Guest Editorial

Special Issue on Big Data Inspired Data Sensing, Processing and Networking Technologies

Big data is the current big thing in computing after cloud computing and it is in par with software defined network (SDN) in communications and networks field. The phrase "big data" refers to large, diverse, complex, typically distributed data sets generated from various sources such as sensors, emails, social network websites, online transactions, and/or all other digital sources available today and in the future. It is evident that big data is making big impact to diverse walks of our everyday life, ranging from increasing understanding of digital system or human/social behaviours/processes to promoting economic growth and eventually to improving human being's health and quality of life.

A long-term strategy to address various challenges introduced or inspired by big data shall be a systematic one that brings together the three key aspects of a big data system, namely, data sensing, data processing and data communications. This special issue of *Elsevier Ad Hoc Networks* create a forum for researches, developers and practitioners from both academia and industry to publish the key results and to disseminate the state-of-the-art concepts and techniques in all aspects of big data advancement.

In total, 18 papers have been submitted to this special issue. After a rigorous reviewing process, we accepted 11 of the submissions to appear in this special issue. The accepted papers cover new research and developments in big data inspired technologies on sensor and ad hoc networks, big data system, cyber-physical system, MapReduce application, Cloud computing, security and trust, and energy-efficient routing.

The paper *New Efficient Velocity-Aware Probabilistic Route Discovery Schemes for High Mobility Ad hoc Networks*, by Bani Khalaf et al., considers the velocity vector probabilistic route discovery while broadcasting in MANETs. Specifically, this study proposes two new velocity aware probabilistic route discovery models, which can exclude unstable nodes while constructing routes between the source and its destination. Simulations confirm the superiority of the proposed schemes against their counterparts in terms of RREQ packet overhead and link stability.

In the paper entitled *Ensembles of Incremental Learners to Detect Anomalies in Ad Hoc Sensor Networks*, Bosman et al. present a general-purpose, online learning, decentralized anomaly detection framework that includes a heterogeneous set of local anomaly detection algorithms in Wireless Sensor Networks (WSNs). Results show that the ensemble performs better than each individual decentralized classifier and that it can match the performance of the online alternatives, thus demonstrating that the proposed approach is a viable solution to detect anomalies, even in environments with little a priori knowledge.

The paper *An Empirical Study on Implementing Highly Reliable Stream Computing Systems with Private Cloud*, by Liu et al., defines reliability as an architectural metric for cloud based business services. A methodology was designed to use the defined metrics to evaluate cloud based services components. The solution was used to resolve business challenges in high volume transactions in Chinese top banks. Novel reliability analysis techniques, queuing theory, and software rejuvenation management techniques were combined to build a framework for supporting stream data with low latency and fault tolerance.

In the paper entitled *A Trustworthy Access Control Model for Mobile Cloud Computing Based on Reputation and Mechanism Design*, Lin et al. propose a reputation and mechanism design based trustworthy access control model (RMTAC) to provide secure and privacy-aware big data access control in Mobile Cloud Computing (MCC). The RMTAC integrates the access control scheme with Vickrey-Clark-Groves (VCG) based adaptive reputation mechanism, the distributed multi-level security scheme and the hierarchical key management protocol to provide secure access control and defend against the internal attacks in MCC.

The paper *Towards a Big Data System Disaster Recovery in a Private Cloud*, by Chang et al., proposes a "multi-purpose" approach that allows data to be restored to multiple sites with multiple methods to ensure the organization recovers a very high percentage of data close to 100%. This paper has conducted various experiments to test the Disaster Recovery process for the Big Data and investigate the results used by different methods. The use of multi-purpose approach can be adopted in data centers for hosting Private Clouds or Big Data services to ensure the restoration of a large quantity and volume of data as efficient as possible.

In the paper entitled *GreeDi: An Energy Efficient Routing Algorithm for Big Data on Cloud*, Baker et al. propose a network-based routing algorithm to find the most energy efficient path to the cloud data centre for processing and storing big data. The algorithm is formalized by the situation calculus. The linear, goal and dynamic programming approaches are used to model the algorithm. The algorithm is evaluated against the baseline shortest path algorithm with minimum number of nodes traversed, using a real Italian ISP physical network topology.

In the paper entitled *A Big Data Inspired Chaotic Solution for Fuzzy Feedback Linearization Model in Cyber-Physical Systems*, Liu et al. developed a Cyber Physical System (CPS) model in light of fuzzy feedback linearization. The chaotic time prediction algorithm is applied to solve the chaotic control problem in CPS. The developed algorithm takes both tracking control problem and synchronization control problem into account. The numerical results suggest that the developed method is feasible and efficient in tracking control and synchronizing of two chaotic CPS.

The paper *Efficient Subspace Skyline Query based on User Preference using MapReduce*, by Li et al., introduces a system model that can support subspace skyline query in mobile distributed environment. An efficient algorithm for processing the Subspace Skyline Query using MapReduce (SSQ) is also presented in order to obtain the meaningful subset of points from the full set of skyline points in any subspace. Extensive experiments on real and synthetic data indicate that the proposed algorithm is efficient and the pruning strategy can further improve the efficiency of the algorithm.

In the paper entitled *A MapReduce based Parallel Niche Genetic Algorithm for Contaminant Source Identification in Water Distribution Network*, Hu et al. study a Contamination Source Identification (CSI) in a water distribution network with limited number of monitoring sensors. They developed a MapReduce based Parallel Niche Genetic Algorithm (MR-PNGA) that is not only able to achieve high identification accuracy but also to explore the cloud resources for performance improvement. The accuracy and efficiency of MR-PNGA is extensively validated on an 8-server cluster.

In the paper entitled *Intelligent Photovoltaic Monitoring based on Solar Irradiance Big Data and Wireless Sensor Networks*, Hu et al. utilize Wireless Sensor Networks (WSN) to efficiently deliver the monitoring data of the photovoltaic (PV) modules from power stations to the monitoring center located in Cloud data center. In order to evaluate the performance of the proposed methods, a comprehensive experimental platform is set up. The experimental results show that the predicted values match well with the theoretical values of power generation.

In the paper entitled *Distributed In-Memory Vocabulary Tree for Real-time Retrieval of Big Data Images*, Duan et al. develop a new scheme of the distributed in-memory vocabulary tree based on MapReduce model for massive image training and retrieval. This distributed vocabulary tree strategy can support massive image training in memory. Therefore, a similar image can be retrieved in a distributed manner based on MapReduce model. Besides, the training time and memory overhead of the proposed scheme are analyzed in detail.

Acknowledgements. We would like to thank all the authors, reviewers and editors involved in the elaboration of this special issue. We are especially grateful to Prof. Ian F. Akyildiz, Editor-in-Chief of the *Elsevier Ad Hoc Networks* for approving this special issue and for his help along the process of its preparation.

Guest Editors

Jia Hu
Department of Mathematics and Computer Science
Liverpool Hope University, UK
Email: huj@hope.ac.uk

Kun Yang
School of Computer Science & Electronic Engineering
University of Essex, UK
Email: kunyang@essex.ac.uk

Chirag Warty
Ahilya Technologies, Los Angeles, California, USA
Email: chiragwarty@ieee.org

Ke Xu
Department of Computer Science, Tsinghua University, Beijing, China
Email: xuke@tsinghua.edu.cn