Editorial for the special issue on *advanced information systems for the Web*

Alex Delis · X. Sean Wang

1 Editors' introduction

For more than twenty years, we are witnesses to novel approaches, intense development and continuous deployment of Web Information Systems that by now empower nearly all aspects of daily life. The fast-paced advancement of the Web and its supporting systems and tools have been assisted by efficient methods for accessing data placed in aggregations of web-servers, databases, and collections, by novel techniques for the provision of high-availability for data, and finally, by policies for the orderly operation of the co-existing and collaborative virtual ecosystems. Global virtual communities are being formed on the web and systems and tools supporting their operations constitute a vibrant and rather rich source for both contemporary research and development on the World-Wide Web.

The *WISE* series of Conferences have provided for almost fifteen years a global forum in which novel ideas and proposals are disseminated. The *13th International Conference of Web Information Systems Engineering (WISE'12)* was held in Paphos, Cyprus in November 2012. The conference attracted 194 submissions of which 44 were finally presented.

In this special issue, we offer extended versions of selected works from the *WISE'12* conference. Six papers make up this issue of the *World Wide Web Journal* and examine a diverse range of topics in contemporary Web research.

The first paper, entitled "A Web-based Approach to Data Imputation" outlines `WebPut`, an approach that addresses the data imputation problem. As data incompleteness is widespread in *WWW* information systems, `WebPut` serves as an effective information extraction method to offer plausible answers using either syntactic patterns or common context terms; they are used as auxiliary information to overcome the problem of incomplete

A. Delis (✉)
University of Athens, 15784 Athens, Greece
e-mail: ad@di.uoa.gr

X. S. Wang
Fudan University, Shanghai, China
e-mail: xywangCS@fudan.edu.cn

data on the Web. Experimentation with pragmatic data sets point out to the effectiveness of the proposed method as well as the efficiency of `WebPut` algorithms.

The second paper, entitled "Like-minded Communities: Bringing the Familiarity and Similarity Together" exploits a social networking movie–rating dataset to experimentally show that people who are interested in an item are socially "better connected" than the overall graph. Based on this, the authors propose a method whose main goal is for locate communities in a social graph demonstrating like-mindedness. Small-size tight groups with common interests are first identified and are subsequently used as the pillars for creating broader groups. The paper shows that communities formed using this approach demonstrate higher similarity if compared with social groups solely identified based on interaction information of their members.

The third paper entitled "Efficient Execution of Web Navigation Sequences" proposes methods that jointly seek to create a web navigation component capable of efficiently executing web navigation sequences. The proposed techniques detect both specific elements and scripts that are part of pages accessed during a navigation sequence. Those page elements and scripts deemed required for the correct execution of a navigation sequence make up the means to identify parts of pages that are no longer considered in the execution sequence. The approach yields optimized navigation sequences that run faster while consuming less resources than baseline competing methods.

The forth paper entitled "Improving the Performance of Pipelined Query Processing with Skipping" investigates the effect of inverted index skipping on the performance of pipelined query processing for Web search engines. The proposed method is based on *Max-Score* pruning within the pipelined query processing as well as on a new term assignment heuristic. Experimental results indicate gains over conventional approaches and lead to additional enhancement when it comes to dynamic load balancing, intra-query concurrent processing and the use of a hybrid combination of pipelined and non-pipelined query execution.

The fifth paper entitled "Performance vs. Freshness in Web Database Applications" outlines a novel method for efficiently materializing dynamic web-pages in e-commerce environments. The key requirement of user satisfaction gets quantified by response time and content freshness especially in heavy workloads settings. The proposed method considers dependencies among content fragments in web dynamic pages. Semantic-based data freshness metrics are used to capture content dependencies and algorithms that balance quality-of-service and quality-of-data freshness are suggested. Evaluation with a prototype ascertain gains over existing methods when it comes to server-side response time, scalability as well as data freshness.

The sixth paper entitled "How Useful are Social Features for Learning to Rank YouTube Videos?" investigates the social characteristics that are associated with the popular videos found on `YouTube`. The paper considers both raw and derived social features and examines the effectiveness these features have on video retrieval. The paper suggests an new feature selection strategy based on the *Maximal Marginal Relevance* method. Using popular and rare queries to annotate hundreds of query-video pairs on `YouTube`, the proposed approach demonstrates much promise in offering improved retrieval performance for both popular and rare queries.

We wish to thank the authors for their contributions in this special issue of the *World–Wide Web Journal*, the Program Committee (PC) Members and the external reviewers of the *WISE'12* Conference. Last but not least, we are indebted to 14 reviewers who diligently offered comments and suggestions for the considered manuscripts and helped us with the selection of articles for this special issue.

Alex Delis and X. Sean Wang
Editors