

# Personalized Web Page Recommendation Using Ontology

Arundhati Patil

ME II (Computer),

SKN, Sinhgad Institute of Technology and Sciences,  
Lonavala,

Email: arundhati9999@gmail.com

Prof. Supriya Sarkar

Computer Dept.

SKN, Sinhgad Institute of Technology and Sciences,  
Lonavala,

Email: supriya.sarkar@rediffmail.com

**Abstract:-** In this network era, Web Page Recommendation and web page Recommendation systems can take advantage of semantic network reasoning-capabilities to overcome common limitations of current systems and improve the recommendations' quality. This paper presents a personalized-web-recommendation system, a system that makes use of representations of items and user-profiles based on ontology in order to provide semantic applications with personalized services. The recommender uses domain ontology to enhance the personalization: on the other hand, user's interests are modeled in a more effective and accurate way by applying a domain-based inference method; on the other hand, the stemmer algorithm used by our content-based filtering approach, which provides a measure of the affinity between an item and a user, is enhanced by applying a semantic similarity method. Web Usage Mining plays an important role in web page recommender systems and web personalization system. In this paper, we propose an effective personalized web recommendation system based on ontology and Web Usage Mining. The proposed approach integrates semantic knowledge into Web Usage Mining and personalization processes.

**Keywords:** Web Usage Mining, Ontology, History Knowledge, Semantic Network, Domain Knowledge and Web-Page Recommendation.

\*\*\*\*\*

## 1. INTRODUCTION

Web-Page Recommendation Systems are becoming more useful and popular. Network is getting wider and wider. Websites are growing in numbers. So it is challenging task of the webmasters to organize the contents of the particular website in the network to gather the needs of the users. This paper provides a better solution to the Web Mining Process for recommendation purpose. The integration of semantic knowledge system is, in fact, the primary challenge for the next generation of personalization systems. It is now common for Web users to encounter sites that provide dynamic recommendations for products and services, targeted banner advertising, and individualized link selections. Indeed, nowhere is this technique is more apparent as in the business-to-consumer in e-commerce arena.

The reason is that, in today's highly competitive e-commerce environment, the success of a site often depends on the site's ability to retain visitors and turn casual browsers into potential customers. An Automatic personalization and recommender system technologies have become critical tools, because they help engage visitors at a deeper and more intimate level by tailoring the site's interaction with a visitor to her needs and interests. Web personalization can be defined as any action that tailors the Web experience to a particular user, or set of users. Before getting deep into the main concept, let's understand the additional concepts which are going to be used throughout the paper.

The concepts which will be used in the whole paper are given below:

- Web Mining Process
- Semantic Network
- Ontology
- Domain Knowledge
- Web Page Recommendation
- Knowledge Base Construction

The goal of this technique is to find the instances of the concepts and relations from the Web site's domain, because of that they can be used to perform further data mining tasks in the system. Learning algorithms play an important role in this technique.

## 2. LITERATURE SURVEY

Classical Techniques [2] of web page recommendation includes the result which is generated in the form of web pages, so we can say that there is no semantic meaning of common navigation profile system. There are many problems which are included in the classical that is old Systems. Now one of these problems is NEW PAGE PROBLEM in the network. The classical systems failed to recommends the newly added pages[3,5,6] or the products to the visitors since these pages or products are not in the current common navigation profiles.

So, to overcome the New page Problem the common navigation profile can be extracted in terms of semantic meaning or information. For that particular purpose ontology should be used. But in classical systems ontology was not used. Another problem is clustering [4, 6] in the

existing systems. In this clustering number of recommended pages get increased. Non Useful or Unrelated pages or links are recommended which user never prefers[2, 7].

Previous techniques used the RDF[9] and LCS algorithms.

LCS: Longest Common Subsequence

RDF: Resource Description Framework

The comparison of two sequences to determine their similarity is one of the fundamental problems in pattern matching technique. The Longest Common Subsequence method generates a sequence or a list of recommended products to the user. It is useful for the latest trend shopping that is online shopping purpose but what about the other users who does not use such online shopping websites but they still want the personalization in their web page recommendation system. The Numerical or numbering features should be represented by single components of vectors representing items. These components hold the exact value of that feature. Unfortunately, classifiers of all types tend to take a long time to construct. For example, if we wish to use decision trees, we need one tree per user. Constructing a tree not only requires that we look at all the item profiles, but the problem is we have to consider many different predicates, that could involve complex combinations of features. Hence, this approach tends to be used only for relatively small problem sizes.

The navigation of the users is largely driven by the semantics. Every time while user is searching he actually aims at finding some information concerning a particular different subject. There are many methods to extract keywords that characterize the web content. Which should be the exact matching between the terms determines the similarity between documents. Previous systems use this approach to find the similarity between documents. But by using this approach, only the binary similarity is achieved. That means no actual semantic similarity is considered. Semantic similarity is very important as far as similarity between documents is considered.

Many number of research approaches integrate other information sources. We require more abstract representation which enables a more flexible and uniform document matching process. It uses the semantic web structures, such as ontology. Early systems possess a very common problem of caching of web pages. When a web user searches for an already cached page, this action is not recorded in the web site's log.

### 3. PROBLEM STATEMENT

The problem in the existing systems is that, there was no more personalization based on ontology and domain or history knowledge. So that the personal recommendation or

personalized recommendation system. This system uses domain ontology and history knowledge to enhance the personalization should be there to use for web page recommendation purpose.

### 4. PROPOSED MODEL

Modules

- Creating Search history
- Query clustering
- Query reformulation
- History grouping

These modules are explained in detail as follows:

- Creating Search history

Any personal documents such as browsing history and emails on a user's computer could be the data source for user profiles. This focus on frequent terms limits the dimensionality of the document set, which further provides a clear description of users' interest. This module allows the search engine to better understand a user's session and potentially tailor that user's search experience according to the needs. Once query groups, which have been identified, that the search engines can have a very good representation of the search context behind the current query using queries and clicks in the corresponding query group.

- Query clustering

User's queries can be classified into different query clusters. The Concept-based user profiles are employed in the clustering process to achieve personalization effect. The most similar pair of concept nodes and after that, merge the most similar pair of query nodes, and so on. Each individual query submitted by each user is treated as an individual node and each query with a user identifier. We perform the grouping in a similar dynamic fashion, and whereby we first place the current query and clicks into a query group

- Query reformulation

To ensure that each query group contains closely related and relevant queries and clicks, so it is important to have a suitable relevance between the current query groups. We can assume that users generally issue very similar queries and clicks within a short period of time. Then search history of a large number of users contains signals regarding query relevance, like which queries tend to be issued closely together. This technique captures the relationship between queries frequently leading to clicks on similar URLs. Query reformulation graph and the query click graph from search logs, as well as how to use them to determine relevance between queries or query groups within a user's history.

- History grouping

Query groups is to first treat every query in a user’s history as a query group, and then merge these query groups in an iterative fashion (in a k-means). However, this is impractical in our scenario for two reasons. The first is, it may have the undesirable effect of changing a user’s existing query groups, which are used for undoing the user’s own manual efforts in organizing the history. Second, it involves a high-computational cost, for that we would have to repeat a large number of query group similarity computations for every new query.

### 5. SYSTEM DESIGN AND IMPLEMENTATION

This paper proposes a model for personalized web page recommendation based on Ontology, Semantic Network and Domain Knowledge [10]. There will be semantic web usage knowledge which an integration of both domain knowledge and web usage knowledge. The previous implementation of this model will demonstrate the results which produces significantly higher performances than the previous web usage mining and web page recommendation techniques.

This paper proposes a formal framework for integrating full domain ontology with personalization process based on

Web usage mining. It is because the classical process was not preferable by the user for the personalization purpose. The Semantic Web Mining Process will provide an infrastructure that enables not just web pages but databases, programs, services, sensors, personal devices and even household appliances to both produce and consume data on the web. The system proposed here, which is used to represent the items and user-profiles based on ontology in order to provide semantic applications with personalized services. The proposed recommender system is Domain Independent system and it is implemented as a Web Service. This kind of web service system uses both Implicit Feedback method and Explicit Feedback Collection Method which is used to obtain the user’s interest in particular search area.

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 builds the Ontology for the website usage by using the extracted features from web documents. After completing this, the semantic similarity of web documents is taken into consideration to cluster the web documents into different Semantic Themes the different themes may contain different preferences.

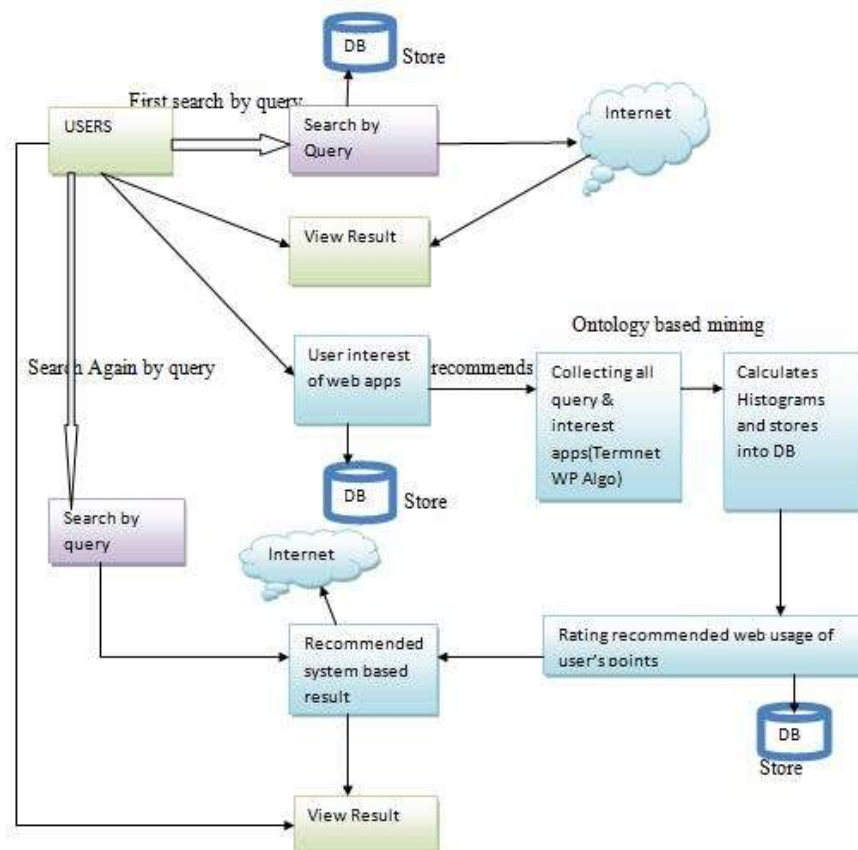


Fig1. Proposed Architecture.

For this purpose basically two algorithms are used. They are as follows:

- A. TermNetWP
- B. PrefixSpan

A. Algorithm:

Automatically construct a TermNetWP[1]

Input:

d (PageID)

O (TermNetWP)

Output:

T (list of domain terms)

Process:

Traverse through O to get the WPage instance whosePageID is d

Set,T = term instances associated with Web-page di via the ‘belongto-Instance’ object property

Sort T in descending order of their occurrences

Return T

Three Steps are involved in this algorithm:

1. Collect the titles of visited Web-pages(TSC)
2. Extract term sequences from the Web-page titles
3. Build the semantic network – TermNetWP

B. Algorithm:

Prefix-Span

Input:

The Sequencedatabase S and the minimum threshold MinSup.

Output:

Complete set of Sequential Patterns.

Process:

Step 1:

To find length of sequence patterns for the Sequence database S considering the minimum support that has been specified.

Step 2:

Divide the search space into the prefixes to search for the ontology.

Step 3:

The subsets of sequential patterns can be searched by constructing projected databases of the prefixes supported. The searching is either can be depth first or breadth first search.

Step 4:

Repeat the steps for sequential patterns until frequent sequences of the database are not found.

## 6. SYSTEM OVERVIEW

General and overall framework of web personalization for Web Usage Mining consists of 3 phases:

1. Data Preparation
2. Transformation
3. Pattern Discovery and Recommendation

These phases are the raw phases of the Semantic Network[1,8] which is used to make relation between concepts for the ontology. The first phase that is Data Preparation phase is used to transform raw Web Log Files into transaction data that can be processed by data mining techniques. Ontology phase takes domain information and web pages as input and generates site ontology. This is the system which is used to represent the items and user-profiles based on ontology in order to provide semantic applications with personalized services.

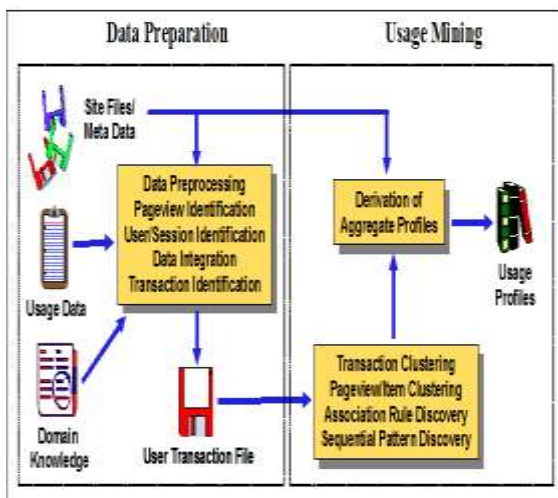


Fig. 2. General Framework for Web Personalization[11]

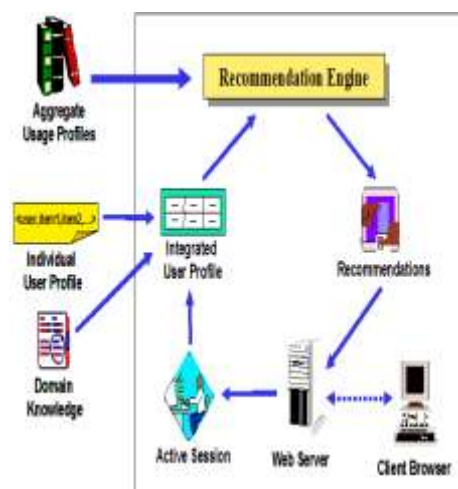
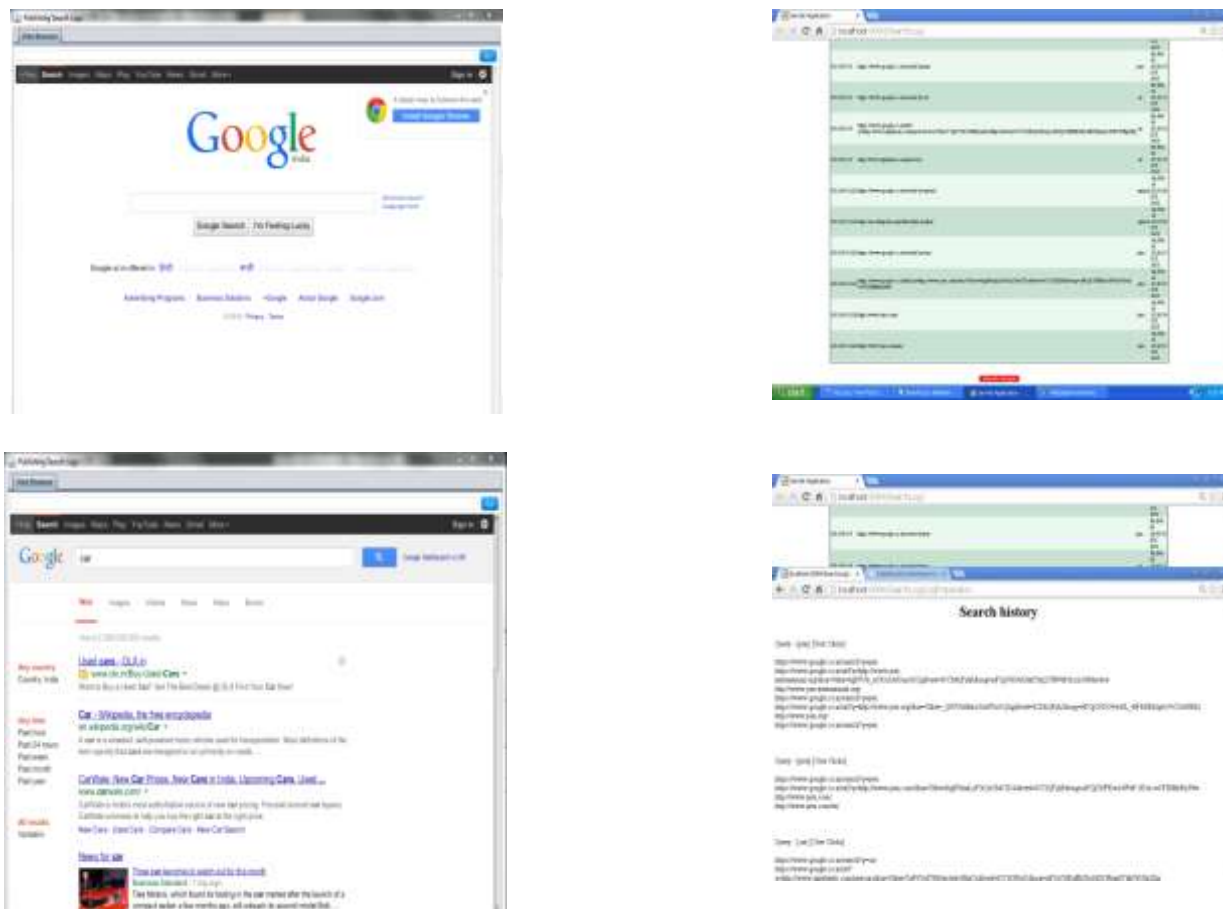


Fig. 3. Online Personalization Component [11]

**ACTUAL IMPLEMENTATION**



**Fig 4. Actual Implementation**



**7. CONCLUSION**

In conclusion, we can say that this paper proposes a new method to offer a better Web Page Recommendation System by using Semantic Enhancements and ontology, which are called as Knowledge Representation Models for Recommendation Purpose. One model is for Domain Knowledge, second model is for ontology. This system is a conceptual prediction model which is used to integrate the web usage and domain knowledge. It is used to form a

weighted semantic network of frequently viewed pages or terms. This system is proposed to predict next Web Page requests of users through querying the knowledge.

**REFERENCES**

[1] ThiThanh Sang Nguyen, Hai Yan Lu, Jie Lu, “Web-page Recommendation based on Web Usage and Domain Knowledge”, Knowledge and Data Engineering IEEE Transaction on September 2014  
 [2] B. Liu, B. Mobasher, and O. Nasraoui, “Web Usage Mining, in Web Data Mining: Exploring Hyperlinks, Contents, and

- Usage Data”, Ed.: Springer-Verlag Berlin Heidelberg, 2011, pp. 527-603.
- [3] S. Grimm, A. Abecker, J. Völker, and R. Studer, "Ontologies and the Semantic Web" , in Handbook of Semantic Web echnologies,J. Domingue, D. Fensel, and J. A. Hendler, Eds.: Springer-Verlag Berlin Heidelberg, 2011, pp. 507-580.
- [4] B. Liu "Information Retrieval and Web Search", in Web Data Mining: Exploring Hyperlinks, Contents, and Usage Data, B. Liu, Ed.: Springer-Verlag Berlin Heidelberg, 2011, pp.183-236.
- [5] 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, SA: IGI Global, 2005, pp. 276 - 306.
- [6] Brindha.S and Sabarinathan.P "An effective search on web log from most popular downloaded content",International Journal of Distributed and Parallel Systems (IJDPS) Vol.5, No.1/2/3, May 2014.
- [7] Mahendra Thakur, Yogendra Kumar Jain, GeetikaSilakari, "Query based Personalization in Semantic Web Mining", (IJACSA) International Journal of Advanced Computer Science and Applications,Vol. 2, No.2, February 2011
- [8] Sneha Y. S, G Mahadevan, and Madhura Prakash, "A Personalized Product Based Recommendation System Using Web Usage Mining and Semantic Web", International Journal of Computer Theory and Engineering Vol. 4, No. 2, April 2012.
- [9] C.Ramesh1, Dr. K. V. Chalapati Rao, Dr. A.Goverdhan, "A Semantically Enriched Web Usage Based Recommendation Model", International Journal of Computer Science & Information Technology (IJCSIT) Vol 3, No 5, Oct 2011.
- [10] V. SithaRamulu, Ch. N. Santhosh Kumar, K. Sudheer Reddy,"A Study of Semantic Web Mining: Integrating Domain Knowledge into Web Mining", International Journal of Soft Computing and Engineering (IJSCE)ISSN: 2231-2307, Volume-2, Issue-3, July 2012.
- [11] Honghua Dai and Bamshad Mobasher, “Integrating Semantic Knowledge with Web Usage Mining for Personalization”, School of Computer Science, Telecommunication, and Information Systems.