2009 International Conference on Information Management and Engineering

Deriving Semantic Sessions from Semantic Clusters

Kuala Lumpur, Malaysia
April 03-April 05
ISBN: 978-0-7695-3595-1

Banafsheh Safarkhani
Mojde Talabeigi
Mehran Mohsenzadeh
Mohammad Reza Meybodi

DOI Bookmark: http://doi.ieeecomputersociety.org/10.1109/ICIME.2009.131

A important phase in any web personalization system is transaction identification. Recently a number of researches have been done to incorporate semantics of a website in representation of transactions. Building a hierarchy of concepts manually is time consuming and expensive. In this paper we intend to address these shortcomings. Our contribution is that we introduce a mechanism to automatically improve the representation of the user in the website using a comprehensive lexical semantic resource and semantic clusters. We utilize Wikipedia, the largest encyclopedia to date, as a rich lexical resource to enhance the automatic construction of vector model representation of user sessions. We cluster web pages based on their content with Hierarchical Unsupervised Fuzzy Clustering algorithms ,are effective methods, for exploring the structure of complex real data where grouping of overlapping and vague elements is necessary. Entries in web server logs are used to identify users and visit sessions, while web page or resources in the site are clustered based on their content and their semantic. Theses clusters of web documents are used to scrutinize the discovered web sessions in order to identify what we call sub-sessions. Each subsession have consistent goal. This process engendered to improving deriving semantic sessions from web site user page views. Our experiments show that proposed system significantly improves the quality of web personalization process.

Index Terms:
Semantic vectors, Semantic sub-session, Semantic cluster, Wikipedia

Citation:
Banafsheh Safarkhani, Mojde Talabeigi, Mehran Mohsenzadeh, Mohammad Reza Meybodi, "Deriving Semantic Sessions f