Editorial

Magdalini Eirinaki* and Jerry Zeyu Gao

Computer Engineering Department,
San Jose State University,
One Washington Square,
San Jose, CA 95192, 408-924-1000, USA
Email: magdalini.eirinaki@sjsu.edu
Email: jerry.gao@sjsu.edu
*Corresponding author

Iraklis Varlamis

Department of Informatics and Telematics,
Harokopio University of Athens,
Omirou 9, Athens, 17778, Greece
Email: varlamis@hua.gr

Biographical notes: Magdalini Eirinaki is an Associate Professor at the Computer Engineering Department of San Jose State University. Her research interests cover the areas of web mining and recommender systems and, in particular, on mining of social networks, aspect-based recommendations, interactive database exploration, and personalisation. She has over 45 publications and 1,900 citations in peer-reviewed international conferences and journals. She has served in the program committee of numerous conferences and as a reviewer in IEEE/ACM journals. She has served as the Track Chair of IEEE BigDataService 2015 and 2016, Special Tracks on Big Data Intelligence and Knowledge Engineering. She will also be in the editorial board of the new Springer series books titled *Advances in Learning, Analytics and Intelligent Computing Systems*.

Jerry Zeyu Gao is a Professor at the Department of Computer Engineering at San Jose State University. He has over 15 years of academic research and teaching experience and over ten years of industry working and management experience on software engineering and IT development applications, focusing on cloud computing and services, software/testing as a service, mobile sensor technologies, and smart cities. He has published over 180 publications in IEEE/ACM journals, magazines, international conferences and workshops. He has co-authored three published technical books and edited numerous books in software engineering, software testing, and mobile computing. In the last ten years, he has played as one of leaders in organising many international conferences and workshops as a conference co-chair, program co-chair, and workshop co-chair.

Iraklis Varlamis is an Assistant Professor at the Harokopio University of Athens, Department of Informatics and Telematics. His research interests are on the use of semantics in web mining tasks and his research work applies to knowledge management and mining in social networks. He has more than 90 publications and more than 1,300 citations in international conferences and journals. He co-edited a collective book on *Certification and Security in Health-Related Web Applications: Concepts and Solutions* and has co-organised the IEEE BigDataService 2015 and 2016 Special Tracks on Big Data Intelligence and Knowledge Engineering.

1 Introduction

The fields of computational intelligence and knowledge management have made significant advances over the past decades. The potential ability to derive intelligence from the analysis of raw data has been successfully applied to a broad array of (diverse) areas such as marketing, manufacturing, sciences, humanities, social media, etc. The success of such applications has been accompanied by a continuous evolution in the characteristics and size of data: they are more dynamic and spatiotemporal, distributed and volatile, while constantly increasing in size. These changes have created new challenges for researchers in the computational intelligence, machine learning, data mining and knowledge engineering fields that relate to the effective harvesting, storage, curating, integration, management, and analysis of big data.

In the digital-world era, social networks have contributed immensely to the creation of large amount of data, with various formats, types and meanings. Big data and social networks become interrelated in the modern computing agenda creating new opportunities and

challenges for data management, security, analytics, etc. and giving space to new applications and technologies that efficiently handle the huge amounts of data.

This special issue focuses on big data applications that extract knowledge from social networks and media and technological solutions and algorithms implemented over big data frameworks for facilitating knowledge management from big data. In particular, two of the featured articles focus on integrating data from various social media sources and extracting useful knowledge by processing the user-generated content, employing text mining and sentiment analysis techniques. The challenge of integrating data in a lower-level, that of NoSQL key-value stores, is addressed in the third featured article.

2 Articles

The first article entitled 'PaloPro: a platform for knowledge extraction from big social data and the news' by Makrynioti et al. presents a platform that aggregates textual content from social media and news-sites in different languages, analyses them using a series of text mining algorithms and provides advanced analytics to journalists and social media marketeers. The platform capitalises on the abundance of social media sources and the information they provide for persons, products and events. In order to handle the increasing load and be able to handle multilingual content, authors have adopted language independent techniques at all levels and from an engineering point of view, have designed a system that takes advantage of parallel distributed computing technologies and cloud infrastructure.

The second article 'Optimising column family for OLAP queries in HBase' by Yang et al. presents a new evolutionary algorithm which aims in grouping HBase columns into families in an optimal way, taking into account the underlying query load. The algorithm seeks for the optimum column family schema for the given user queries using a Map-Reduce for learning rates selection of the designed algorithm. Empirical evaluation on a real dataset, which contains millions of rows of aggregated tracking data and a million OLAP queries submitted by real users, showed that the use of the optimised column family schema improves the reading performance significantly and is slightly worse than the theoretically optimal solution, which however cannot be used in practice due to the current limitations of HBase.

The third article 'A knowledge-based integrated framework for increasing social management intelligence' by Dimitrakopoulos et al. presents an integrated platform, capable of:

a aggregating at real time huge amounts of data from social media and online publications

b intelligently processing this data through sophisticated sentiment analysis algorithms

c proposing novel structured approaches for the exploitation of these data through applying technologically innovative and efficient marketing practices.