

Exploiting Semantic Web Technology for Personalized Recommendation System

Preeti Deshmukh¹, Shivani Deosthale², Lukesh Kadu³

Information Technology, Shah & Anchor Kutchhi Engineering College, Mumbai

Abstract:

World Wide Web makes a revolution on information sharing. People could search any kinds of information from Internet and create their own web sites. However, with the increasing of information published on the Internet, the problems of redundant data, error data and incomplete data come up, which leads to the challenge of information retrieval accuracy. Users need help to search and to filter out irrelevant information based on their demand. In this Paper an approach for Semantic Recommendation system which uses Page rank based on links between Web pages, to compute relevance to user search query. It works on the assumption that a well-designed Webpage should have an informative title. The user search terms are mapped to the refined URL's extracted from the Google Server. These terms in a user query have a different degree of importance based on whether they occur inside title, description or anchor text of the document. Recommendation system has to find the relevant Web pages and prioritize them. Therefore an optimal weighting scheme for these components is required for search engines to prioritize relevant documents near the top for user searches. Personalized Recommendation System provides other user with a recommendation results for based on the common shared search queries.

Keywords — Term Extraction, Concept Extraction, Semantic web, Web Usage and Domain Knowledge.

I. INTRODUCTION

Nowadays, Intelligent recommender systems on the web intends to recommend web pages for individual users by discovering useful knowledge from Web usage data and web content data. Knowledge representation for the web contents and integrating with web usage knowledge are the challenging issues to make Web page recommendations effective.

Differently from traditional search engines, a semantic search engine stores semantic information about Web resources and is able to solve complex queries, considering as well the context where the Web resource is targeted. Semantic search integrates the technologies of Semantic Web and search engine to improve the search results gained by current search engines and evolves to next generation of search engines built on Semantic Web. Adding semantics to the web involves two things: allowing documents which have information

in machine-readable forms, allowing links to be created with relationship values.

Recommending web pages are based on web usage mining so that can serve a set of web pages which impress the individual user based on the preferences formed via usage patterns. Web recommendation became a key need in web services, where in the websites, the links to the recently viewed web pages is shown, is also performs efficiently and effectively, in which some of the machine learning algorithms can also be used. Personalization is a factor which is to be given much importance as far as E-learning Recommender systems are concerned [1].

II. LITERATURE SURVEY

The related work and literature review covers the background, latest development of and related techniques for semantic-enhanced recommender systems.

Recommender systems differ from each other mainly through their filtering method. Distinctions between types of filtering systems are made, namely collaborative, content-based. Collaborative filtering systems generate recommendations obtained from persons having similar interests. Content-based filtering only takes into account descriptions of products, based upon metadata and extracted features. Semantic data mining research has attested the positive influence of domain knowledge on data mining. For example, the pre-processing can benefit from domain knowledge that can help filter out the redundant or inconsistent data [2].

Types of Recommendation Systems

An intelligent recommender system on the web intends to recommend web pages for individual users by discovering useful knowledge from Web usage data and web content data. An effective method is to integrate the domain knowledge and web usage knowledge of website through semantics [3].

Integrating semantic information in the Web usage mining process is used where sequential pattern Mining technique is applied over the semantic space to discover the frequent sequential patterns. The frequent navigational patterns are extracted in the form of ontology instances instead of Web page views and the resultant semantic patterns are used for generating Web page recommendations to the user [4].

Recommend documents that contain the keywords that were included in user preferred documents by using keyword-matching techniques. This type of recommendation method based on the frequency of keywords that were contained in previous user queries. This type of a recommender learns from the user preferences and suggests items that are similar to the likes of the user [5].

The approach used for web content mining use ontology which specifies the domain specific knowledge and finding its relevant content, by applying the various clustering techniques [6]. It is a process that assigns items to groups so that the items in the same groups are more similar than items in different groups.

In knowledge-based The Semantic Web mechanisms incorporate additional knowledge about the user's preferences into recommendation process, specifically in the user profile, by elicitation information from the pages user visited and rated during his practice with the system, and therefore provides more accuracy and flexibility to the personalization processes [7].

The ontology can be built by enterprise ontology development which describes the group of steps and tasks that is to be done for building an ontology if no priori information or ontology. This method of building ontologies starts from scratch, which does not use a-priori knowledge [8].

A tool, Onto Learn aimed at the extraction of domain ontologies from websites. Onto Learn extracts a domain terminology from available documents [9]. The complex domain terms are semantically interpreted and arranged in a hierarchical fashion. Finally, a general-purpose ontology is enriched with the detected domain concepts.

A method to extract domain ontology from web sites without using a priori knowledge. This approach takes the web pages structure into account and defines a contextual hierarchy [10]. The data pre-processing is an important step to define the more relevant terms to be classified. Weights are associated to the terms according to their position in this conceptual hierarchy. Then, these terms are automatically classified and concepts are extracted. Semantic information is embedded into the web page recommendation techniques, which generates the ontology of websites, which improved the performance of the systems significantly [3]. Sometimes system such as web personalization system, can reuse the existing ontology, and then combine web usage mining technique, where concepts are from the web logs enhanced with semantics. In this system, the ontology classes are retrieved from the text documents.

III. PROBLEM STATEMENT

The biggest issue faced by recommendation systems is that they need a lot of data to effectively make recommendations. There are other problems like content visited, behaviour of users, content search, content uploaded. A major problem in content based recommendation is with its Limited contents. There was no more personalization based on Semantic and domain or history knowledge. It had faced "New user problem" where users did not had sufficient information before, to get accurate recommendations.

Recommendation system which uses domain knowledge and history knowledge to enhance the personalization, where the users were navigated to relevant pages can be an effective way of recommendation. A personalized-web-recommendation system that made use of representations of user and search word based on semantics in recommendation. Users of Web search usually submit short and ambiguous queries to specify their requirement. After using search engine lots of data gets accumulated, from the queries provided by users, to retrieve documents the ones with the highest relevance are stored as query log. These query logs provide valuable information to extract relationships between queries and documents that can be used for future use. Relevant or most suitable expansion terms are selected from the database with the analysis of relation between queries and documents.

Specific Problems.

- A. *Meaningful Buzz Word: Everybody makes up his own meaning (like 'Artificial Intelligence', 'Chaos Theory').*
- B. *The Hype: People without intricate understanding involved prior to prove results.*

IV. WEBPAGE RECOMMENDATION SYSTEM.

Web-page recommendation plays an important role in intelligent Web systems. Useful knowledge discovery from Web usage data and satisfactory knowledge representation for effective Web-page recommendations are crucial and challenging. A novel method to efficiently provide better Web-page recommendation by integrating the domain A number of effective queries have been developed

to query about these knowledge bases. Based on these queries, a set of recommendation strategies have been used to generate Web-page candidates.

V. SEMANTIC RECOMMENDATION SYSTEM

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 infer how closely the Web-pages are semantically related to each other. User search query on Webpages and the relevant pages for a given search query infer the semantics of Web-pages to achieve semantic enhanced Web-page recommendations.

In this recommendation system Semantic Domain knowledge about relations between terms and Webpages for user Search query is use, the relevant pages for a search is obtained, the key terms for a given page, and the pages for given terms is extracted , to infer the semantics of Web-pages which enhances Semantic Web-page recommendations. Where each term instance is specified to relevant keywords in the Semantic Network.

VI. IMPLEMENTATION

This paper proposed a model for personalized web page recommendation based on Semantic Network and Domain Knowledge [1]. The semantic web usage knowledge is integrated with both domain knowledge and web usage knowledge. The implementation of Recommendation model is demonstrate and the results produces significantly higher performances than the Webpage Recommendation System.

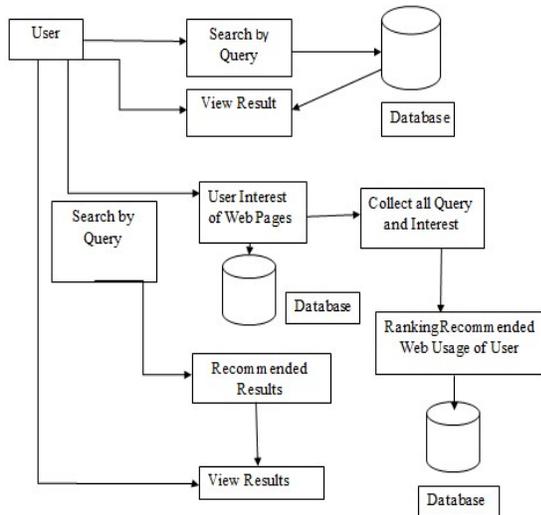


Fig 1: Semantic Web Recommendation System.

The problem in contextual semantic search systems resides on building a new domain ontology which has not been defined before. Knowing that the Web is an enormous information source and a dynamic, we have the idea to integrate ontology learning process in the search process [11].

The first step of this proposed system is Feature Extraction. Feature extraction technique approach extracts features from web documents and constructs relevant concepts. Then it extracts features from web documents. After completing this, the semantic similarity of web documents it taken into consideration to fetch the information on web.

A novel method to efficiently provide better Web-page recommendation through semantic enhancement by integrating the domain and Web usage knowledge of a website has been discussed. In this dissertation, the system is based on online data fetched from the Google's search engine.

Using the current visited Web-page (referred to as a state) and k previously visited pages (the previous k states), the Web-page(s) that will be visited in the next navigation step can be predicted. With the help of Domain Knowledge 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 is represented.

VII. SYSTEM OVERVIEW

In order to obtain the semantic Web usage knowledge that is efficient for semantic-enhanced Web-page recommendation, a conceptual prediction model (CPM) is proposed to automatically generate a weighted semantic network of frequently viewed terms with the weight being the probability of the transition between two adjacent terms based on FVTP.

From the result, we discover that some words which have no distinct domain feature and maybe their domain labels are ambiguous. Larger Domain knowledge can make the results more reliable.

Its evaluation in content classification indicates the proposed method is more effective method for extraction of domain-specific terms. Also the proposed method achieves better performance by adding the semantic meaning. This indicates replacing features by domain-specific terms is an effective way to content classification.

The Semantic Web Recommendation System is an engine that takes user query as input. Extracts tokens from the User Query and stores the tokens in an SQL database. Before storing it also performs some processing on the tokens by adding the Semantics and tries to get useful domain terms related to user from the User Query.

It also gives the frequently used Domain term count, the number of terms that where used to extract the relevant information across the collection of documents.

RESULTS

In order to prove the effectiveness of the Semantic Web Recommendation System, we used the Google Database to crawl the User Search query and crawled the pages from the site.

We then find the co-occurrence of the user search query and the semantic terms from the documents

to provide Semantic Web Recommendation System and obtained the following results. The two Recommendation system Webpage Recommendation System and Semantic Web Recommendation System had the same useful data but the files from the Semantic system have additional useful information added to it.

The recommendation systems had experimented with the variations of user query and in one of such experiment where three times a different User Search Query was collected and the results of those results came out as shown in Fig 2.

Following where the Users search queries:

Query1: College in Mumbai

Query2: College of Engineering in Mumbai

Query3: College of Engineering and Diploma in Mumbai.

Table 1: Comparison between the Two Recommendation Systems.

A set of distinct user search queries were fired and the Average execution time for Recommendation was measured and the result shows that Semantic Recommendation system takes less time as compared to Webpage Recommendation system.

Here to get fast and relevant results we have use semantic network which had helped in providing recommendation system, a powerful technique to finding data over web.

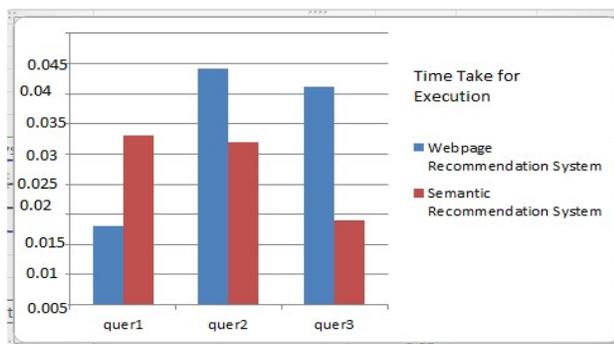


Fig 2. Comparison between the Two Recommendation Systems.

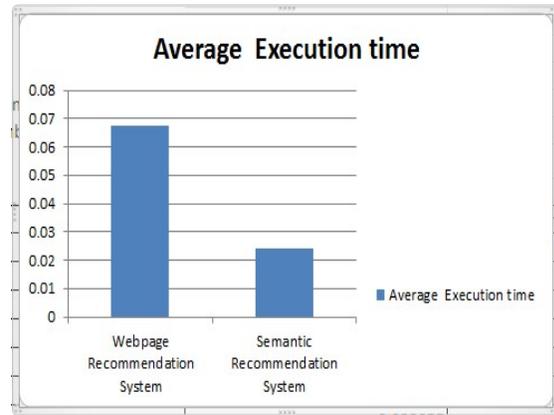


Fig 3. Average Execution Time in Milliseconds for Both Recommendation Systems.

VIII. CONCLUSION

The System discussed is a conceptual prediction model which had integrated the web usage and domain knowledge. The underlying idea was to integrate usage data with content semantics, in

Different User Search Pattern	Webpage Recommendation System	Semantic Recommendation System
Query1	0.0130082 secs	0.0280398 Secs.
Query2	0.0390550 Secs.	0.0270374 Secs.
Query3	0.360563 Secs.	0.0140173 Secs.

order to produce a semantically enhanced navigational pattern which was subsequently useful in producing valuable recommendations [12]. To automatically generate a weighted semantic network of frequently viewed terms with the weight was implemented where the probability of the transition between two adjacent terms based on Frequently Viewed Termed Patterns was focused.

In this paper, a semantic web personalization framework, which had enhanced the recommendation process with content semantics was discussed. The framework exploited the inherent semantic similarities between the domain terms in order to group web documents together and semantically expand the recommendation set. It has formed a weighted semantic network of frequently viewed pages and terms. This system had predicted next Web Page requests of users

through querying the knowledge. The experimental results are promising and indicates usefulness of the proposed models. Semantic recommendation system provided quick and relevant results for more précised search queries. Here, more the number of patterns matched to the Webpages, less time taken for execution of the search queries. Despite the difference in the execution time taken by these systems the performance of both the recommendation system in context in finding the relevant webpages is high. Time Execution required for Webpage Recommendation System was more as compare to Semantic Recommendation System. Semantic Recommendation System makes use of Domain Knowledge and Semantic network and improves the search results for user query. Recommendation Systems help users in sharing experience over the search results by providing recommended results in the Log file without disclosing the identity of the user whose recommendation where used.

In future scope, we can use different sizes of dataset and evaluate performance of recommendation system accordingly.

ACKNOWLEDGMENT

My thanks to the Principal Dr.V. C. Kotak, Guide Ms. Shivani Deosthale and Mr. Lukesh Kadu who all provided me constructive and positive feedback during the preparation of this project work.

REFERENCES

- [1] S. Kanetkar, A. Nayak, S. Swamy and G. Bhatia, "Web-based Personalized Hybrid Book Recommendation System", IEEE International Conference on Advances in Engineering & Technology Research (ICAETR), 2014.
- [2] D.Dou,H.wang and H.Liu,"Semantic Data Mining: A Survey of Ontology-based Approaches", IEEE International Conference on Semantic Computing (ICSC), 2015.
- [3] M. Venu Gopalachari ,Po Sammulal ,"Personalized Web Page Recommender System using integrated Usage and Content Knowledge", IEEE International Conference on Advanced Communication Control and Computing Technologies,2014.
- [4] T. Thanh, S. Nguyen, H. Lu, and J. Lu,"Web-Page Recommendation Based on Web Usage and Domain Knowledge", IEEE Transactions on Knowledge and Data Engineering, 2014.
- [5] S. Kanetkar, A. Nayak, S. Swamy and G. Bhatia, "Web-based Personalized Hybrid Book Recommendation System", IEEE International Conference on Advances in Engineering & Technology Research (ICAETR), 2014.

- [6] H. Liao," A new web service model of hybrid personalized recommendation", Ninth International Conference on Natural Computation (ICNC), 2013.
- [7] N. Madhak, C. Varnagar and S.Chauhan," A Novel Approach for Improving the Recommendation System by Knowledge of Semantic Web in Web Usage Mining", International Journal of Engineering Research and Development, 2014.
- [8] M. Craven,D.DiPasquo,D. Freitag,A. McCallum,T. Mitchell,K. Nigam and S.Slattery, "Learning to construct knowledge bases from the World Wide Web". Artificial Intelligence.
- [9] R.Navigliand P.Velardi,"Learning Domain Ontologies from Document Warehouses and Dedicated Web Sites"Computational Linguistics.
- [10] L.Karoui, , M.-A.Aufaure, , and N.Bennacer,"Ontology Discovery from Web Pages: Application to Tourism", Workshop on Knowledge Discovery and Ontologies (KDO), collocated with ECML/PKDD.
- [11] N.Mustapha, H.Baazaouighal, M.Aufaure, and H.Ghezala,"Survey on Ontology learning from Web and open issues", Laboratory RIADI-GDL, National School of Computer Sciences, University of Manouba, 2010.
- [12] M.Eirinaki ,D.Mavroeidis,G.Tsatsaronis and M.Vazirgiannis,"Introducing Semantics in Web Personalization: The Role of Ontologies, Semantics, Web and Mining".Springer,2006.