Special issue on big data networking-challenges and applications

Zhixiang Hou [a,*] and Chong Shen [b]

[a] *Changsha University of Science and Technology, Changsha 410076, China*
[b] *College of Information Science and Technology, Hainan University, Haikou 570228, China*

Big data refers to the data set that cannot be captured, managed and processed by conventional software tools within a certain period of time. It is a massive, high growth rate and diversified information asset that requires new processing patterns to have stronger decision-making power, insight and process optimization capabilities.

Big data is more than a matter of size, and it is an emerging paradigm of data of very large size (volume) and fast in/out (velocity), from various sources (variety), and of high value for knowledge extraction and decision making. Technological advances in data gathering have led to a rapid proliferation of big data in diverse areas such as remote sensing, medicine, the Internet, and social sectors. Such data brings opportunities and challenges to scientists and engineers. In order to make use of this massive amount of data, the new data management and computational approaches are needed to permit scientists and engineers to analyze the data in (nearly) real time. In addition, big dataset cannot be stored in one location, and massively parallel processing databases and scalable storage systems are being designed to store the large datasets. What is more, big data generates an industry of supporting architectures. Many cloud computing platforms and frameworks are developed to handle the big data operations, such as MapReduce. To deal with different properties of big data, different algorithms also need to be developed. Overall, big data is an opportunity to find insights in new and emerging types of data and content. The objective of this special issue is to invite authors to submit original manuscripts that demonstrate and explore current advances in various aspects of big data networking. We solicit novel papers including but not limited to:

– Big data theory, applications, algorithm and challenges
– Recent development in big data and MapReduce
– Big data mining, analytics, and visualization
– Large data stream processing, infrastructure and incremental datasets on cloud
– Data management, distributed data systems, NoSQL data stores and DB scalability
– Security, privacy, trust, fault tolerance, reliability and risk in big data
– Big data placement, scheduling, and optimization
– Mobile systems and development for handheld devices such as mobile phones.

This special issue is dedicated to exploring recent advances in big data networking challenges and applications. Papers were selected using the following criteria:

– Quality and originality in theory and methodology
– Relevance to big data for Big Data Networking-Challenges and Applications
– Application-oriented papers exhibiting originality with reasonable theoretical content
– The novelty of theory and application focusing on Big Data Networking-Challenges
– The papers that were eventually selected represent and characterize the outstanding and active studies in the big data networking and illustrate clearly the advantages of developing big data networking.

In the first paper [2], Xiao-ling Jiang and Hui Zhang provide an intelligent mining acceleration algorithm based on particle swarm optimization. Particle swarm

---

*Corresponding author. E-mail: zhixiangh@foxmail.com.

optimization (PSO) algorithm is used to select the optimal feature subset, and average classification error is used as fitness function to complete intelligent mining of communication data. The experimental results show that the algorithm has lower memory consumption and faster mining speed.

In the second paper [3], in order to overcome the inaccuracy recognition results of the current network redundancy data, Dr. Ying Li provides a new method based on attributes and relations for evaluating the redundancy of massive heterogeneous data in the Internet of Things. The experimental results show that the proposed method has high accuracy and strong anomaly recognition performance, and the integrated database of the Internet of Things is security.

In the third paper [4], Ruijing Liu and Xiaoting Luo provides a new Bayesian classification based heterogeneous network anomaly big data intelligent detection. The method obtains the data frequency information and uses the Bayesian network classification method to realize the intelligent detection of abnormal big data in heterogeneous networks. The experimental results show that compared with the traditional method, the proposed method greatly improves the detection accuracy, convergence and anti-interference.

In the fourth paper [7], in order to overcome the low efficiency of traditional data mining algorithms without considering weighted association rules, Dr Dong Xu proposes a distributed data flow mining algorithm based on matrix weighted association rules. The experimental results show that the proposed algorithm has high efficiency and the highest accuracy of 98.9%.

In the fifth paper [5], Hui Sun provided an interest reading recommendation method based on big data technology for intelligent library. The experimental results show that compared with the traditional recommendation method, the proposed method can achieve high-precision reading recommendation with low response delay.

In the sixth paper [6], Dr. Hao Wang studied the integration of library E-Book borrowing history data based on big data technology. Based on big data technology, this method builds an integrated platform for library e-book borrowing. The data integration layer solves the local distortion and realizes the integrated processing of borrowed historical data. Experimental results show that the proposed method has the advantages of short time delay and low integration error rate.

In the seventh paper [1], in order to overcome the problem of poor recovery and stability of traditional big data fault-tolerant technology, a fault-tolerant tech-nology of big data cluster in distributed flow processing system is proposed. By using spark application framework, a fault-tolerant model of big data cluster is built to realize the fault-tolerant of big data cluster in distributed flow processing system. The experimental results show that the recovery rate of big data cluster fault-tolerant technology proposed is higher.

We very much hope that readers of this special issue will find the ideas presented in it of interest to them. We also hope that the material will be sufficiently stimulating to encourage many readers to explore the big data networking-challenges and applications further. This special issue is directed toward practicing engineers, researchers and industry operations managers, who wish to enhance or deepen their knowledge on the big data networking-challenges and applications. We do hope that these carefully selected papers will inspire and stimulate the industrial effort to use big data networking.

Acknowledgements

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

References

[1] Z.C. Jia, Fault-tolerant technology for big data cluster in distributed flow processing system, *Web Intelligence* **18**(3) (2020).

[2] X.L. Jiang and H. Zhang, Accelerated algorithm for intelligent mining of communication data in cloud computing environment, *Web Intelligence* **18**(3) (2020).

[3] Y. Li, Redundancy evaluation method of massive heterogeneous data in Internet of things based on attributes and relations, *Web Intelligence* **18**(3) (2020).

[4] R.J. Liu and X.T. Luo, Intelligent detection method for abnormal big data in heterogeneous networks based on Bayesian classification, *Web Intelligence* **18**(3) (2020).

[5] H. Sun, Research on interest reading recommendation method of intelligent library based on big data technology, *Web Intelligence* **18**(3) (2020).

[6] H. Wang, Research on the integration of library e-book borrowing history data based on big data technology, *Web Intelligence* **18**(3) (2020).

[7] D. Xu, Research on distributed data stream mining algorithms based on matrix weighted association rules, *Web Intelligence* **18**(3) (2020).