CrossMark

Guest editorial: special issue on advanced data mining and applications

Quan Z. Sheng[1] · Yongrui Qin[2]

1 Introduction

A growing attention has been paid to the study, development and application of data mining. As a result, there is an urgent need for sophisticated techniques and tools that can handle new subfields of data mining, e.g., smartphone and social network data mining, spatial data mining in the context of spatial-temporal characteristics, streaming data mining, green computing data mining and biomedical data mining. Our expertise in data mining also has to be expanded to new application areas such as the Internet of Things.

This special issue showcases some recent research results, which were presented at the 12th International Conference on Advanced Data Mining and Applications (ADMA 2016) held in Gold Coast, Australia, on 12–15 December 2016. This has been a premier annual event on research and applications of data mining. ADMA 2016 brought together the experts on data mining from around the world, and providing a leading international forum for the dissemination of original research findings in data mining. Attendees included academics and practitioners from around the world, and this issue includes a set of selective best research papers that were presented at the conference and extended after the conferences based on the feedback from the conference. All the papers invited for this special issue have undergone two rounds of rigorous review process. Based on the reviewers' feedback, six papers were selected for publication.

✉ Quan Z. Sheng
  michael.sheng@mq.edu.au

[1] Department of Computing, Macquarie University, Sydney, Australia

[2] School of Computing and Engineering, University of Huddersfield, Huddersfield, UK

2 In this special issue

This special issue begins with a paper by Duan et al. entitled "Mining Distinguishing Customer Focus Sets From Online Customer Reviews". This paper looks at effective online shopping decision support. Due to the typically large amount of online customer reviews, how to assist customers in gaining personalized and useful information of such online reviews has become interesting and challenging. This paper investigates a novel data mining problem of mining distinguishing custom focus sets from customer reviews. A novel mining method, dFocus-Miner, is proposed to effectively obtain interpretable ad user-friendly mined results on a large set of online customer reviews.

The second paper by Xu et al. entitled "A New Paradigm of Software Service Engineering in Big Data and Big Service Era" focuses on big data and service oriented engineering. It develops a new paradigm, named RE2SEP, to seamlessly integrate service oriented requirement engineering, domain oriented service engineering, and the development approach of software services. Based on RE2SEP, adaptive and personalized service solutions can be efficiently developed and implemented.

The third paper by Yao et al. entitled "WITS: An IoT-endowed Computational Framework for Activity Recognition in Personalized Smart Homes" studies smart home care and develops an end-to-end Web-based in-home monitoring system, named WITS, for convenient and efficient care delivery. It mainly leverages shared-structure dictionary learning and rule-based reasoning for continuous daily activity tracking and abnormal activity detection.

The next paper by Al-Maskari et al. entitled "Bio-inspired learning approach for electronic nose" particularly looks at improving performance of e-nose (electronic nose) in odour recognition, thereby extending human sensing ability dramatically. E-nose technology enables new ways to sense and monitor the environment for humans, where chemical analysts and odour molecules may exist. This paper proposes AORCM, which is inspired by neural circuits of the vertebrate olfactory system, to improve e-nose's stability and performance.

The fifth paper by Wang et al. entitled "Clustering by Differencing Potential of Data Field" studies hierarchical clustering in data field. The concept of data field is inspired by physical field that describes the interaction between particles in physics. Data field depicts the interacted relationships between data objects in a digitalized data space. This paper leverages the difference potential techniques to assist in the clustering process of data field to handle scenarios that are with high dimensions, complex distribution shapes and noise identification.

The last paper by Liu et al. entitled "Discovering Pan-correlation Patterns from Time Course Data Sets by Efficient Mining Algorithms" focuses on the efficient and effective discovery of correlation patterns in time-course data. Such research has potentially significant impact on many real-world applications, such as finance and health-care, where time-course changes are critical. This paper investigates the discovery of correlation patterns with a novel theory that obtains a generalized form of positive correlation patterns and with a novel single-scan algorithm for mining correlations.

to express our deepest gratitude to the referees who provided very useful and thoughtful feedback to the authors. Finally, our sincere thanks go to the Editor-in-Chief, Professor Schahram Dustdar, of Computing for his kind support, advice, and encouragements throughout the preparation of this special issue.