# WEB USAGE AND DOM GNOSTICISM USING WEB PAGE RECOMMENDATION

**S.Sindhuja[1], C.Hariram[2]**

*[1]M.Tech IT PG Scholar, [2]Assistant Professor, Department of Information Technology,*
*Dr. Sivanthi Aditanar College of Engineering Tiruchendur (India)*

## ABSTRACT

*A new framework for a semantic-enhanced Web-page recommender (WPR) system, and a suite of enabling techniques which include semantic network models of dom Gnosticism and Web practiceerudition is described in this paper. On web different kind of web recommendation are made available to user every day that includes Image, Video, Audio, query suggestion and web page. Conceptual and structural characteristics of a website is important for the quality of recommendations provided by a recommendation system. Resources like Google Directory, Yahoo! Directory and web-content management systems attempt to organize content conceptually. Semantic Web Mining aims at combining the two fast-developing research areas Semantic Web and Web Mining. In this paper, we discuss the interplay of the Semantic Web with Web Mining, with a specific focus on usage mining.There are two new models that are proposed to represent the dom Gnosticism. Ontology is used to represent the dom knowledge. Then the semantic networks are automatically generated for the semantic network to represent domain terms, Web-pages and the relations between them. Another new model, the conceptual prediction model, is proposed to automatically generate a semantic network of the semantic Web usage knowledge, which is the integration of dom Gnosticism and Web usage knowledge. The experimental results demonstrate that the proposed method produces significantly higher performance than the WUM method.*

***Index Terms: Web Usage Mining, Web-Page Recommendation, Domain Ontology, Semantic Network, Knowledge Representation.***

## I. INTRODUCTION

This paper develops a prototype of the new semantic-enhanced Web-page recommender system (SWRS) utilizing these models to leverage recommendations produced by a community of users to deliver recommendations to an active user. Firstly, the system needs to learn users' Web usage experience which is Web minutes of given websites. The Web minutes are the records of users' Web browsing behaviours daily, that is Web usage data. By mining Web usage data, useful knowledge can be discovered and represented in the models, i.e. Domain On toWP or Term Net WP, WPN avNet, and Term Nav Net, which can facilitate making Web-page recommendations. Web page recommendations are becoming very popular, and are shown as links to related web page, related image, or popular pages at websites. When user sends request to web server, session is created for the user. During session when user browses a website the list of page that user visits is stored as a session data.

- Web content mining (WCM) is used to mine Web content, such as HTML or XML documents.
- Web structure mining (WSM) focuses on Web structure, such as hyperlinks on Web-pages.

- Web usage mining (WUM) is applied to Web usage data, such as Web minutes or clickstreams, from a website.

In summary, making recommendations is the main application of WUM in recommender systems as WUM can obtain the actual user behaviour rather than the behaviour expected from the Web design. And pattern discovery methods play an important role in mining user behaviour sequences.An ontology may include individuals (instances), classes, attributes, relations, restrictions, rules, and axioms. Based on these components, we can build an object-oriented model for an application domain and use this model for sharing and reusing domain information on the Web. Such an ontology model allows human- and machine-understandable content and human-machine interaction.

This paper presents a different method to provide better Web-page recommendation based on Web usage and dom Gnosticism, which is supported by three new knowledge representation models and a set of Web-page recommendation strategies. Ontology-based model is used to represents the dom Gnosticism of a website. The construction of this model is semi-automated so that the development efforts from developers can be reduced. Semantic network is used to represents dom Gnosticism, whose construction can be fully automated. This model can be easily incorporated into a Web-page recommendation process because of this fully automated feature. The third model is a conceptual prediction model, which is a navigation network of domain terms based on the frequently viewed Web-pages and represents the integrated Web usage and dom Gnosticism for supporting Web-page prediction. The construction of this model can be fully automated. To a great extent, this new method has automated the knowledge base construction and alleviated the new-page problem. This method yields better performance compared with the existing Web usage based Web-page recommendation systems.

## II. RELATED WORK

Web Mining techniques arose as a tool for helping managers and web masters in the enhancements task. However, many tools provide not very useful results to perform off-line web site enhancements. A big problem is how the people in charge of a Web site modifies the site based on a web mining tools' results. The issue of analyzing huge amount of subjective information, hundreds of features, and mil-lions of user sessions; to extract useful information to enhance a web site is not straightforward. Moreover, most tools or processes use terms as features, making even harder to see a good solution in a reasonable time. Therefore, the development of new methods which may lead to good results in relatively short processing times is quite important.

This papers' idea is the development of a new methodology to perform better Web Usage Mining (WUM), based on the introduction of concepts into the mining process. This way we can take advantage of the organizations' business knowledge and use it in the mining task. This process is called concept-based web usage mining. A strong difficulty occurs to evaluate the proposed technique. Usually papers in data mining field doesn't provide any means of comparison with other techniques. In these case, the problem of evaluation is bigger since we need to evaluate a mining process. Based on [1] we developed an evaluation schema to be applied in a real web site.

On the other hand, by mapping Web-pages to domain concepts in a particular semantic model, the recommender system can reason what Web-pages are about, and then make more accurate Web-page recommendations [3, 8]. Alternatively, since Web access sequences can be converted into sequences of ontology instances, Web-page recommendation can be made by ontology reasoning [2, 9]. In these studies, the Web usage mining algorithms find the frequent navigation paths in terms of ontology instances rather than normal Web-page sequences.

Generally, ontology has helped to organize knowledge bases systematically and allows systems to operate effectively.

## III. DOMAIN ONTOLOGY MODEL

Domain ontology can be obtained by manual or automatic construction approaches. Depending on the domain of interest in the system, we can reuse some existing ontologies or build a new ontology, and then integrate it with Web mining. Web minutes in a Web personalization system. Ontology is a knowledge representation technology whose implementation can be machine-understandable using the ontology language, such as OWL. Ontology defines the concepts and their associations in an application domain. In the context of Web-page recommendation, it is necessary to have an ontology that expresses the meaning of Web-pages for better understanding Web usage patterns and discovering frequently viewed domain terms for supporting more effective Web-page recommendations.

The Web usage knowledge can be discovered from Web usage data through unsupervised learning processes, such as sequential pattern mining techniques, but without the semantics of Web-pages, the discovered knowledge are limited in supporting Web-page recommendation, such as no alleviation to the "new page" problem. A domain ontology is really useful to enhance a Web-page recommendation process by adding semantics to Web-pages, but how to build an effective domain ontology for Web-page recommendations is always a big challenge. The study presented in this chapter builds a domain ontology of Web-pages of a website that can be used to interpret the semantics of Web-pages. This chapter proposes a domain ontology model that represents the domain concepts, Web-pages, and the relations among them for a given website to support semantic-enhanced Web-page recommendation and also presents a novel method to build such a domain ontology for a website.

In the context of Web-page recommendation, we build the domain ontology of Web-pages of a given website based on the visited Web-pages to represent the domain concepts (general domain terms), the relationships between the concepts with constraints, the instances of concepts (specific domain terms), Web-pages, and the links between Web-pages and specific domain terms.

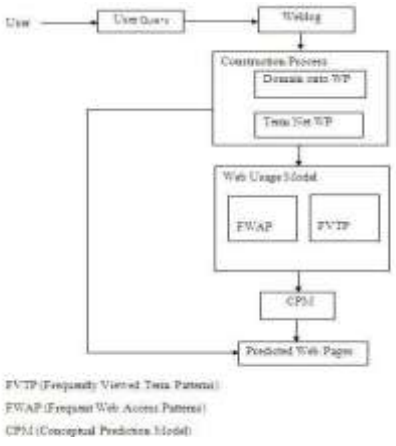


**Fig 1 Architectural Design**

**Definition 3.1 (Domain ontology model of Web-pages - DomainOntoWP)** A domain ontology structure of a website is defined as a four-tuples: $O_{man}$:= < C, D, $P_{MAN}$, A>, where C represents terms extracted from the Web-page titles within the given website, D represents the Web-pages of the website, $P_{MAN}$ represents properties defined in the ontology, and A represents axioms, such as, an instantiation axiom assigning an

instance to a class, an assertion axiom assigning two instances by means of a property, a domain axiom for a property and a class, and a range axiom for a property and a class. In details, C, D, and $P_{MAN}$ are further divided into sets:

C= C ∪$T_{man}$ comprises a set of general domain terms (concepts) C, and a set of specific domain terms (instances of the concepts) $T_{man}$, D= SemPage∪ D comprises class SemPage which represents Web-page instances, and a set of Web-pages D, $P_{MAN}$ =$R_{man}$∪Aman comprises a set $R_{man}$ of the relations between terms ($R_c$) and the relations between terms and Web-pages ($R_p$), and a set of attributes $A_{man}$defined in the ontology. In particular, $R_c$ will be specified depending on the application domain. Rp= haspage⊔isAbout , where the 'hasPage' relation states that a domain term may have some Web-pages, and the 'isAbout' relation is the inverse of the 'hasPage' relation. That means each domain concept class has the 'hasPage' object property referring to class SemPage, and class SemPagehas the 'isAbout' object property referring to the domain concept classes.

## IV. SEMANTIC DOMAIN TERM GENERATION

One of the big challenges that these approaches are facing is the semantic dom Gnosticism acquisition and representation.Kearny et al. [10] also investigate how Web usage data may be combined with se-mantic dom Gnosticism to provide a deeper understanding of user behavior. The Semantic Web is based on a vision of Tim Berners-Lee, the inventor of the WWW. The great success of the cur-rent WWW leads to a new challenge:

A huge amount of data is interpretable by humans only; machine support is limited. Berners-Lee suggests to enrich the Web by machine-processable information which supports the user in his tasks. For instance, today's search engines are already quite powerful, but still too of-ten return excessively large or inadequate lists of hits. Machine-processable information can point the search engine to the rele-vant pages and can thus improve both precision and recall.

## V. CONCEPTUAL PREDICTION MODEL

With the given dataset, meaningful terms are extracted by removing stop words, e.g., "the", "a", and "for", or invalid words from the Web-page titles. For example, terms which are extracted from the sample dataset in Table 4-2 are "MS", "Access", "Support", "SQL Server", "Office", "News", "PowerPoint", "Project", and "Excel". It is possible for some extracted terms to share same features, so they are better to be the instances of a concept rather than standalone concepts. As the scope of the domain has been stated in the requirements analysis, the considered domain concepts of the MS website are Manufacturer, Application, Product, Category, Solution, Support, News, Misc, and SemPage. In which, the concept SemPagerefers to Web-pages, and other concepts refer to terms used in the MS website.

## VI. CONCLUSION AND FUTURE ENHANCEMENT

This paper aims to address the challenges and/or problems in developing Web-page recommender systems, such as the "new page" problem, the challenges of manual knowledge construction, and the "heterogeneous knowledge bases" problem. We discussed how Semantic Web Mining can improve the results of Web Mining by exploiting the new semantic structures in the Web; and how the construction of the Semantic Web can make use of Web Mining techniques. The example provided in the last section shows the potential benefits of further research in this integration attempt.

Three important directions for further interdisciplinary cooperation between mining and application experts in Semantic Web Usage Mining have been identified: (1) the development of ontologies of complex behaviour, (2) the deployment of these ontologies in Semantic Web description and mining tools, and (3) continued research into methods and tools that allow the integration of both experts' and users' background knowledge into the mining cycle.

Finally, we concluded with the description of the recommendation engine's operation and presented an algorithm for making effective recommendations. Compared with one of the most advanced Web usage mining method, i.e. PLWAP-Mine, the proposed method can substantially enhance the performance of Web-page recommendation in terms of precision and satisfaction. More importantly, this method is able to alleviate the "new-page" problem mentioned in the introduction because it based on not only the Web usage knowledge, but also the semantics of Web-pages.

Future work will focus on further experiments with different combinations of the system's functionalities, further contextualization possibilities from the Semantic Web Mining area, and an evaluation of the proposed approach with respect to learning support and to open-corpus learning.

**REFERENCES**

[1] A. Geyer-Schulz and M. Hahsler.Comparing Two Reco-mender Algorithms with the Help of Recommendations by Peers, volume 2703 of Lecture Notes in Computer Science. Sept. 2003.

[2] S. Salin and P. Senkul, "Using Semantic Information for Web Usage Mining based Recommendation," in 24th International Symposium on Computer and Information Sciences, 2009., 2009, pp. 236-241.

[3] A. Bose, K. Beemanapalli, J. Srivastava, and S. Sahar, "Incorporating Concept Hierarchies into Usage Mining Based Recommendations," in Proceedings of the 8th Knowledge discovery on the web international conference on Advances in web mining and web usage analysisPhiladelphia, PA, USA: Springer-Verlag, 2007, pp. 110-126.

[4] ModrajBhavsar Int. Journal of Engineering Research and Applications "Web Page Recommendation Using Web Mining", ISSN : 2248-9622, Vol. 4, Issue 7( Version 2), July 2014, pp.201-206.

[5] ThiThanh Sang Nguyen "Semantic-enhanced web-page recommender systems" December, 2012.

[6] Bettina Berendt, Andreas Hotho, and GerdStumme "Semantic Web Mining and the Representation, Analysis, and Evolution of Web Space".

[7] "Mining Web Log Sequential Patterns with Position Coded Pre-Order Linked WAP-Tree*", C.I. EZEIFE, YI LU School of Computer Science, University of Windsor, Windsor, Ontario, Canada, Received July 2002; Revised June 2004.

[8] N. R. Mabroukeh and C. I. Ezeife, "Semantic-Rich Markov Models for Web Prefetching," in 2009 IEEE International Conference on Data Mining WorkshopsMiami, Florida, USA, 2009, pp. 465-470.

[9] M. O'Mahony, N. Hurley, N. Kushmerick, and G. Silvestre, "Collaborative recommendation: A robustness analysis," ACM Transactions on Internet Technology, vol. 4, pp. 344-377, 2004.

[10] Kearney, P., Anand, S.S., Shapcott, M.: Employing a domain ontology to gain insights into user behaviour. In: Proceedings of the 3rd Workshop on Intelligent Techniques for Web Personalization, at IJCAI 2005, Edinburgh, Scotland (August 2005).

[11] Sebastian A. Rios and Juan D. Velasquez "Semantic Web Usage Mining by a Concept-based Approach for Off-line Web Site Enhancements", 2008 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology.

[12] GerdStumme, Andreas Hotho, Bettina Berendt "Semantic Web Mining: State of the art and future directions", Received 20 May 2005; accepted 2 February 2006. www.sciencedirect.com

**ACKNOWLEDGEMENT**

My Sincere thanks to Mr.C.Hariram, Assistant Professor/IT for his innovative ideas and proper guidance for the successful completion of my project.

## BIOGRAPHY

| | |
|---|---|
|  | **S.Sindhuja** has completed the B.E degree in Computer Science in Syed Ammal Engineering College, Ramanathapuram in 2013. Currently she is pursuing M.Tech degree in Information Technology in Dr. Sivanthi Aditanar College of Engineering, Tiruchendur. Her research ares of interest include Algorithm design, Analysis and Implementation, Data Mining and Knowledge Engineering. |
|  | **C.Hariram** has completed B.Tech in Information Technology in Sethu Institute of Technology, Madurai in 2006 and M.E in Computer Science in Manonmaniam Sundaranar University in 2010. He is now working as an Assistant Professor for 6 year as in Dr. Sivanthi Aditanar College of Engineering, Tiruchendur. His research areas are Data Mining and Databases. |