

---

## Editorial

---

### Ching-Hsien Hsu

Department of Computer Science and Information Engineering,  
Chung Hua University,  
707, Sec. 2, WuFu Rd.,  
Hsinchu, 30012, Taiwan  
and  
School of Computer and Communication Engineering,  
Tianjin University of Technology,  
391 Binshui Xidao,  
Xiqing District Tianjin 300384, China  
E-mail: chh@chu.edu.tw

**Biographical notes:** Ching-Hsien (Robert) Hsu is a Full Professor in the Department of Computer Science and Information Engineering at Chung Hua University, Taiwan; and Distinguished Chair Professor in School of Computer and Communication Engineering at Tianjin University of Technology, China. His research includes high performance computing, cloud computing and big data analytics, parallel and distributed systems, ubiquitous/pervasive computing and intelligence. He has published around 200 papers in refereed journals, conference proceedings and book chapters in these areas. He was awarded six times on Distinguished Award for Excellence in Research and Annual Outstanding Research Award through 2005 to 2012 from Chung Hua University. He has been serving as an executive committee and standing Director of Taiwan Association of Cloud Computing (TACC, 2008–2012), and executive committee of the IEEE Technical Committee of Scalable Computing (2008–2012). He is an elected member of the Phi Tau Phi Scholastic Honor Society and an IEEE senior member.

---

Welcome to the *International Journal of Big Data Intelligence (IJBDI)*. On behalf of the editorial board of *IJBDI*, I would like to thank the readers and the big data community world wide for their interest in *IJBDI*.

Big data is a rapidly expanding research area spanning the fields of computer science and information management, and has become a ubiquitous term in understanding and solving complex problems in different disciplinary fields such as engineering, applied mathematics, medicine, computational biology, healthcare, social networks, finance, business, government, education, transportation, telecommunications, etc. Today's IT technologies offer great opportunities to make the era of big data a reality. To gain value and insight from big data analytics, organisations need the ability not just to process the vast quantities of data being generated, but also to blend the right datasets together to give context and meaning. Big data intelligence goes beyond business intelligence; it gives significant potential for increasing productivity, competition and innovation. This revelatory exploration of the hottest trend in technology and the dramatic impact it will have on the economy, science, and society at large.

*IJBDI* aims to host and foster scientific dialogue in an international level by publishing and promoting excellent ideas, latest innovations, significant research and high-quality, peer-reviewed papers in the field of big data.

The inaugural issue of *IJBDI* consists of nine papers. Articles in this issue of *IJBDI* are selected after several rounds of meticulous review process. The papers accepted

for the first issue of *IJBDI*, cover hot topics in big data ranging from theoretical research to practical and applied development on variety of big data applications. I hope you will enjoy reading them and get inspired to innovate further.

The first paper entitled 'Big Data (lost) in the cloud' by Beniamino Di Martino, Rocco Aversa, Giuseppina Cretella, Antonio Esposito and Joanna Kołodziej done an excellent survey on the most recent developments in cloud computing support for solving Big Data problems. This study presents a comprehensive critical analysis of the existing solutions and shows further possible directions of the research in this domain including new generation multi-datacenter cloud architectures for storage and management of the huge Big Data streams.

The second paper entitled 'Designing and implementing a cloud-hosted SaaS for data movement and sharing with SlapOS' by Walid Saad, Heithem Abbes, Mohamed Jemni and Christophe Cérin proposes a cloud-hosted data grid using the SlapOS cloud. Through a software as a service (SaaS) solution, users can request and install automatically any data movement and sharing tools without any intervention of a system administrator. The proposed solution has been validated on Grid'5000 and is now running in production into the SlapOS cloud at Paris 13 University.

The third paper entitled 'Multi-source streaming-based data accesses for MapReduce systems' by Jiadong Wu and Bo Hong presented a novel multi-source streaming enabled Hadoop system for improving performance on remote data

access. With the help of such adaptive multi-source streaming capability, the Hadoop system is able to read remote data with higher aggregated throughput, and becomes less dependent on static topology information to explore locality. The proposed system was verified on multiple heterogeneous clusters, and the result shows significant improvements over native Hadoop system in the overall throughput and also the robustness to imbalanced system congestion.

The fourth paper entitled ‘A new approach for accurate distributed cluster analysis for Big Data: competitive K-means’ by Rui Máximo Esteves, Thomas Hacker and Chunming Rong presented an efficient and scalable clustering method, competitive K-means, for large dataset analytics. Comparing with several well-known methods, such as K-means, K-means++ and streaming K-means, the proposed CK-means can largely improve the accuracy of cluster analysis; decrease variance and the running time; and benefits the use of Hadoop and MapReduce.

The fifth paper entitled ‘Peculiarities of numerical algorithms parallel implementation for exa-flops multicomputers’ by Victor E. Malyshkin investigates methodologies for parallel implementation of large-scale numerical models and simulations on exa-flops multicomputer. A fragmented programming method for optimising dynamic load balancing of data-intensive computation was also suggested. This work presents indepth theoretical analysis. The suggested method is very useful for big data modelling and simulation.

The sixth paper entitled ‘Towards quality-of-service driven consistency for Big Data management’ by Álvaro García-Recuero, Sérgio Esteves and Luís Veiga proposed a data store architecture that serves custom levels of consistency to applications, providing consistency guarantees driven by data semantics. This work presents an optimisation for non-relational data store HBase that uses quality-of-data (QoD) to envision a very well defined goal: a flexible and adaptive consistency model to provide applications with the required information at the right time. This study brings insights about how data stores can handle consistency without requiring intrusion to the data schema and avoiding extra middleware overhead.

The seventh paper entitled ‘D-CEP4CMA: a dynamic architecture for cloud performance monitoring and analysis via complex event processing’ by Afef Mdhaffar, Riadh Ben Halima, Mohamed Jmaiel and Bernd Freisleben presents a dynamic monitoring and analysis architecture for cloud computing environments. Based on the methodology of complex event processing (CEP), the proposed system architecture was designed to dynamically switch between different centralised and distributed CEP architectures

depending on the load/memory of the CEP machine and network traffic conditions. The experiments demonstrated the merits of the D-CEP4CMA and very good results, which lead interest to further research.

The eighth paper entitled ‘An extended analytical study of arabic sentiments’ by Nawaf A. Abdulla, Mahmoud Al-Ayyoub and Mohammed Naji Al-Kabi performed a comprehensive study on interdisciplinary sentiment analysis (SA) for automatically identifying the polarity of comments/reviews in social network. The authors collected and annotated a large dataset of Arabic comments from Yahoo!-Maktoob, which contains diverse domains (arts, politics, science and technology, and social) for empirical experimentations.

The ninth paper entitled ‘Health Big Data analytics: current perspectives, challenges and potential solutions’ by Mu-Hsing Kuo, Tony Sahama, Andre W. Kushniruk, Elizabeth M. Borycki and Daniel K. Grunwell discusses the characteristics of health Big Data as well as the challenges and solutions for health Big Data analytics (BDA). The authors designed and evaluated a pipelined framework for use as a guideline/reference in health BDA. Findings and results of this study may trigger further related research and technology improvements in application of BDA and could serve as a landmark source to advancing Big Data analytics in healthcare.

*IJBDI* owes much to many people who contributed their time and ideas to define the scope and vision of this publication. Thanks are due first to the strong support offered by all members of the international and distinguished Editorial Board with extensive academic qualifications in getting this journal up and running. I am very thankful to all the reviewers who supported the journal with their valuable comments and suggestions. I would like to thank all researchers in the field of big data who accepted our invitation to submit their scholarly work for the inaugural issue of *IJBDI*. I express my sincere thanks to the Inderscience Publishers who provided me the opportunity to edit an international journal with such a reputed publisher. It is my privilege and honor to serve the world wide big data community.

Big data is reshaping the future and bringing huge possibilities for tomorrow. Without doubt, big data intelligence will transform how we think, work and live. I request your continued support for establishing *IJBDI* as the premier forum for big data research and innovation. High quality submissions and special issue proposals devoted to emerging topics in big data intelligence are definitely very welcome. Your feedback and comments on further improving *IJBDI* are highly appreciated.