Guest Editorial: Big Data Analytics and the Web

Quan Z. Sheng, *Member, IEEE*, Athanasios V. Vasilakos, Qi Yu, *Member, IEEE*, and Lina Yao, *Member, IEEE*

LAST few years have seen the rapid increase of sheer amount of data produced and communicated over the Web. Such Big Data are generated from all kinds of sources and applications such as social network services, cloud services, knowledge bases, and intelligent terminals, and often in a wide variety of formats such as unstructured, semi-structured, and structured. A particular recent trend around the web is to connect and communicate between billions of physical objects (also called "Things"), i.e., Web of Things (WoT). WoT offers the capability of integrating both physical and virtual worlds and massive volumes of real-time data are expected to be produced by these connected things and their associated sensors.

While it is widely believed that Big Data holds the potential to revolutionize many aspects of our modern society (e.g., smart cities), many technical challenges need to be addressed before this potential can be realized. Indeed, Big Data requires a revisit of data analysis systems in fundamental ways at all stages from data acquisition and storage to data transformation and interpretation. Services should be ideally provisioned in a way that speeds up data processing, scales up with data volume, improves the adaptability and extensibility over data diversity and uncertainties, and finally turns low-level data into actionable knowledge towards better understanding and manipulation of Big Data.

This special section aims at presenting the latest developments, trends, and solutions of Big Data analytics on the web. There were 16 submissions and two papers were selected to be included in the first part of this special section after several rounds of rigorous review by the guest editors and invited reviewers. The two papers cover important topics and present some of the key directions in this vibrant and rapidly expanding area of research and development. We hope that the set of selected papers provides the community with a better understanding of the current directions and areas to focus in future, and inspires your own work.

The first paper by Zhao, Liu, Zhou, Jian and Yang, "LS-AMS: An Adaptive Indexing Structure for Realtime Search on Microblogs", considers the challenging issues of real-time microblog search where new microblogs are created at a fast pace and user query requests constantly change. The authors propose an efficient Log-Structured index structure with Adaptive Merging Strategy (LS-AMS) that can improve the query performance and improve the self-adaptability of microblog search in dynamic environments. The proposed LS-AMS structure includes an inverted index buffer and a sequence of dynamically adjustable index packages. These packages manage their indexes by exploiting an adaptive merging strategy to reduce the merging overhead and improve the query performance.

The second paper by Cheng and Kotoulas, "Scale-Out Processing of Large RDF Datasets", proposes efficient methods for processing RDF data using dynamic data re-partitioning for rapid analysis of large, distributed RDF datasets. The authors propose a two-tier index architecture on each of the distributed computation nodes, including a light-weight primary index to keep loading time low and a series of dynamic, multi-level secondary indexes to decrease or remove inter-machine data movement for subsequent queries that contain the same graph patterns. The authors also propose to replace certain secondary indexes with distributed filters to decrease the memory consumption. Using their approach, it is possible to improve loading speeds while remaining competitive in terms of performance while dealing with large distributed RDF data. For example, the authors demonstrate that their approach can load a dataset of 1.1 billion triples at a rate of 2.48 million triples per second and provide good performance to RDF-3X and 4store for expensive queries.

ACKNOWLEDGMENTS

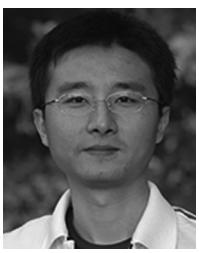
The authors thank all the authors for considering this special section as an outlet to publish their research results in the area of Big Data analytics and the web. They also would like to thank the referees who provided very useful and thoughtful feedback to the authors. Finally, they express their gratitude to the editor-in-chief, Professor Qiang Yang, for his kind support, advice, and encouragement throughout the preparation of this special section.

- *Q.Z. Sheng is with the University of Adelaide, Adelaide, SA, Australia. E-mail: michael.sheng@adelaide.edu.au.*
- *A.V. Vasilakos is with the Luleå University of Technology, Luleå, Sweden. E-mail: th.vasilakos@gmail.com.*
- *Q. Yu is with the Rochester Institute of Technology, Rochester, NY. E-mail: qi.yu@rit.edu.*
- *L. Yao is with the UNSW, Sydney, NSW, Australia. E-mail: lina.yao@unsw.edu.au.*

For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org, and reference the Digital Object Identifier below.
Digital Object Identifier no. 10.1109/TBDATA.2016.2518778

Quan Z. Sheng received the PhD degree in computer science from the University of New South Wales (UNSW Australia) in 2006. He is a professor and the head in Advanced Web Science Research Group, School of Computer Science, University of Adelaide. His research interests include Internet of Things, Web of Things, big data analytics, service-oriented computing, distributed computing, Internet computing, and pervasive computing. He received Australian Research Council (ARC) Future Fellowship in 2014, Chris Wallace Award for Outstanding Research Contribution in 2012, and Microsoft Research Fellowship in 2003. He is the author of more than 220 publications. He is a member of the ACM and IEEE.

Athanasios V. Vasilakos is currently a professor with the Luleå University of Technology, Sweden. He has authored or coauthored over 400 technical papers in major international journals and conferences. He is a author/coauthor of five books and 20 book chapters in the area of communications. He has served as a general chair/technical program committee chair for many international conferences. He served or is serving as an editor and/or guest editor for many technical journals, such as the *IEEE Transactions on Network and Service Management*, the *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, the *IEEE Transactions on Information Technology in Biomedicine*, the *IEEE Transactions on Computers*, the *ACM Transactions on Autonomous and Adaptive Systems*, *IEEE JSAC* special sections of May 2009, January 2011, and March 2011, *IEEE Communications Magazine*, *ACM/Springer Wireless Networks (WINET)*, and *ACM/Springer Mobile Networks and Applications (MONET)*. He is the founding editor-in-chief of the *International Journal of Adaptive and Autonomous Communications Systems (IJAACS)* and the *International Journal of Arts and Technology (IJART)*. He is a general chair of the Council of Computing of the European Alliances for Innovation.



Qi Yu received the PhD degree in computer science from Virginia Polytechnic Institute and State University (Virginia Tech). He is an associate professor and graduate program director in the Department of Information Sciences and Technologies at the Rochester Institute of Technology. His current research interests lie in the areas of service computing, data mining, and machine learning. His publications have mainly appeared in well-known journals (e.g., *VLDB journal*, *IEEE TKDE*, and *ACM TWEB*) and conference proceedings (e.g., ICSOC and ICWS). He is a guest editor of the *IEEE TSC* special section on service query models and efficient selection. He frequently serves as a program committee member on service computing and database conferences (e.g., IEEE Cloud, WISE, CollaborateCom, ICSOC, and IRI). He is also a reviewer for various journals (e.g., the *VLDB Journal*, *ACM TOIT*, and *ACM TWEB*). He is a member of the IEEE.

Lina Yao received the PhD degree in computer science from the University of Adelaide in 2014. She is currently a lecturer at the School of Computer Science and Engineering, University of New South Wales (UNSW Australia). Her research interests include Internet/Web of Things analytics, web mining, service computing, and ubiquitous and pervasive computing. She received an Australian Research Council (ARC) Discovery Early Career Researcher Award (DECRA) in 2015. She is the author of more than 30 publications. She is a member of the IEEE.