J Intell Inf Syst (2016) 47:165–192
DOI 10.1007/s10844-016-0398-3

Improving the prediction of page access by using semantically enhanced clustering

Erman Sen[1] **· I. Hakki Toroslu**[1] **· Pinar Karagoz**[1]

Received: 22 July 2015 / Revised: 24 January 2016 / Accepted: 14 March 2016 /
Published online: 20 April 2016

Abstract There are many parameters that may affect the navigation behaviour of web users. Prediction of the potential next page that may be visited by the web user is important, since this information can be used for prefetching or personalization of the page for that user. One of the successful methods for the determination of the next web page is to construct behaviour models of the users by clustering. The success of clustering is highly correlated with the similarity measure that is used for calculating the similarity among navigation sequences. This work proposes a new approach for determining the next web page by extending the standard clustering with the content-based semantic similarity method. Semantics of web-pages are represented as sets of concepts, and thus, user session are modelled as sequence of sets. As a result, session similarity is defined as an alignment of two sequences of sets. The success of the proposed method has been shown through applying it on real life web log data.

Keywords Ontology · Concept set similarity · Session similarity · Sequence alignment

1 Introduction

The World Wide Web is emerging as a major environment for business transactions and user-organization interactions such as e-Commerce and e-Government applications, since it is convenient, fast and inexpensive. However, since the Web is a huge collection of resources, users often suffer from information overload. Therefore, it becomes more important to be able to recognize the web users' behaviours on the web sites, and improve the performances of web sites and the qualities of user experiences on these sites by using

✉ Pinar Karagoz
karagoz@ceng.metu.edu.tr

[1] Computer Engineering Department, Middle East Technical University, 06800, Ankara, Turkey