Elysium Journal
Engineering Research & Management

P-ISSN: 2347-4408
E-ISSN: 2347-4734

RECOMMENDATION OF SEMANTIC NETWORK OF WEB PAGES USING ONTOLOGY BASED MODEL

Mr. V.Prasanna[1] , Ms. B.Kamala[2]

Post Graduate Student, Department of Computer Applications, Sri Sairam Engineering College, West Tambaram, Chennai, Tamilnadu, India.

Assistant Professor, Department of Computer Applications, Sri Sairam Engineering College, West Tambaram, Chennai, Tamilnadu, India.

Abstract- Web-page recommendation is one of the unleashing features in intelligent Web systems. It is a crucial task to recommend a web page without any satisfactory knowledge. The proposed semantic enhancement model integrates the domain and web usage knowledge for a website. There are three models are used to represent the domain knowledge such as ontology based model, generation model and Conceptual prediction model. First, the ontology based model represents the domain knowledge by using the ontology. Then, generation model in automatic generated network represents the domain terms like web pages and relation between them. Conceptual prediction model generates a semantic network contains domain knowledge and web usage knowledge. The queries developed to search the knowledge bases. A set of recommendation strategies generates web page candidates on the basis of queries. The comparison between the recommendation results and results from an advanced Web Usage Mining (WUM) method confirms that proposed method provides the higher performance higher performance than the WUM method.

Key Terms: Web usage mining, Web page recommendation, semantic network, knowledge representation.

1. INTRODUCTION

Webpage recommendation is popular division in intelligent web systems. The recommendation provides the links to stories, books and most viewed pages. A sequence of many web pages is generated from starting to closing session of browsing of website by the user. This sequence is organized into a Web session S = d, d2. . dk, where di (i = [1 . . . k]) is the page ID of the i th visited Web-page by the user. The objective of a Web-page recommender system is to effectively predict the Web-page or pages that will be visited from a given Web-page of a website.

There are a number of issues in developing an effective Web-page recommender system, such as how to effectively learn from available historical data and discover useful knowledge of the domain and Web-page navigation pat- terns, how to model and use the discovered knowledge, and how to make effective Web-page recommendations based on the discovered knowledge.

A great deal of research has been devoted to resolve these issues over the past decade. It has been reported that the approaches based on tree structures and probabilistic models can efficiently represent Web Access Sequences (WAS) in [1] queries the web usage data. The transition links between the web pages built by using the training sets. The query process based on WAS from web usage data. Hence, predicted pages are limited within the WAS.

The remainder of this paper is described as follows. Section 2 describes the related work done in web page recommendation, section 3 describes the system architecture of web page recommendation model, section 4 carries out the proposed framework, section 5 & 6 deals with conclusion and references respectively.

2. RELATED WORK

It has been roughly classified the research work related to Web-page recommendation as below.

In applying sequence learning models to Web-page recommendation, association rules and probabilistic models have been commonly used. Some models, such as sequential modeling, have shown their significant effectiveness in recommendation generation

In order to model the transitions between different web pages in Web sessions, Markov models and tree-based structures are strong candidates. Some surveys have shown that tree based algorithms, particularly Pre Order Linked WAP-Tree Mining (PLWAP-Mine for short), are outstanding in supporting Web-page recommendation, compared with other sequence mining algorithms[2][3][4]. Furthermore, the integration of PLWAP-Mine and the higher-order Markov model can significantly enhance mining performance.

3. SYSTEM ARCHITECTURE

In the context of web-page recommendation, the input data is web logs that record user sessions on a daily basis. The user sessions include information about user's web-page navigation activities. Each web-page has a title, which contains the keywords that embrace the semantics of the web-page. Based on these facts, aim is to discover domain knowledge from the titles of visited web-pages at a web-site and represent the discovered knowledge in domain ontology to support effective web-page recommendation [5].

E-ISSN: 2347-4734
P-ISSN: 2347-4408

Domain ontology is defined as a conceptual model that specifies the terms and relationships between them explicitly and formally, which in turn represent the domain knowledge for a specific domain. The three main components are listed as follows.

1) Domain terms (concepts),
2) Relationships between the terms (concepts), and
3) Features of the terms and relationships.

Ontology's are often implemented in a logic-based language, such as OWL/RDF, to become understandable to software agents or software systems [6]. Therefore, ontology-based knowledge representation allows sharing and inter-changing semantic information among Web systems over the Internet. It also enables the reuse of the domain knowledge, and reasoning the semantics of Web-pages from the existing facts. Furthermore, ontological representation of discovered knowledge from different sources can be easily integrated to support Web-page recommendation effectively. The following figure shows the web page recommendation model.

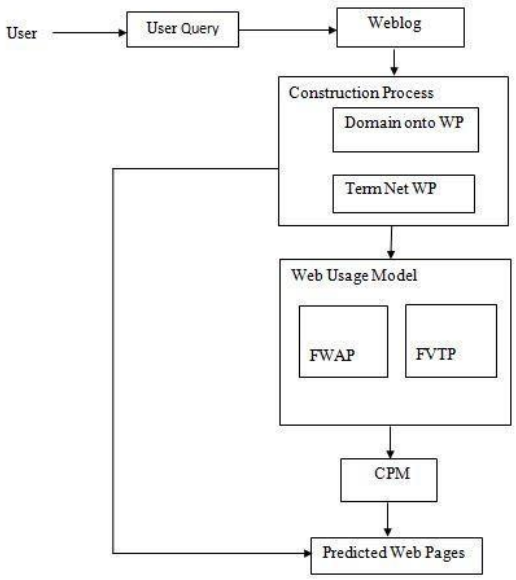

Fig.1 Webpage recommendation model

There are three steps in the procedure for constructing the domain ontology.

1) Collect the terms
2) Define the concepts
3) Define taxonomic and non-taxonomic relationships

3.1 COLLECT THE TERMS

In order to collect the terms, user has to:

- Collect the Web log file from the Web server of the website for a period of time (at least seven days),
- Run a pre-processing unit to analyze the Web log file and produce a list of URLs of Web-pages that were accessed by users,
- Run a software agent to crawl all the web-pages in the URL list to extract the titles, and
- Apply a mechanism to extract terms from the retrieved titles.

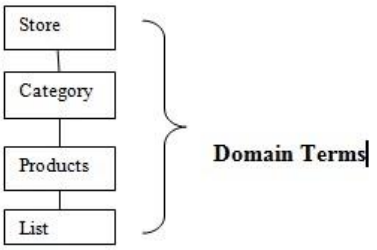

Fig.2 Illustration of Domain Terms

3.2 DEFINE THE CONCEPTS

It is possible for some extracted terms to share the same features, so it is better for them to be instances of a concept, rather than standalone concepts. In this step, the domain concepts will be defined for the given website based on the extracted terms. In this paper, MS web-site is presented as an example. This website focuses on the application software, such as MS Office, Windows Operating System, and Database.

Table.1 Sample MS Web Dataset

Page	Title	Path
d1	MS Word	/msword
d2	MS Word support	/mswordsupport
d3	MS Access	/msaccess
d4	MS Access support	/msaccesssupport
d5	MS Education	/mseducation
d6	Visual FoxPro	/visual FoxPro

3.3 DEFINE TAXONOMIC AND NON-TAXONOMIC RELATIONSHIPS

To develop the taxonomic relationships there are three possible approaches such as,

- Top-down development process starts from the general concepts in the domain and then identifies the subsequent specialization of the general concepts[7][8].
- Bottom-up development process starts from the most specific concepts as the leave nodes in the concept hierarchical structure/tree structure, then groups these most specific concepts into more general concepts.
- Combination of the top-down and bottom-top approaches termed as hybrid development process.

E-ISSN: 2347-4734
P-ISSN: 2347-4408

At first, the core concepts behind the domain are identified. Then, the concepts are generalized and specialized.

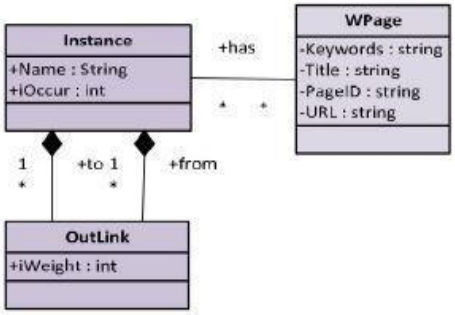

Fig.3 Taxonomic and Non Taxonomic relationship

4. PROPOSED FRAMEWORK

Web page recommendation is proposed based on the web usage and domain knowledge using three new knowledge representation models. The three models are ontology-based model, semantic domain terms generation model, and conceptual prediction model.

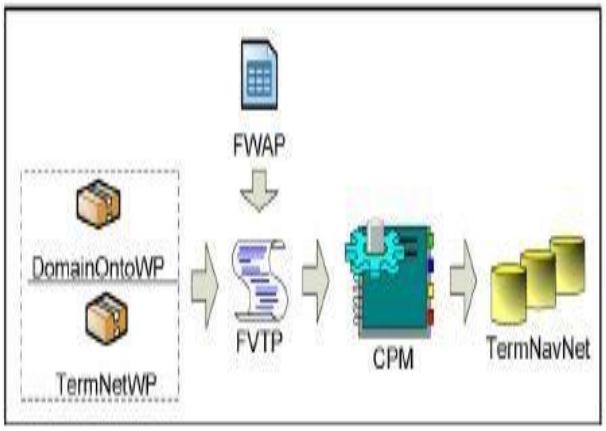

Fig.4 Term prediction process

4.1 ONTOLOGY-BASED MODEL

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.

4.2 SEMANTIC DOMAIN TERMS GENERATION MODEL

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[9]. Based on the relations between the terms and Web-pages, it can be inferred how closely the Web-pages are semantically related to each other.

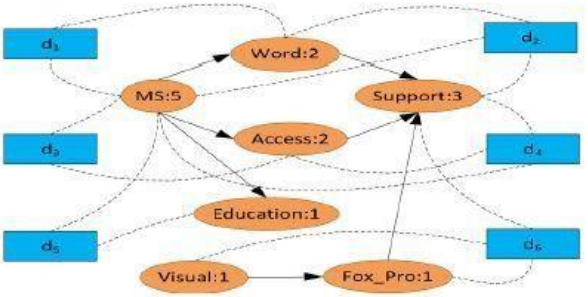

Fig.5 Illustration of terms

4.3 CONCEPTUAL PREDICTION MODEL

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.

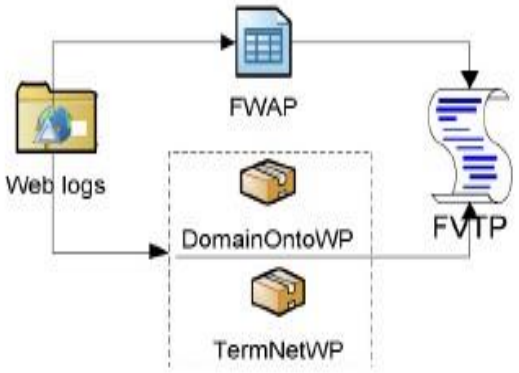

Fig.6 FVTP Discovery

5. IMPLEMENTATION

Implementation phase of web page recommendation is as follows.

Elysium Journal
Engineering Research & Management
E-ISSN: 2347-4734
P-ISSN: 2347-4408

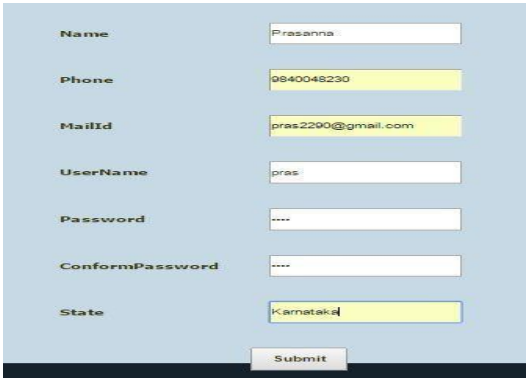

Fig.7 Registration page for an user

Registration page will carry out the fields as above.

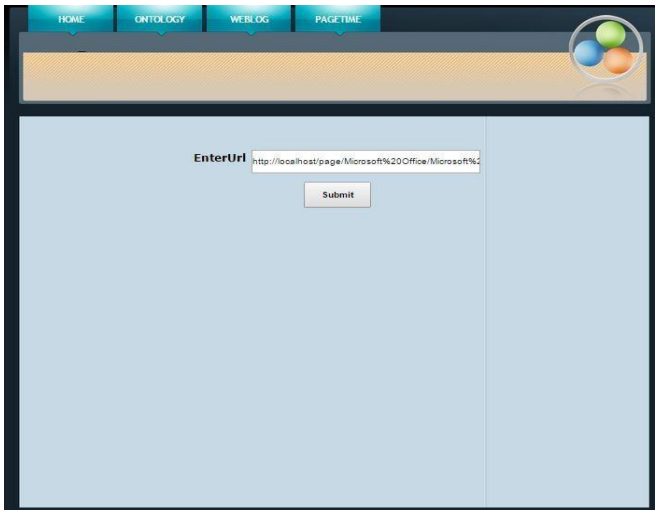

Fig.8 Crawling a page from server

Domain will be crawled by admin through their ontology link.

Fig.9 Web history of all users

Above figure represents the weblog of all users.

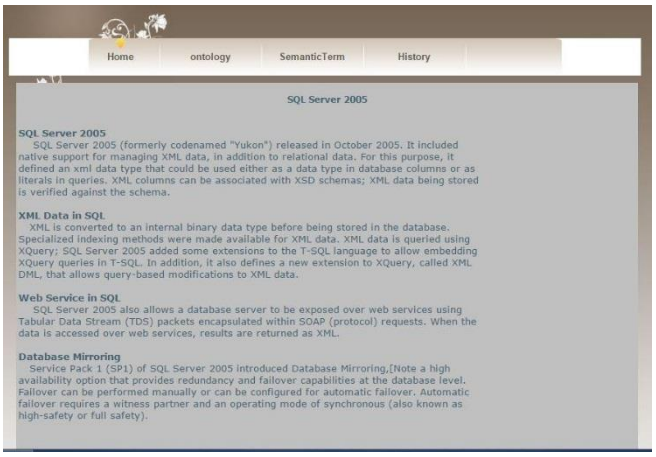

Fig.10 Web page hosted by admin

Hosted web page in IIS server.

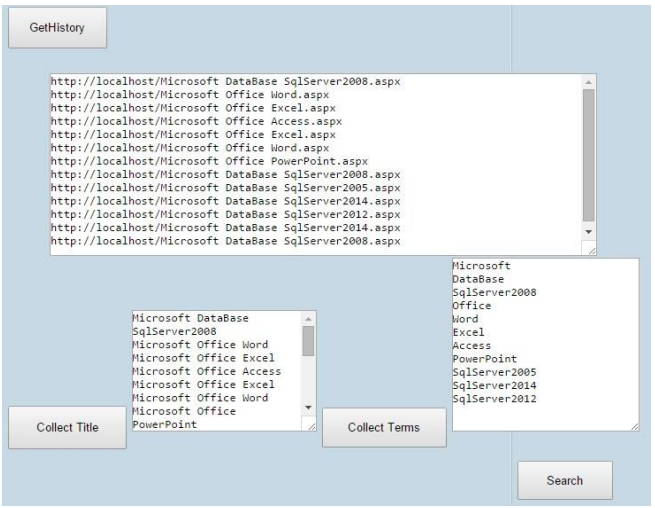

Fig.11 Web log of an user

Web history of a user based on their previous search

6. CONCLUSION

A new method to offer better web-page recommendations through semantic enhancement by three new knowledge representation models has been proposed. 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 web-pages, which can be automatically constructed, namely TermNetWP. The weighted semantic network of frequent viewed terms from the integrated information of web usage and domain knowledge by using proposed conceptual prediction model termed as TermNavNet. A number of web-page recommendation strategies have been proposed to predict next web-page requests of users through querying the knowledge bases.

Compared with earlier web usage mining method, 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.

For the future work, a key information extraction algorithm will be developed to compare with the term extraction method in this work, and also intense comparisons with the existing semantic web-page recommendation systems will be performed.

REFERENCES

[1] B. Liu, B. Mobasher, and O. Nasraoui, "Web usage mining," in *Web Data Mining: Exploring Hyperlinks, Contents, and Usage Data*, B. Liu, Ed. Berlin, Germany: Springer-Verlag, pp.527–603, 2011.

[2] B. Mobasher, "Data mining for web personalization," in *The Adaptive Web*, vol.4321, P. Brusilovsky, A. Kobsa, and W. Nejdl, Eds. Berlin, Germany: Springer-Verlag, pp.90–135, 2007.

[3] G. Stumme, A. Hotho, and B. Berendt, "Usage mining for and on the Semantic Web," in *Data Mining: Next Generation Challenges and Future Directions*. Menlo Park, CA, USA: AAAI/MIT Press, pp.461–480, 2004.

[4] H. Dai and B. Mobasher, "Integrating semantic knowledge with web usage mining for personalization," in *Web Mining: Applications and Techniques*, A. Scime, Ed. Hershey, PA, USA: IGI Global, pp.205–232, 2005.

[5] S. A. Rios and J. D. Velasquez, "Semantic Web usage mining by concept-based approach for off-line web site enhancements," in *Proc. WI- IAT'08*, Sydney, NSW, Australia, pp.234–241.

[6] S. Salin and P. Senkul, "Using semantic information for web usage mining based recommendation," in *Proc. 24th ISCIS*, Guzelyurt, Turkey, pp.236–241, 2009.

[7] Bose, K. Beemanapalli, J. Srivastava, and S. Sahar, "Incorporating concept hierarchies into usage mining based recommendations," in *Proc. 8th WebKDD*, Philadelphia, PA, USA, pp.110–126, 2006.

[8] N. R. Mabroukeh and C. I. Ezeife, "Semantic-rich Markov models for Web prefetching," in *Proc. ICDMW*, Miami, FL, USA, 465–470, 2009.

[9] M. O'Mahony, N. Hurley, N. Kushmerick, and G. Silvestre, "Collaborative recommendation: A robustness analysis," *ACM Trans. Internet Technol.*, vol.4, no 4, pp.344–377, Nov. 2004.

[10] G. Stumme, A. Hotho, and B. Berendt, "Semantic Web mining: State of the art and future directions," *J.Web Semant.*, vol.4, no.2, pp.124–143, Jun. 2006.