

Research Journal of Pharmaceutical, Biological and Chemical Sciences

Recommendation of Web Pages by Utilizing Web Usage Knowledge.

Sterose S*.

Assistant Professor, Sathyabama University, Faculty of Computing, Tamil Nadu, India.

ABSTRACT

Recommendation of Web Pages by Utilizing Web Usage Knowledge, which is supported by three new knowledge representation models and a set of Web-page recommendation strategies. It is based on three models. One of the model is an ontology-based model which has the information about the domain knowledge of a website. The creation of this model is semi-automated, so that it reduces the development efforts from developers. Another model is a semantic network that consisting domain knowledge, whose creation can be fully automated. This model is fully automated feature so that it can be integrated easily into a Web-page recommendation process. The last model is a conceptual prediction model, which is a routing network of domain terms based on the frequently viewed Web-pages and represents the integrated Web usage and domain knowledge for supporting Web-page prediction.

Keywords: Web usage mining, Recommendation of web-pages, Domain ontology, Semantic network, Connectional prediction

**Corresponding author*

INTRODUCTION

Recommendation web-pages play a significant character in intellectual Web systems. It gives practical knowledge discovery from Web usage data and suitable knowledge representation for effective Web-page recommendation is essential and not easy. It proposes a novel method to provide enhanced Web-page recommendation through semantic enhancement by integrating the domain and Web usage knowledge of a website. It consist two models are projected to represent the domain knowledge. One of the models uses ontology to represent the domain knowledge. Another model uses automatically generated semantic network to represent domain terms, Web-pages and the association between them. One more new model, the conceptual prediction, in which a semantic network generated automatically. Web usage knowledge, which is the combination of domain knowledge and Web usage knowledge. A number of effective questions have been developed using these knowledge bases. Depends upon these questions, a set of related technique have been planned to generate Web-page candidates. The proposal results compared with the results attain from an advanced existing Web Usage Mining (WUM) method.

The web is Dynamic and very huge and diverse, which has the knowledge repository. The popularity of WWW is rapidly developing. Web consists various and non scalable data so extracting the useful knowledge from web is very tedious process. Web mining has been categorized in three categories: Web content mining, Web Structuring mining and Web usage mining. Web content mining is used to extract useful knowledge from web pages which is stored in repository. Web structuring mining focuses on analyzing the association between web pages using web structure. The Objective of Web Usage Mining is used to discover the useful information from web data. It is also used to increase the usability of the web information and to implement the technology on the web applications, for example, pre-fetching and caching. Forecasting the user's browsing behavior is one of web usage mining issue. In order to achieve the purpose, it is required to understand the customer's browsing behavior during analyzing the web data or web log files. Predicting the most possible user's next requirement is based on the previous similar behavior. There are many advantages to implement the prediction. Generally, three kinds of information have to be handled in a web site: content, structure and log data. Content data contains information about the application in a web page. Structure data and usage data is the organization and usage patterns of web sites.

RELATED WORK

The primary purpose of using a recommendation of web-pages for web usage and domain knowledge. Recommends that (i) frequently web accessed pages retrieve from web log and (ii) frequently viewed pages from the users retrieve from web log, stored on RDF file on html format. Based on the queries candidate web page has been generated.

A user is visiting a Web-page that is not in the discovered Web Access sequence, and then these approaches cannot offer any recommendations to this user. We refer to this problem as "New-page problem" for this we are using RDF (Resource Description Framework) storing the information on HTML format. Webpage recommendation based on Web usage knowledge, which is supported by three new knowledge representation models and a set of Web-page recommendation strategies. The first model is an ontology-based model that gives information about the domain knowledge of a website. The construction of this model is semi-automated so that the development efforts from developers can be reduced. The second model is a semantic network that represents domain knowledge, whose construction can be fully automated. This model can be easily incorporated into a Web-page recommendation process because of this fully automated. The new method to offer better Web-page recommendations through semantic enhancement by three new knowledge representation models. Two new models have been proposed for representation of domain knowledge of a website. One is an ontology-based model which can be semi-automatically constructed, namely DomainOntoWP, and the other is a semantic network of WebPages, which can be automatically constructed, namely TermNetWP. A conceptual prediction model is also proposed to integrate the Web usage and domain knowledge to form a weighted semantic network of frequently viewed terms, namely TermNavNet. A number of Web-page recommendation strategies have been proposed to predict next Web-page requests of users through querying the knowledge bases. The experimental results are promising and are indicative of the usefulness of the proposed models.

J. Borges, [1-3], proposed Higher-order Markov model for web usage mining. There are various problems associated with lower-order Markov model. The low accuracy is the major limitations of lower-order Markov model. As Higher-order Markov model suffer from the state space complexity, K-mean clustering technique has been used to reduce the state space complexity. The experimental result shows that the accuracy is improved by introducing the clustering technique in Markov model. Author has proposed the Higher-order Markov model with clustering technique to improve the effectiveness of Markov model and to reduce the state space complexity

Srivastava, J., Cooley, R., Deshpande, M., and Tan, P. N., [4], Internet usage continues to grow at a tremendous pace as an increasing number of activities are performed online. Computers on the Internet that host websites, the web servers, are capable of collecting information about website usage. Given the popularity of the Internet the volume of such data is enormous. To store all this information in web as feasible, the enhance in secondary storage size and reduce in their cost. This information is a profitable repository for mining and discovering interesting patterns. On the Researchers point of view using the extracted patterns to predict the next user request during an online session with the website. Similar systems are called Recommended Systems and to predict user requirements using tools. For increasing overall usability of the website use predictive ability has application in areas like pre-fetching of pages.

Cooley, R., Mobasher, B., and Srivastava, J., [5-7], Web Mining is described as the application of data mining techniques to extract patterns from usage information. Various data mining methods have been used to generate models of usage patterns. For designing Models based on algorithms such as clustering algorithms, sequential analysis, association rules. The inputs to the system are the web-usage logs from the web-servers and the concept hierarchy of the website content.

B. Berendt, A. Hotho, G. Stumme, [8-11], Parts of this substantially revised and extended survey were presented at the First International Semantic Web Conference Application events are defined with respect to the application domain and the site, a significant task that amounts to a detailed formalization of the site's business/application model. To attain the concept, Web data (usage, content, structure) are represented by using developing model of representation, ontologies. This representation having the difference between Semantic Web and Web Mining areas, to create a research areas such as privacy and security, e-activities, health care, bioinformatics and knowledge management and information retrieval. It suggested that great opportunities in finance, business, marketing, commerce, education, research and development.

Sebastian A. Rios, [12] gave a concept based approach for offline web site enhancement. Because obtain useful information from sites based on list of keyword searching techniques is somewhat difficult. Before this approach semantic WUM process used a concept-based approach in mining process of semantic web. The technique discussed in this paper can enhance the contents and structure of a web site offline. He compared his work with four other WUM methods. After that a quality of enhancement was calculated based on survey to 100 visitors of the site, giving the effective result. This web mining process generates more semantically fit result according to the user's query. This approach is based on user's interest and correlation measurement, so by the result analysis it was proved that this approach obtain closer results for users according to their browsing preferences. Resultant performance is useful for 74% visitors which is better than the other methods (previously existed methods that were discussed in this paper) which are giving less than 50%. Classical WUM process took approximate 11 hours to complete the task whereas this approach took only 15 minutes for same task. Developed model is a very powerful method for solving many real time problems of classical WUM processes.

L. Wei and S. Lei [13], There are studies that aim to generate patterns in terms of semantic information. (B.Y. Zhou, S.C. Hui, This algorithm and not producing candidate frequent patterns. Accordingly, WAP-mine algorithm is an order of magnitude faster than Apriori algorithm (B.Y. Zhou, S.C. Hui, & A.C.M. Fong (2004)) put forward by Agrawal at earlier stage.

The previous work is Web Usage Mining (WUM). This type of web mining allows for the collection of web access information for web pages. This information is often gathered automatically into access logs. It is recommends only accessed pages for the users. Information is stored on database. Its recommends only accessed web page. User entering new keyword its recommend the database, the page should not available.

PROPOSED WORK

Proposed system is web page recommendation based on the web usage and domain knowledge using three new knowledge representation models. Ontology construct the web-pages on sever log using web usage mining technique, recommended domain terms are stored on RDF file when user giving queries candidate web page will be generated.

In semantic domain terms are collocations of terms which are determined by the co-occurrence relations of terms in web-pages. Relations between terms and web pages. Conceptional prediction are future work based on web-page recommends. Based on web usage knowledge that can be discovered from web log files using server log. It recovers the “New page problem”.

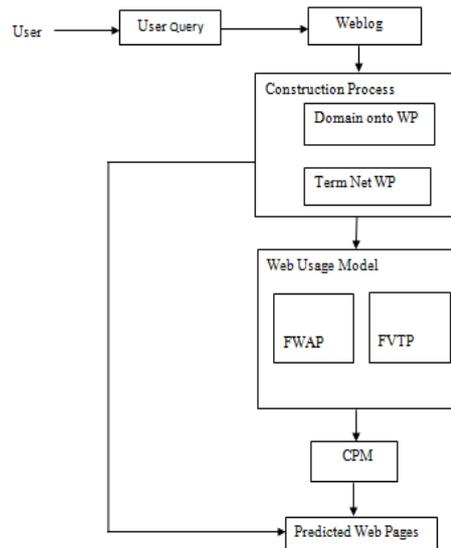


Fig: 3.1 Recommend Webpage

The framework of Web-page recommendation, Web logs stored user session based on the user query on a daily basis. The user session contains the web page navigation activities. Here to check the content or web page based on the user queries within system. The system has drive had the files such as documents, text files, excel sheets and xml. Based on the type of file it easily finds out the file and display to the user. In Web logs the files are separated based on the type of file.

Ontology’s are often implemented in a logic-based language, such as OWL/RDF, to become understandable to software agents or software systems. Therefore, ontology based knowledge representation allows sharing and interchanging semantic information among Web systems. Ontological representation of discovered knowledge from different sources can be easily integrated to support Web-page recommendation effectively.

Semantic domain terms are collocations of terms which are determined by the co-occurrence relations of terms in Web-page titles the associations between terms and Web-pages. In addition, the domain terms and co occurrence relations are weighted to provide a rough indication of how much these terms are associated with each other semantically. Based on the relations between the terms and Web-pages, we can infer how closely the Web-pages are semantically related to each other.

Recommend the web pages based on the Web usage knowledge that can be discovered from Web log files using a Web usage mining technique. Discover the Web usage knowledge, which is in the Form of frequent Web access. (**FWAP**) and we integrate FWAP with frequently viewed term patterns (**FVTP**), predict the efficiently and recommend the URLs.



Fig 3.1 Crawl the webpage from Internet Information Server

If, enter login and password click the submit button ontology page will be display here, we are crawl the web page .

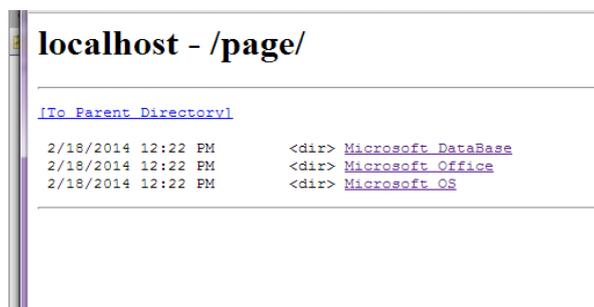


Fig 3.2 Localhost webpage of domain creation .

Here we are creating three domain models, if click any domain in that it will display the directory the particular domain page

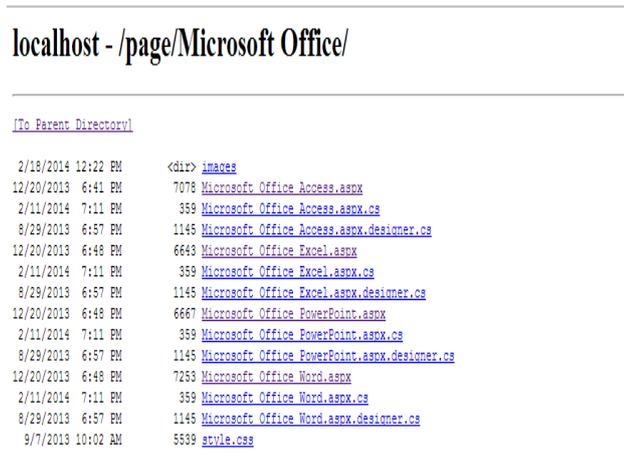


Fig 3.3 Local host webpage of sub domain creation

Here, to create sub domain of the domain page, click any option in that it will display the information about sub domain of web page, using URL only we are crawling the web page. It's stored on the RDF file it display the message is "Domain created" otherwise Already domain created"



Fig 3.4 Searching webpage of recommends process

Here we entering given domain page of related keywords it will display click any option on the related pages, and then it displays the information about required pages.

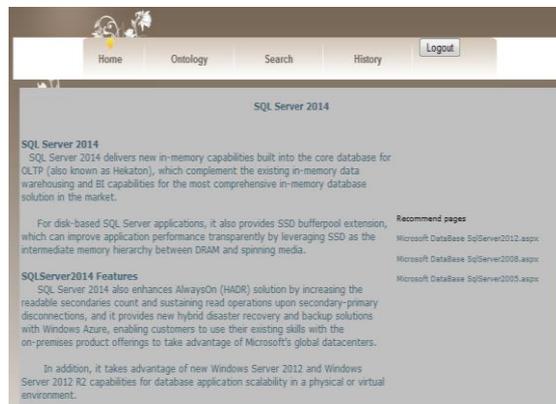


Fig 3.5 Recommend page for related webpage

Recommend the web pages based on the Web usage knowledge that can be discovered from Web log files using a Web usage mining technique. Discover the Web usage knowledge, which is in the Form of frequent Web access. (**FWAP**) and we integrate FWAP with frequently viewed term patterns (**FVTP**), predict the efficiently and recommend the URLs.

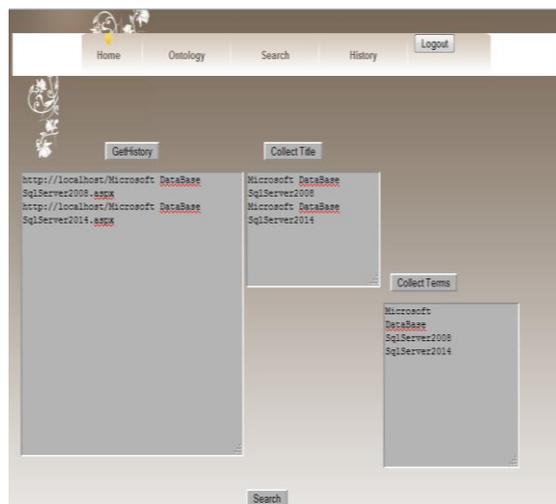


Fig 3.6 Collecting terms from Ontology construction

If you are clicking the existing user option it will display about the user visited page of the URL and keyword terms.

After that click the search option give some related keyword of domain then some related option will display, in that select particular option it display then information about that pages.

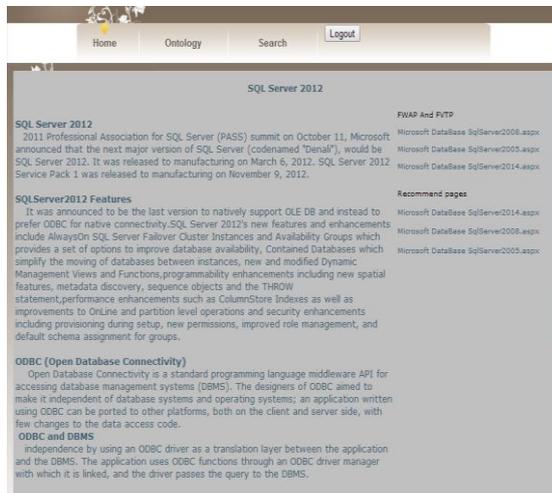


Fig 3.7 Recommends on Frequently accessed and viewed page

Here it displaying recommended pages and frequently visited pages and frequently web accessed pages.

CONCLUSION AND FUTURE WORK

Web usage knowledge is an emerging technology that can help in producing personalized a Web-based system. A number of systems following this approach have been developed, using methods and techniques from Web usage mining, in order to realize a variety of Web page functions. Semantic and ontologies for improving performance of recommender system. Web-pages for effective Web-page recommendations by using semantic knowledge of both Web-pages and Web usage Future work is to find out the better support for frequent web navigation patterns generated by using web usage mining techniques. The web log records to discover user access patterns of web pages.

REFERENCES

- [1] Borges J and Levene M. "A clustering-based approach for modeling user navigation with increased accuracy", Proc. Second Int'l Workshop Knowledge Discovery from Data Streams, Oct. 2005.
- [2] Borges J and Levene M. "Generating dynamic higher-order Markov models in web usage mining," Proc. Ninth European Conf. Principles and Practice of Knowledge Discovery in Databases (PKDD), eds. A. Jorge, L. Torgo, P. Brazdil, R. Camacho, and J. Gama, Oct. 2005;pp. 34–45.
- [3] Borges J and Levene M. "Data Mining of User Navigation Patterns," Web Usage Analysis and User Profiling, eds. B. Masand and M. Spiliopoulou, LNAI 1836, Springer, 2000;pp. 92–111.
- [4] Srivastava J, Cooley R, Deshpande M and Tan PN. Web usage mining: Discovery and applications of usage patterns from web data. SIGKDD Explorations 2000;1(2):12–23.
- [5] Cooley R, Mobasher B, and Srivastava J. Data preparation for mining World Wide Web browsing patterns. Journal of Knowledge and Information Systems 1999;1:1.
- [6] Cooley R, Srivastava J and Mobasher B. Web Mining: Information and pattern discovery on the world-wide web. 9th IEEE International Conference on Tools with Artificial Intelligence (November 1997).
- [7] Mobasher B, Cooley R and Srivastava J. Creating adaptive web sites through usage-based clustering of URLs. Knowledge and Data Engineering workshop, 1999. International Conference on Data Mining (ICDM'02) (Maebashi City, Japan, December 2002).
- [8] Berendt B, Hotho A, Stumme G. Towards semantic web mining, in: I. Horrocks, J.A. Hendler (Eds.), The Semantic Web—ISWC 2002. First International Semantic Web Conference, Proceedings, LNCS. Springer, 2002; 2342:264–278.

- [9] Berendt B, Hotho A, Stumme G. Usage mining for and on the semantic web, in: H. Kargupta, A. Joshi, K. Sivakumar, Y. Yesha (Eds.), *Data Mining: Next Generation Challenges and Future Directions*, AAAI/MIT Press, Menlo Park, CA, 2003.
- [10] Berendt B, Hotho A, Stumme G. *Towards Semantic Web Mining Proceedings of the First International Semantic Web Conference on the Semantic Web* Springer-Verlag, 2002; pp. 264-27.
- [11] Berendt B, Hotho A, Mladenic D, Someren M, Spiliopoulou M and Stumme G. *G.A Roadmap for Web Mining: From Web to Semantic Web* Berendt, B., Hotho, A., MladeniÄ? D., Someren, M., Spiliopoulou, M. & Stumme, G. (ed.) *Web Mining: From Web to Semantic Web* Springer Berlin Heidelberg, 2004;3209:1-22.
- [12] Sebastian A Rios and Juan D Velasquez. "Semantic Web Usage Mining by a Concept-based Approach for Off-line Web Site Enhancements," 2008 IEEE.
- [13] Wei L and Lei S. "Integrated Recommender Systems Based on Ontology and Usage Mining", *Active Media Technologies*, 5820, Springer-Verlag, Berlin Heidelberg, 2009; pp. 114-1.
- [14] Zhou BY, Hui SC and Fong ACM. CS-mine: An Efficient WAP-tree Mining for Web Access Patterns. In *Proceedings of the 6th Asia Pacific Web Conference (APWeb'04)*, Hangzhou, China, Lecture Notes in Computer Science 3007, Springer, 2004; pp. 523-532.