners become self-learners and explorers of knowledge [4]. E-learning platforms should be able to find the individual differences of learners and build a personalized learning environment to meet the learners' individual requirements. Day by day a large number of E-learning platforms are being designed to provide useful information to the learners [6]. Well designed E-Learning platforms are always considered as successful learning platforms. It must be designed such a way that it can always gratify the needs of the learners. An optimized, updated and structured E-learning platform always attracts learners.

II. AN OVERVIEW OF WEB USAGE MINING

Web Usage Mining is the process of obtaining precious information or knowledge from web data access information. The process of web usage mining can be

divided into the following four stages as shown in the figure1: Data Collection, Data Preprocessing, Pattern Discovery and Pattern Analysis [7].

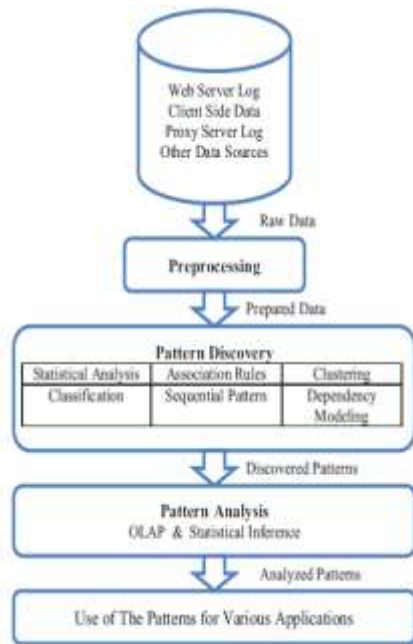


Figure1. Web Usage Mining Process

A. Data collection

Web usage mining consists of many data sources major sources are web server logs, client side data, proxy server log and other data sources. Among these data sources the most important source of data are web server logs. These are the log records of the learners who are visiting the E-learning website. Every time when the learner visiting the E-learning websites the server records their visiting pattern and saves these records in the log files. These log records mainly contains the learners IP address, date of their visit, the type of browser they are using, operating system they are using and so on. Along with server site log records there are other sources of data that are the web site backend database which contains the registered information and academic information of the learners.

B. Data Preprocessing

In the data preparation process the data preprocessing is the initial step that aims to reformat the original web server logs to identify learners session [8]. As the log files which are received from the web server are not suitable for the direct mining of data. There are many pre-treatment methods to make them in proper format so that we can deal with them accordingly. There are four steps of pre-treatment process they are data cleaning, user identification, session identification and path added.



Figure2. Various phases of data preprocessing

As shown in the figure2 preprocessing is done generally in four steps they are data cleaning, identification of user, session and path completion.

C. Pattern Discovery

Pattern discovery is the key process of the web usage mining, pattern discovery covers the relevant data mining algorithms and techniques from several research areas, such as data mining, machine learning, statistics and pattern recognition [6]. The main mining techniques such as statistical analysis, association rules, clustering, classification, sequential pattern and dependency modelling are used to determine rules and patterns. The knowledge that can be revealed is represented in the form of rules, tables, charts, graphs, and other visual presentation forms for characterizing, comparing, predicting, or classifying data from the web access logs.

D. Pattern Analysis

Pattern Analysis is the final stage of web usage mining. The objective of pattern analysis is to reassign the rules, models and statistics found in the pattern to knowledge and then a valuable model will be achieved after the analysis of patterns. There are two most common approaches for the patter analysis. The first one is to use the knowledge query mechanism such as SQL, while another is to construct multi-dimensional data cube before perform OLAP operations [9].

III. RECOMMENDER SYSTEMS

Recommender systems are intended to cover the space between information collection and analysis by filtering all the available data, and presenting the most suitable items to the learners [10]. Currently the recommender systems are extensively adopted in various fields of data mining for the recommendation as recommendation of music, article, movies and even research papers. The learning material can be recommended based on the top overall learners visiting on an e-learning website, or on an analysis of the past learners visiting behaviour as a prediction for future visiting behaviour. The forms of recommendation include suggesting useful e-learning sources to the learners and

providing personalized learning information [11]. Generally, these recommendation techniques are the component of personalization on a site because they help the site adapt itself to each learner.

A. Types of Recommender Systems

- *Content-based Recommendations*

This approach uses a set of distinct features of an item to recommend more items with similar properties. Content-based filtering is based on item description and user preference. In this type of recommendation system, a learner is recommended substances similar to the ones the student preferred in the past. For instance, if a learner named 'Mr. Xyz' reads an E-learning contents like many online research papers on the topic web data mining, then the content-based recommendation techniques will be able to recommend other web data mining research papers to 'Mr. Xyz'.

- *Collaborative Filtering Recommendations*

In this approach of recommender system, a learner is recommended learning materials that other learners with similar tastes and preferences selected in the past. For example, many book recommendation systems assists learners in finding appropriate information on the internet. There are two major approaches for collaborative filtering algorithms: They are Model-based approaches and Memory-based approaches.

- *Hybrid Approaches*

Hybrid approaches combine both collaborative and content based methods. In several cases, combining the results of any two recommendation technique like collaborative and content-based filtering proves to be more competent. It can be done by making collaborative and content based recommendations separately and then integration them or by adding content based to collaborative approach or vice versa.

- *Preference-based Filtering*

Preference-based filtering is a method of predicting the comparative preferences of learners. For instance, in a social E-learning recommendation system, preference-based filtering techniques would focus on predicting the accurate comparative nature of the courses instead of their individual scoring [12, 13].

IV. RELATED WORK

Many authors presented different recommendation systems and techniques to improve the suggestions or recommendations for learners. This section examines the current existing approaches that attempt to address personalised E-learning platforms, recommendation systems and brief mention about some methods proposed by different authors.

K. Reddy et al. [8] proposed an effective data preprocessing method to increase the quality of data by reducing the size of server log files. Several data preprocessing techniques are presented and these techniques are applies on the server log files before the mining process resulting to improve the performance of data preprocessing to identify the unique sessions and unique learners.

Rong Shan et al. [14] have proposed an architecture of Browsing Behaviour Personalized Information Recommendation System (BBIRS) which consists of web usage mining, user interest analysis and personalized recommendation. The proposed system mines the web server logs and learners interactive data by offline and online separately, and the learners interesting degree on the page been calculated by their browsing speed. The system uses web usage mining and analyzing the learners' interest database to recommend the user according to their browsing interest.

Hadeel Alharbi et al. [10] described the merits of various data mining techniques and reviewed the problems related with their limitations in E-learning and social mining. Also discussed about personalized recommender approaches and techniques that represents personalization services which aim to predict learners interest in some elements that existing in the E-learning application systems. Also it is discussed there about the data mining techniques that are used to examine the data that is existed in the form of data formulated by learners in term of posts and comments on social networking sites. The data mining techniques and various tools can be used to embark upon the research question that focuses on mining social networking to personalise and boost the E-learning services.

Chanle Wu et al. [15] proposed a personalized learning service framework that is based on semantic web technologies. The semantic web technologies provide a depiction of the learning contents and semantic user profiles that can outline a good representation of the learning context that improve the quality of the learning service.

Rozita Jamili Oskouei [16] explored an resemblance in learners interest usage patterns, their academic routine, their personal activities, categorize factors for usage patterns on the internet in terms of their average time spend per day and percentage of learners visiting E-learning websites along with their time spend on these E-learning websites.

Nandita Agarwal et al. [17] proposed an algorithm which focuses on the learner based data that can be fetched from the web server log. It is shown there that the work gives the learner based thought concerning the web server logs that are utilized. The entire web log is determined on and the

outcomes are collected on each and every learner basis. The learner based computation makes it more precise and functional for the learner. The final results are calculated on time dependent and independent domains with deliberation of the factors like page visits, time spend on each page, operating system used and the browser used.

Jian-Wen Zhao et al. [18] presented a new approach to clustering learner access pattern based on the transitive closure. The learners' access pattern mining is an emerging technique that can be utilised to not only to divulge learners' access pattern but also advance the web page recommendation. This will provide benefits for the enhancement of E-learning systems.

V. THE PROPOSED RECOMMENDER SYSTEM BASED ON WEB USAGE MINING

The proposed system is based on the recommender model proposed by Xinjin Li et al. [7]. In this model it is shown the function of web usage mining. It can mine through storing the server log files and back-office database of E-learning platforms, such as the classification of the users and the gathering and classification of the page frequently visited by the learners.

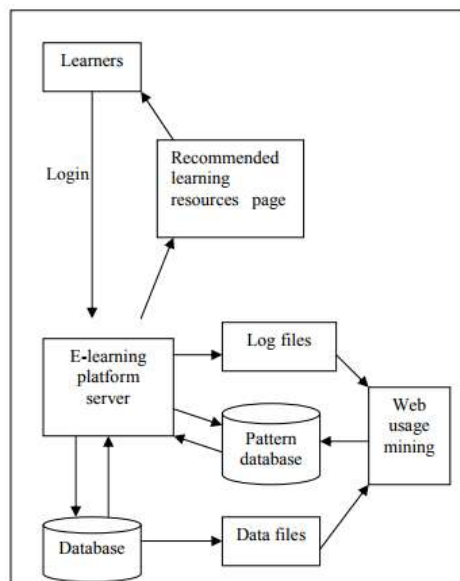


Figure3. Model of recommender system of learning resources based on web usage mining

The figure3 shows the model of recommender system of learning resources that is based on web usage mining. In this system when users login the E-learning platform, the server reads the classified information from the database according to the login information for the learners also it can classify the learners and read the recommended resources about the learners.

Meanwhile, it can read the relevant requesting data from the background database and ultimately a recommended

learning resource page visited by the learners will be established. This page includes not only the public learning content, but also the recommended information content preferred by the learners, which is convenient for the learners' navigation pattern.

The combination of web data mining and technologies results an elegant web agent technology. Intelligent web agent technology can provide navigation to the learners and results the improvement in the learning capabilities of learners and recommend them good resources of learning. The recommendation of learners' access pattern technique can be utilised to not only to reveal learners' access pattern but also advance the web page recommendation. This will also suggest benefits to the enrichment of E-learning systems.

VI. CONCLUSION

Web usage mining is appropriate and lively concerning field of research that capably used to improve E-learning system that can be advantage to develop intelligent and personalized environment for the learners. Web usage mining techniques within E-learning systems is a hastily growing trend. Web mining is the sequence of task used for mining or extracting useful information from the E-learning websites. Based on the analysis of the conception and methods of the web usage mining, this paper combines the web usage mining technology and E-learning platform to build a model with functions of recommending learning resources. Making a web usage mining analysis of the quality courses E-learning website so as to improve the structure of the E-learning website, classifying the learners, clustering the page so as to reach good results and provide specific learners with the recommendation of learning resources.

REFERENCES

- [1] Reema Sikka, Amita Dhankar and Chaavi Rana, "A Survey Paper on E-learning Recommender System", International Journal of Computer Applications (0975 – 888), Volume 47–No.9, June 2012.
- [2] Liu MinLi, Yu Chao, Mu Dan, "The Method of Dynamic Services in E-learning System", International Forum on Information Technology and Application, 2010,pp. 259-262.
- [3] Mohamed Koutheaïr Khribi, Mohamed Jemni and Olfa Nasraoui, "Automatic Recommendations for E-Learning Personalization Based on Web Usage Mining Techniques and Information Retrieval", Educational Technology & Society, 12 (4), 30–42.
- [4] Manasi Chakurkar, Prof. Deepa Adiga, "A Web Mining Approach for Personalized E-learning System", International

Journal of Advanced Computer Science and Applications (IJACSA), Vol. 5, No. 3, 2014, pp. 51-56.

- [5] Sunil, Prof. M.N. Doja, "A Review Paper on Identifying Students Interest in E-learning using Web Usage Mining", International Journal of Latest Trends in Engineering and Technology (IJLTET) Volume 8, Issue 1, pp.520-525, e-ISSN:2278-621X, doi:http://dx.doi.org/10.21172/1.81.067.
- [6] Sunil, Prof. M.N. Doja, "Recommender System Based On Web Usage Mining For Restructuring Of E-Learning Websites And Blogs", International Journal of Research in Engineering & Advanced Technology (IJREAT), Volume 5, Issue 1, Feb - Mar, 2017, ISSN: 2320 – 8791.
- [7] Xinjin Li, Sujing Zhang, "Application of web usage mining in E-learning platform", International Conference on E-business and E-government, IEEE 2010, doi:10.1109/ICEE.2010.353., pp. 1391-1394.
- [8] K. Sudheer Reddy, M. Kantha Reddy and V. Sitaramulu, "An effective Data Preprocessing method for Web Usage Mining", International Conference on Information Communication and Embedded Systems (ICICES), IEEE-2013. doi: 10.1109/ICICES.2013.6508197, pp. 7-10.
- [9] O. Zaiane, M. Xin, J. Han., "Discovering Web Access Patterns and Trends by applying OLAP and Data Mining Technology on Web Logs", In Advances in Digital Libraries, pages 19-29, Santa Barbara, CA, 1998.
- [10] Hadeel Alharbi, Ashoka Jayawardena and Paul Kwan, "Social recommender system for predicting the needs of students/instructors: Review and Proposed Framework", International Conference on Future Internet of Things and Cloud, IEEE 2014, doi: 10.1109/FiCloud.2014.93.
- [11] Harshit Patel and Prof. Pooja Jardosh, "A Survey on Recommendation System Using Web Usage Mining" International Institution for Technological Research and Development Volume 1, Issue 1, 2015.
- [12] Cleomar, B., Jr., & Oliveira, M., "Recommender systems in social networks. JISTEM - Journal of Information Systems and Technology Management", 2011, 8(3), 681–716.
- [13] Adomavicius, G., & Tuzhilin, A., "Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions", 2005, Knowledge and Data Engineering, IEEE Transactions, 17(6), 734–749.
- [14] Rong Shan and Zhibin Ren, "Research on Personalized Recommendation System in E-learning", 2nd International Conference on Education Technology and Computer (ICETC), IEEE-2010, Vol. 4, pp. 182-184, doi: 10.1109/ICETC.2010.5529705.
- [15] Chenle Wu, Jiyian Wu, Yibo Chen, Ming Xie, Xiaojun Guo, Weichen Xiong, "Personalized Learning Service Framework Based on Semantic Web Technologies", International Conference on Computer Science and Service System (CSSS), IEEE-2011, pp. 2396 - 2399, doi: 10.1109/CSSS.2011.5974494.
- [16] Rozita Jamili Oskouei, "Identifying Students' Behaviour Related to Internet Usage Pattern", International Conference

on Technology for Education, IEEE-2010, pp. 232-233, doi: 10.1109/T4E.2010.5550104.

- [17] Nandita Agarwal and Prof. Anand Jawdekar, "User-Based Approach For Finding Various Results in Web Usage Mining", Symposium on Colossal Data Analysis and Networking (CDAN), IEEE-2016, pp.1-6, doi:10.1109/CDAN.2016.7570867.
- [18] Jian-Wen Zhao, Shen-Ming Gu and Ling He, "A Novel Approach to Clustering Access Pattern in E-learning Environment", 2nd International Conference on Education Technology and Computer, IEEE-2010, Vol.1, pp. 393-397, doi:10.1109/ICETC.2010.5529224.

BIOGRAPHIES

Sunil is Assistant Professor in the Computer Engineering Section, Department of University Polytechnic, Faculty of Engineering and Technology, Jamia Millia Islamia (A Central University), New Delhi, India. He has more than 10 years of teaching experience. He is pursuing Ph.D. in



Computer Engineering from Department of Computer Engineering, Jamia Millia Islamia, New Delhi, India. He did M.Tech. in Computer Science from Department of Computer Science, Faculty of Management Studies and Information Technology, Jamia Hamdard, New Delhi, India, in 2012, B.Tech in Computer Science and Engineering from Northern India Engineering College, Lucknow, Uttar Pradesh, India, affiliated to Uttar Pradesh Technical University, Lucknow, in 2006 and Diploma in Computer Engineering from Government Polytechnic, Unnao, Uttar Pradesh, India, affiliated to Board of Technical Education, Uttar Pradesh, Lucknow, in 2003. His research interests are in the areas of Computer Networks and Data Mining. He has taught various papers including Computer Networks, Programming Languages, Database Management System and Operating System. He has published several research papers in Journal(s) and Conference(s).



Prof. M. N. Doja is Professor in the Department of Computer Engineering, Faculty of Engineering and Technology, Jamia Millia Islamia (A Central University), New Delhi, India. He has more than two decades of academic, research, training and administrative experiences in the field of Computer Engineering and Information Technology. He received his B.Sc. (Engg.), M.Tech. and Ph.D. degrees from B.I.T, I.I.T. Delhi and Jamia Millia Islamia, New Delhi, India respectively. His areas of research are Software Engineering, Networks, Security, Simulation, Operating System,

Data Mining and Soft Computing. Prof. Doja is an active researcher. He has more than 100 publications in referred journals and conferences of international and national repute. He has been referee for a number of journals in the area of Computer Engineering and Information Technology. Several students have been awarded Ph.D under his able guidance and several students are working under his supervision. He is also author of a number of books in the area of Information Technology. He has successfully completed several projects including AICTE sponsored projects. Prof. Doja is a member of Academic Council, Board of Research Studies and Board of Studies of a number of Universities including Ambedkar University Lucknow, NSIT New Delhi, A.M.U. Aligarh, Guru Gobind Singh Indraprastha University Delhi, NIT Jalandhar, Hamdard University, New Delhi etc. He has been a member of a number of committee for various Universities in various capacities. He has been expert member of various committee constituted by UGC and AICTE. He has also been associated with the affiliation work of various Universities like Guru Gobind Singh Indraprastha University, Delhi, M.D. University Rohtak, U.P. Technical University Lucknow and Utrakkhand Technical University, India.