



PERSONALIZED RECOMMENDATION SYSTEM FOR ENHANCED WEB SEARCH

Mrs. Therasa M¹, Mr. Hari Krishnan K A², Mr. Kovarthanan R³, Mr. Rakesh A R⁴

¹ Assistant Professor, Department of Computer Science and Engineering, Panimalar Institute of Technology,
Chennai –

Tamilnadu, India.

therasamic@gmail.com

² UG Scholar, Department of Computer Science and Engineering, Panimalar Institute of Technology, Chennai -
Tamilnadu, India.

harikrishnanrider@gmail.com

³ UG Scholar, Department of Computer Science and Engineering, Panimalar Institute of Technology, Chennai –
Tamilnadu, India.

kovavishal@gmail.com

⁴ UG Scholar, Department of Computer Science and Engineering, Panimalar Institute of Technology, Chennai -
Tamilnadu, India.

rakeshravi29296@gmail.com

ABSTRACT--*The undertaken project proposes a novel strategy to proficiently give better Web-page suggestion through semantic improvement by coordinating the domain and Web utilization knowledge of a website. Two new models are proposed to speak to the domain knowledge. The principal show utilizes ontology to represent the domain knowledge. The second model uses one automatically produced semantic system to speak to domain terms, Web-pages and the relations between them. Another new model, the conceptual prediction model, is proposed to naturally create a semantic system of the semantic Web use knowledge, which is the joining of domain knowledge and Web utilization knowledge. Various powerful inquiries have been produced to inquiry about these knowledge bases.*

Keywords--*Ontology, Semantic Network, Web Usage Mining (WUM), Conceptual Prediction Model, Stemming Algorithm*

I. INTRODUCTION

Web-page recommendation plays an essential part in canny Web frameworks. Valuable knowledge disclosure from Web usage information and agreeable knowledge portrayal for successful Web-page recommendations are essential and testing. This paper proposes a novel technique to effectively give better Web-page recommendation through semantic improvement by coordinating the domain and Web usage knowledge of a website. Two new models are proposed to speak to the domain knowledge. The main model uses ontology to speak to the domain knowledge. The second model uses one naturally created semantic network to speak to domain terms, Web-pages and the relations between them. Another new model, conceptual prediction model, is proposed to consequently create a semantic network of the semantic Web usage knowledge, which is the reconciliation of domain knowledge and Web usage knowledge. Various compelling questions have been produced to inquiry about these knowledge bases. In view of these inquiries, an arrangement of recommendation procedures have been proposed to produce Web-page competitors. The recommendation comes about have been contrasted and the outcomes got from a progressed existing Web Usage Mining (WUM) technique. The test comes about exhibit that the proposed technique creates essentially higher execution than the WUM strategy.

II. DESCRIPTION

Proposed System is web page recommendation in view of the web usage and domain knowledge utilizing three new knowledge portrayal models. This framework utilizes an ontology-based model and semantic network model and conceptual prediction model.

2.1. Domain Ontology Model:

Ontology's are frequently actualized in a logic-based dialect, such as, OWL/RDF, to wind up justifiable to software specialists or software frameworks. In this manner, ontology based knowledge representation permits sharing and interchanging semantic data among Web frameworks. Ontological representation of found knowledge from various sources can be effortlessly coordinated to help Web-page recommendation viably.

2.2. Semantic Domain terms Generation:

Semantic domain terms are collocations of terms which are controlled by the co-event relations of terms in Web-page titles the relationship amongst terms and Web-pages. Furthermore, the domain terms and co event relations are weighted to give a harsh sign of how much these terms are related with each other semantically. In view of the relations between the terms and Web-pages, we can construe how intently the Web-pages are semantically identified with each other.

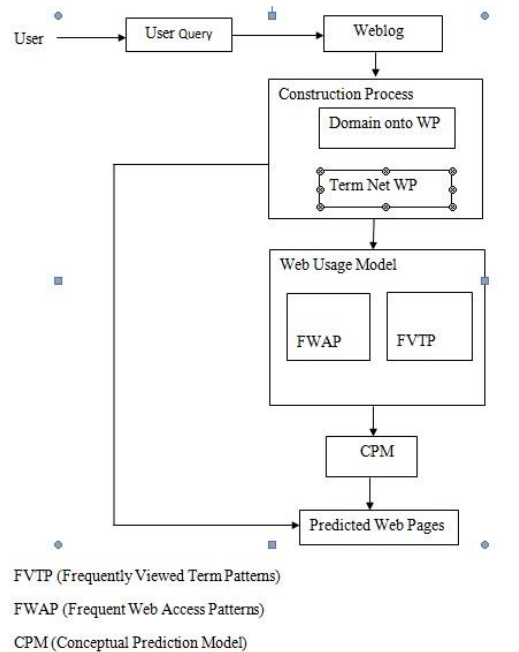


Figure1: Semantic Domain terms Generation

2.3. Conception Prediction Model:

Suggest the web pages in light of the Web usage knowledge that can be found from Web log documents utilizing a Web usage mining strategy. Find the Web usage knowledge, which is as incessant Web get to. (FWAP) and we incorporate FWAP with as often as possible saw term designs (FVTP), foresee the proficiently and suggest the URLs.

III. DOMAIN ONTOLOGY OF A WEBSITE FOR WEB-PAGE RECOMMENDATION

In the context of Web-page recommendation, the information is Web logs that record client sessions regularly. The client sessions incorporate data about clients' Webpage route exercises. Each Web-page has a title, which contains

the catchphrases that grasp the semantics of the Web-page. In light of these actualities, we expect to find domain knowledge from the titles of went by Web-pages at a website and speak to the found knowledge in a domain ontology to help viable Web-page recommendation. A domain ontology is characterized as a reasonable model that indicates the terms and connections between them unequivocally and formally, which thus speak to the domain knowledge for a particular domain. The three principle parts are recorded as takes after: 1) Domain terms (ideas), 2) Relationships between the terms (ideas), and 3) Features of the terms and connections. Ontologies are regularly executed in a rationale based dialect, for example, OWL/RDF, to end up justifiable to programming specialists or programming frameworks. Along these lines, ontology based knowledge representation permits sharing and exchanging semantic data among Web frameworks over the Internet. It likewise empowers the reuse of the domain knowledge, and thinking the semantics of Web-pages from the current actualities. Besides, ontological representation of found knowledge from various sources can be effortlessly incorporated to help Web-page recommendation successfully. Contingent upon the motivations behind ontologies, they can be composed as domain conceptualizations of different degrees of custom and can be as idea plans, scientific categorizations, applied information models, or general coherent speculations. In this area, we will develop a reasonable information display as a domain ontology for a given website. Since this ontology is utilized to help Web-page recommendation, we take a Web-page as a unit and accept each page title is very much characterized to speak to key data about the substance of the page. The method of reasoning behind this suspicion can be seen from two viewpoints. One viewpoint is that a Web-page contains a gathering of items (spoke to by HTML labels) reported in metadata, which is information about information. Metadata grasps the centre components of title, which means, illustrative setting, structure, and general setting of a Web-page. By breaking down the metadata, for example, Web-page title, the importance of a Web-page can be comprehended and caught. The second perspective is from the expert practice in Web advancement.

In all around composed Web-pages, the TITLE tag ought to contain the significant catchphrases which are moderately short and alluring to help Web hunt or creeping. By and by, the terms in page titles are normally given higher weights via web crawlers, for example, Google. Consequently, proficient website engineers need to characterize the Web-page titles truly in light of the fact that they need their Web-pages to be effectively recognized amid Web pursuit or creeping and utilize the Web-page titles to pass on precise data about the Web-page. Due to these certainties, we utilize the Web-page titles as signs to speak to the domain knowledge of a website. It suggests that despite the fact that there are various models for separating themes of Web-pages, making utilization of Web-page titles is basic and simple to actualize. This area now shows a technique for building the domain ontology utilizing the Microsoft (MS) website (www.microsoft.com) for instance. The dataset was downloaded from <http://kdd.ics.uci.edu/databases/msweb/msweb.html>. The ontology will be built in view of the titles of went by Web-pages with the goal that it is the domain knowledge saw by clients. Inquiries are then given in light of this domain ontology.

IV. THE MINING PROCESS FOR PROPOSED SYSTEM

With the help of Domain Ontology Construction, we will collect the terms available in metadata of the web-pages and then depending upon the metadata, they will be categorized in order i.e. define the concepts and after this we will define taxonomic and non-taxonomic relationships between words. By considering above model, semantic knowledge representation model of web usage of website for webpage recommendation will be considered. With the help of this model, the pages can be predicted

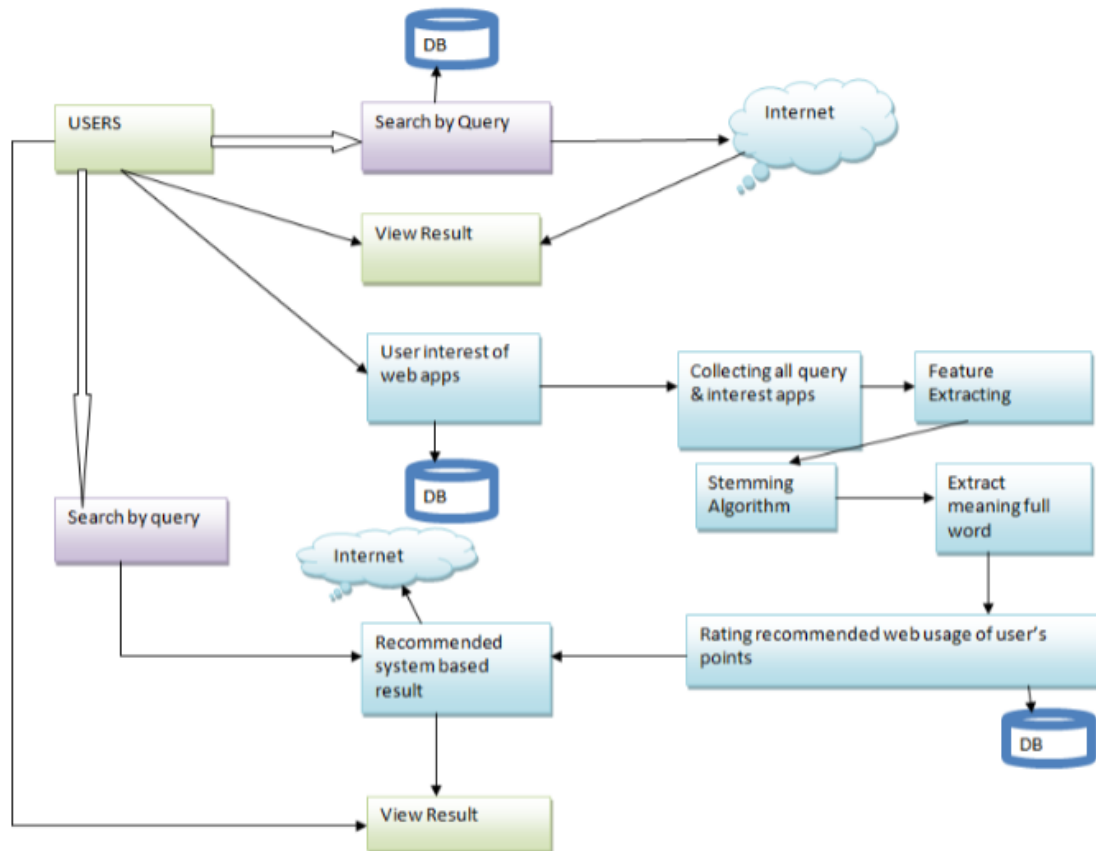


Figure2: Data flow diagram

The flow of work in the recommendation system is illustrated by step by step explanation as follows:

1. Firstly User signs up in the recommendation system to create a username and password.
2. On successive logins user provides his username and password.
3. User enters a query in search box for accessing information.
4. Every query received by recommendation system interface is passed to Google API for extracting websites containing relevant information from the Google. At the same time every query is scanned to analyze whether the same query was clickthrough by the user previously. If so, the already recommended web-pages are provided to users.
5. The web pages extracted from the Google API are given to the Recommendation system where the Stemming Algorithm will be executed and the resultant web pages will be displayed by their priority based on frequency of their domain terms.
6. All information accessed through recommendation system i.e. all the terms searched for or websites visited are maintained in the browsing history list dedicated for the user. Proposed system keeps an eye on every query provided by the user and the information accessed from the response generated by recommendation system. Database records this information in browsing history list maintained for the user and later mines it to draw useful knowledge about user interests.

7. Once the user session is over, proposed system mines the browsing history list, classifies user interests and applies clustering techniques to draw useful knowledge about the user interests.

V. SYSTEM FEATURES

The new technique to offer better Web-page recommendations is through semantic improvement by three new knowledge representation models. Two new models have been proposed for representing of domain knowledge of a website. One is an ontology-based model which can be semi-naturally developed, to be specific DomainOntoWP, and the other is a semantic network of Web-Pages, which can be consequently built, to be specific TermNetWP. A conceptual prediction model is additionally proposed to incorporate the Web usage and domain knowledge to shape a weighted semantic network of much of the time saw terms, to be specific TermNavNet. Various Web-page recommendation techniques have been proposed to foresee next Web-page solicitations of clients through questioning the knowledge bases. The trial comes about are promising and are characteristic of the convenience of the proposed model.

5.1 External Interface Requirements

5.1.1 User Interfaces

- User Interfaces are Graphical User Interfaces in this product.
- Users are communicated with Buttons to clear the content or send data to the destination.
- User can enter the data through the textbox.
- User can interact with text area to enter the multiple line of text.

5.1.2 Hardware Interfaces

5.1.2.1 Ethernet

Ethernet on the AS/400 supports TCP/IP, Advanced Peer-to-Peer Networking (APPN) and advanced program-to-program communications (APPC).

5.1.2.2 ISDN

You can connect your AS/400 to an Integrated Services Digital Network (ISDN) for faster, more accurate data transmission. An ISDN is a public or private digital communications network that can support data, fax, image, and other services over the same physical interface. Also, you can use other protocols on ISDN, such as IDLC and X.25.

5.1.3 Software Interfaces

- This software is interacted with the TCP/IP protocol.
- This product is interacted with the Socket and listening on unused ports.
- This product is interacted with the Server Socket and listening on unused ports.

5.2 Communications Interfaces

The TCP/IP protocol will be used to facilitate communications between the client and server.

5.3 Other Non-functional Requirements

5.3.1 Performance Requirements

Response time is estimated from the time that the client plays out the activity that says "Go" until the point when the user gets enough criticism from the system to proceed with the undertaking. It is the client's subjective hold up time. It isn't from passage to a subroutine until the point that the main compose articulation. On the off chance that the client denies enthusiasm for reaction time and demonstrates that exclusive the outcome is of intrigue, you can ask whether "ten times your present gauge of remain solitary execution time" would be worthy. In the event that the appropriate response is "yes," you can continue to examine throughput. Else, you can proceed with the discourse of

reaction time with the client's complete consideration. The response time that is negligibly worthy whatever remains of the time. A more drawn out response time can make clients think the framework is down. You additionally need to indicate rest of the time; for instance, the pinnacle moment of a day, 1 percent of connections. Response time degradations can be all the more exorbitant or excruciating at a specific time.

5.3.2 Safety Requirements

The software may be safety-critical. If so, there are issues associated with its integrity level. The software may not be safety-critical although it forms part of a safety-critical system. For example, software may simply log transactions. If a system must be of a high integrity level and if the software is shown to be of that integrity level, then the hardware must be at least of the same integrity level. There is little point in producing 'perfect' code in some language if hardware and system software (in widest sense) are not reliable. If a computer system is to run software of a high integrity level then that system should not at the same time accommodate software of a lower integrity level. Systems with different requirements for safety levels must be separated. Otherwise, the highest level of integrity required must be applied to all systems in the same environment.

5.3.3 Security Requirements

- Do not block the some available ports through the windows firewall
- Two machines should be connected with LAN setting

VI. PERFORMANCE EVALUATION

So as to assess the viability of the proposed models of knowledge representation and the recommendation techniques alongside the queries, we execute these models, algorithm and systems to test their execution of Web-page recommendation utilizing an open dataset. In this segment, we initially list the measures for the execution assessment of Web-page recommendation procedures, and after that present the outline of the tests, trailed by the examinations of test comes about.

The performance of Web-page recommendation strategies is measured in terms of two major performance metrics: Precision and Satisfaction according to Hao Chen[1]

The precision is useful to measure how probable a user will access one of the recommended Web-pages. Besides, we also need to consider if a user accesses one of the recommended Web-pages in the near future. Actually, the next page accessed by a user may not be the target page that user wants. In many cases, a user has to access a few intermediate pages before reaching the target page. Hence, the satisfaction is necessary to give the precision that the recommended pages will be accessed in the near future.

VII. CONCLUSION

Our principle goal of this paper is to give a precise recommendation to the client in light of the usage and domain knowledge to influence the client to encounter more solid and productive. Taking everything into account, this paper has exhibited another technique to offer better Web-page recommendations through semantic upgrade by three new knowledge representation models. Two new models have been proposed for representation of domain knowledge of a website. A conceptual prediction model is likewise proposed to incorporate the Web usage and domain knowledge to shape a weighted semantic network. The trial comes about are promising and are characteristic of the helpfulness of the proposed models. Contrasted and one of the most developed Web usage mining technique, (i. e) PLWAP-Mine, the proposed strategy can significantly improve the execution of Web-page recommendation as far as exactness and fulfilment.

For the future work, a key data extraction algorithm will be created to contrast and the term extraction strategy in this work, and we will perform exceptional examinations with the existing semantic Web-page recommendation frameworks.

REFERENCES

- [1] Hao Chen and Susan Dumais, "Bringing Order to the Web: Automatically Categorizing Search Results" SIGCHI Conference on Human Factors in Computing Systems, the Hague, The Netherlands. pages 145-152(April 1-6 2000).

- [2] Uichin Lee, Zhenyu Liu and Junghoo Cho," Automatic Identification of User Goals in Web Search"14th International Conference on World Wide Web, Chiba, Japan .pages391-400(May10-14 2005)
- [3] Ricardo Baeza-Yates, Carlos Hurtado and Marcelo Mendoza," Query Recommendation using Query Logs in Search Engines " 2004 International conference on Current Trends in Database Technology, Heraklion , Greece (March 14-18 2004)
- [4] Steven M. Beitzel, Eric C. Jensen, Abdur Chowdhury, Ophir Frieder," Varying Approaches to Topical Web Query Classification" International SIGAR Conference, Amsterdam, The Netherlands (July 23-27 2007)
- [5] Huanhuan Cao, Daxin Jiang, Jian Pei, Qi He,Zhen Liao, Enhong Chen and Hang Li," Context-Aware Query Suggestion by Mining Click-Through and Session Data" 14th ACM SIGKDD International conference on Knowledge discovery and data mining, Las Vegas, Nevada, USA.Pages:875-883(August 24-27 2008)
- [6] Rui Li, Alex X. Liu, Ann L. Wang and Bezawada Bruhadeshwar," Fast and Scalable Range Query Processing With Strong Privacy Protection for Cloud Computing" IEEE/ACM Transactions on Networking (Volume: 24, Issue: 4) (August 2016)
- [7] Peter Mika,"Ontologies are us: A unified model of social networks and semantics" International Semantic Web Conference(ISWC), Volume 5, Issue 1, Pages 5-15 (March 2007)
- [8] Mohamed Nader Jelassi, Sadok Ben Yahia and Engelbert Mephu," A personalized recommender system based on users' information in folksonomies" 22nd International Conference on World Wide Web, Rio de Janeiro, Brazil. Pages:1215-1224(May 2013)
- [9] Michael G. Noll and Christoph Meinel," Web Search Personalization via Social Bookmarking and Tagging" 6th International The semantic web and 2nd Asian conference on Asian semantic web conference, Busan, Korea. Pages 367-380 (Nov 11-15 2007)
- [10] Arkaitz Zubiaga, V'ictor Fresno, Raquel Mart'inez and Alberto P. Garc'ıPlaza,"Harnessing Folksonomies to Produce a Social Classification of Resources" IEEE Transactions on Knowledge and Data Engineering,(Volume: 25, Issue: 8) Page(s): 1801 - 1813 (Aug 2013)