International Journal of Research
eISSN: 2348-6848 & pISSN: 2348-795X Vol-4 Special Issue-6
National Conference on Innovations in Information and Communication Technology
Held on 17-03-2017, Organized by Department of Information Technology, Meenakshi Sundararajan Engineering College,
363,Arcot Road Kodambakkam, Chennai 600024, India

Improving Web Navigation Usability by Maintaining a Personalized Recommendation System

Mr.R. Umanesan[1], S. Banumathi[2], K. Dhivya[3], P. Haripriya[4], S. Kayalvizhi[5]
[1]Asst.Professor and [2,3,4,5]B.Tech Valliammai Engineering College

Abstract

Recommendation system can take the benefits of semantic reasoning potentiality which is used to overcome the common constraints of the current system and hence to improve the quality of recommendations. In this paper, we present a personalized recommendation system, a system that makes use of representations of items and user-profiles based on ontology to facilitate the semantic applications with personalized services. The recommender uses domain based inference method to mould the user's interests in a more effective and accurate way to enhance personalization. Web Usage Mining plays a significant role in augmenting web personalization and recommendations. We propose an efficient personalized recommendation system based on domain ontology and Web Usage Mining. The primary approach is to extract features from web documents and to structure related concepts. Then construct ontology for the website using the characteristics extracted from the documents. According to the semantic similarity of web documents, cluster them into diverse semantic ideas where the diverse ideas imply different preferences. The proposed approach incorporates semantic knowledge into Web Usage Mining and personalization processes.

Keywords

Web Usage Mining, ontology, Webpage recommendation, user-item subgroups, semantic search.

1. Introduction

Web page recommendation is a method of proactively discovering and recommending appropriate Web services to end users. Widely engaged techniques are Collaborative Filtering (CF), Content-based approach or hybrid of both. Webpage recommendation has become increasingly popular, and is shown as links to related stories, related books, or most viewed pages at websites. This paper presents a method to provide enhanced Webpage recommendation based on Web usage and domain knowledge.

2. Architecture Diagram

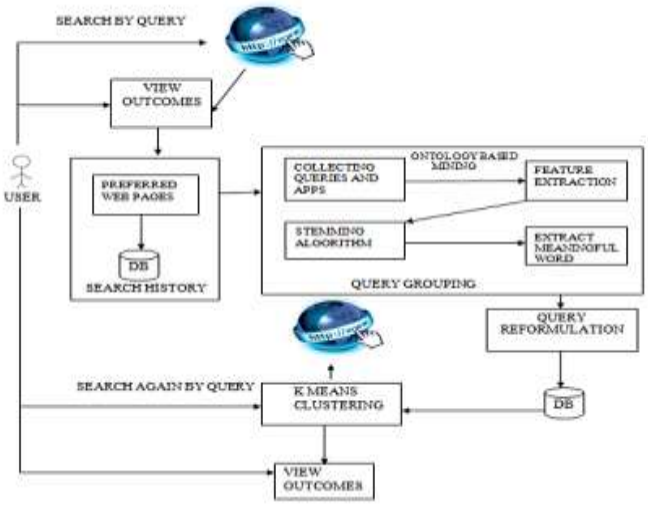

Figure 1. Architecture Diagram

International Journal of Research
eISSN: 2348-6848 & pISSN: 2348-795X Vol-4 Special Issue-6
National Conference on Innovations in Information and Communication Technology
Held on 17-03-2017, Organized by Department of Information Technology, Meenakshi Sundararajan Engineering College,
363,Arcot Road Kodambakkam, Chennai 600024, India

3. Architecture Description

In figure 1the user initially performs the normal web search and the related web applications are displayed on the browse. The specific web application which is of users interests is stored in the search history. Once the user interested applications are collected, query clustering and classification is performed, wherein the keyword and the mine words are extracted and stored in the personalized database under the corresponding user login. Now the Stemming algorithm is applied on the extracted mine words, which removes the grammatical representations of the word and stores the root word in the database. Now when the user searches again using the same keyword, the recommendation system based result i.e. the corresponding link and its sub links for the keyword are displayed right up front using the k means clustering technique.

4. Algorithm

Stemming Algorithm

Stemming is one technique to provide ways of finding morphological variants of search terms. Used to improve retrieval effectiveness and to reduce the size of indexing files. Stemming can have marked effect on the size of indexing files, sometimes decreasing the size of file as much as 50 percent.

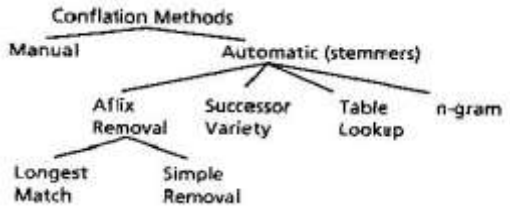

Figure 2. Classification of stemming

Affix removal algorithms as shown in figure 2 is an automatic stemmer which removes suffixes and/or prefixes from terms leaving a stem

1. If a word ends in “ies” but not ”eies” or ”aies ”, Then “ies” -> “y”

2. If a word ends in “es” but not ”aes” , or ”ees ” or “oes”, Then “es” -> “e”

3. If a word ends in “s” but not ”us” or ”ss ”, Then “s” -> “NULL”

Conditions:

1. The measure, denoted m, of a stem is based on its alternate vowel-consonant sequences

Measure	Example
M=0	TR, EE, TREE, Y, BY
M=1	TROUBLE, OATS, TREES, IVY
M=2	TROUBLES, PRIVATE, OATEN

2.*<X> ---the stem ends with a given letter X

3.*v*---the stem contains a vowel

4.*d --- the stem ends in double consonant

5.*o --- the stem ends with a consonant-vowel consonant, sequence, where the final consonant is not w, x or y

Suffix conditions take the form: (current_suffix == pattern)

K-means Algorithm

K-means clustering is a method of cluster analysis which aims to partition n observations into k clusters in which each observation belongs to the cluster with the nearest mean. K-means minimizes within-cluster point scatter:

$$W(C) = \frac{1}{2}\sum_{k=1}^{K}\sum_{C(i)=k}\sum_{C(j)=k}\left\|x_i - x_j\right\|^2 = \sum_{k=1}^{K} N_k \sum_{C(i)=k}\left\|x_i - m_k\right\|^2$$

Where

m_k is the mean vector of the k^{th} cluster

N_k is the number of observations in k^{th} cluster.

International Journal of Research
eISSN: 2348-6848 & pISSN: 2348-795X Vol-4 Special Issue-6
National Conference on Innovations in Information and Communication Technology
Held on 17-03-2017, Organized by Department of Information Technology, Meenakshi Sundararajan Engineering College,
363,Arcot Road Kodambakkam, Chennai 600024, India

K is positive integer number.

Simply speaking k-means clustering is an algorithm to classify or to group the objects based on attributes/features into K number of group. The grouping is done by minimizing the sum of squares of distances between data and the corresponding cluster centroid.

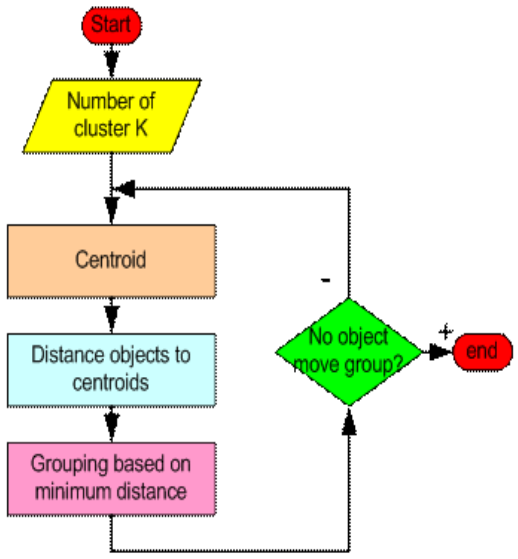

Figure 3. Flow of K-means algorithm

5. Module Description

Creating Search history

Any personal documents for instance browsing history and emails on a user's computer could be the resource for user profiles. This concentrate on repeated terms restricts the dimensionality of the document set, which additionally grant a detailed description of user's interest. By using this module the search engine can aware of the user's session and potential tailor of the user according to their needs.

Query classification

User's queries can be classified into various query clusters. Concept-based user profiles are used in the clustering method to accomplish personalization effect. Most similar pair of concepts or query nodes is merged. Every single query given by every individual user is treated as separate node and each query with a user identifier. We carry out the grouping in a similar dynamic manner.

Query reformulation

To certify that all query groups contains strongly relevant and correlated queries, it is essential to have a appropriate relation between the present query groups. We assume that users commonly give much alike queries. The search history of a huge amount of users contains signals about query relevance, such as which queries tend to be issued closely together. This captures the correlation among queries repeatedly leading to clicks on similar URLs.

K means 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).

6. Bibliography

[i].Guosheng Kang, Mingdong Tang, Jianxun Liu, Xiaoqing(Frank) Liu, and Buqing Cao, "Diversifying Web Service Recommendation Results via Exploring Service Usage History "*IEEE Transactions on Services computing*, (2015).

[ii].Jiajun Bu, Xin Shen, Bin Xu, Chun Chen, Xiaofei He, Deng Cai,"Improving Collaborative Recommendation via User-Item Subgroups", *IEEE Transactions on knowledge and data engineering*, vol-28, (2016).

[iii].Ruili Geng and Jeff Tian,"Improving web navigation usability by comparing actual usage and anticipated usage",*IEEE Transactions on human-machine systems*, Vol-45,(2015).