Special Issue on Mining Associations and Patterns from the Semantic Data

PREFACE

Advancements in technologies for extracting metadata from unstructured content, and the broadening adoption of Web publishing best practices advocated by initiatives like the Linked Open Data and Semantic Web communities, has led to an large amount of metadata associated with Web content, sensor social network data. In order to maximally leveraging the increasing amount of metadata, sophisticated techniques for reasoning about links in metadata in a way that will help expose interesting relationships or associations, expressed as meaningful paths, subgraphs and patterns. Techniques are needed for ascribing meaning or relative importance to associations, for determining the underlying dynamics in these relationships like cause and effect or influence, and for building more advanced search and ranking models based on associations. The papers in this special section of the *International Journal on Semantic Web and Information Systems* go in this direction. This issue features a collection of three papers, selected from 18 submissions that represent recent advances in Mining and Discovering Associations and Patterns from Semantic Data.

In the first paper, Learning to Rank Complex Semantic Relationships" by Na Chen and Viktor K. Prasanna the authors address the problem of ranking the results of a search for complex semantic associations (indirect connectivity relationships) in an RDF graph. The approach is preference-based and uses a feature vector to describe various statistical and semantic features of a semantic association. The authors propose a learning-to-rank algorithm for featured-based preference learning which constructs a personalized ranking function for a user. The ranking function is then used to produce ranked lists of semantic association search results for subsequent queries from the user.

The second paper, "More Spreaders But Less Influence: An Analysis of Indirect Influence on Twitter" by Xin Shuai, Ying Ding, Jerome Busemeyer, Shanshan Chen, Yuyin Sun and Jie Tang, proposes a model for studying indirect influence between users that are not directly connected in social networks. Their investigation was done in the context of Twitter networks with indirect influence quantified as the increased probability of retweeting messages. The authors investigate the impact of parallel spreaders on the intensity of indirect influence and observed that globally, retweeting probability increases non-monotonically, validating the theory of complex contagion. However, the authors also observe some local drops in intensity of influence which they account for using a quantum cognition based probabilistic model.

The third paper, "Diversifying Search Results through Pattern-based Subtopic Modeling" by Wei Zheng, Hui Fang, Hong Cheng and Xuanhui Wang addresses the problem of improving the quality of search results by diversifying results and eliminating redundancy. The underlying goal of the authors' approach is to enable search results go beyond the immediate topic of a given query to include results of subtopics of a query. The diversification model is a pattern-based framework where patterns are extracted using maximal frequent itemset mining techniques. Subtopics are modeled as groups of patterns, grouped using a profile-based clustering technique. The output of the subtopic modeling is used as the basis for search result diversification.

We would like to thank all of the authors of submitted papers to this special section for their high-quality contributions. We also thank the referees for their time and effort and their valuable suggestions. We are grateful to Professor Amit Sheth, the Editor-in-Chief of *IJSWIS* for his strong support for this special issue.

- Ying Ding, Kemafor Anyanwu, Jie Tang (http://keg.cs.tsinghua.edu.cn/jietang/), Philip Yu (http://www.cs.uic.edu/PSYu/)