

Special issue on big data computing, analytics and applications

Chenren Xu¹ · Zhu Han² · Yanyong Zhang³ · Lan Zhang⁴

© Springer-Verlag London 2016

With the rapid growth of data in the past few years, information technologies have made significant impacts on global environments on both positive and negative aspects. Generation of huge amounts of data, called big data, across different sectors such as banking, health care, retail and education, among others, is creating the need for an efficient tool to manage this data. This will foster many interesting problems in the intersection of computing, analytics and applications. Conventional database management tools do not have the capability to manage surging volumes of unstructured data. For example, more than 80% of data is unstructured in the form of videos, tweets, GPS coordinates and emails, which means that decisions need to be made at high velocity.

Data is expected to exponentially grow through data collected via pervasive sensors and/or the Internet. It also leads to new emerging challenges that have the potential to create more accurate solutions for science and technologies. We believe that we need a new way for personal and ubiquitous computing where big data is immensely involved, especially for the data trace collected from ambient sensors, wearable, social media and so on. This special issue is devoted to addressing new challenges in big data computing, analytics and applications in personal and ubiquitous Computing. We intend to bring together

researchers, developers and practitioners interested in this research area from academia, industry and service providers, to share ideas, experiences and practical implementations. Both theoreticians as well as practitioners, including system builders and individuals applying big data in personal and ubiquitous computing domains, are encouraged to submit their papers to this special issue. To that end, papers are solicited from all BigCom related areas emphasizing theoretical foundations, novel modeling, algorithmic methodologies and data-driven applications.

This issue is in collaboration with The 2nd International Conference on Big Data Computing and Communications (BIGCOM) held in Shenyang, Liaoning, China (<http://conf.neu.edu.cn/bigcom2016/>), and addresses on hot topics relevant to big data computing, analytics and applications. Following a review process, we accepted 10 papers from BIGCOM'16 for this theme issue. Each of the papers was peer-reviewed by at least two experts in the field. In the following, we provide a brief introduction to each paper.

The paper “Redundancy Reduction for Indoor Device-Free Localization” by Jinjun Liu, Ning An, Md. Tanbir Hassan, Min Peng, Zheng Cui and Shenghui Zhao presents a method of redundancy reduction. With the goal of using fewer nodes while maintaining high localization accuracy, the proposed method controls the two-level redundancy efficiently and reduces the number of nodes greatly, thereby reducing the needed amount of localization data and the hardware costs. The results show that the proposed method can be used as an efficient tool for indoor device-free localization.

The paper “CondioSense: High-quality Context-aware Service for Audio Sensing System via Active Sonar” by Fan Li, Huijie Chen, Xiaoyu Song, Qian Zhang, Youqi Li and Yu Wang presents CondioSense, a Context-aware service for audio sensing system, which identifies the

✉ Chenren Xu
chenren@pku.edu.cn

¹ Peking University, Beijing, China

² University of Houston, Houston, TX, USA

³ Rutgers University, New Brunswick, NJ, USA

⁴ The University of Science and Technology of China, Hefei, China

current phone context (i.e., pocket, bag, car, indoor and outdoor) and detects the microphone sensing states. Audio sensing has been applied in various mobile applications for sensing personal and environmental information to improve user's life quality. However, the quality of audio sensing is distorted seriously, while the sensing service is working in incorrect context or the ability of the acoustic sensing is limited. The main idea behind context detection is to extract multi-path features from actively generated acoustic signal to identify various contexts since the space size and material among various contexts is different. The sound of physical vibration is explored on microphone sensing state detection, by leveraging that the frequency response of recorded vibration sound changes when the signal propagation in the air is blocked with the microphone covered. CondioSense is prototyped on the smartphones and offers the possibility to recognize various phone contexts with an accuracy exceeding 92% and the accuracy of microphone sensing states detection exceeding 90%.

The paper "Participant Selection for Data Collection through Device-to-Device Communications in Mobile Sensing" by Yu Wang, Hanshang Li and Ting Li proposes to leverage device-to-device communications for data collection in mobile sensing. Relay node selection for such data collection is formulated as an optimization problem, and greedy-based algorithms are proposed to solve the problem. Experiments over a real-life mobile trace confirm the effectiveness of the proposed algorithms.

The paper "Customized Privacy Preserving for Inherent-data and Latent-data" by Zaobo He, Zhipeng Cai, Yunchuan Sun, Yingshu Li and Xiuzhen Cheng generalizes two types of malicious behaviors of adversaries by defining inherent-data privacy and latent-data privacy and develops an optimized data sanitization strategy that can optimize the trade-off between data utility and customized two types of privacy. A collective data sanitization method is proposed to combat against powerful adversaries, which takes advantages of various data manipulating methods to guarantee sanitized data do not sacrifice much data utility. Experiment results toward real dataset show that the developed data sanitization strategy does not much reduce the benefit brought by user data, while the sensitive information can still be protected.

The paper "Big Data Challenges in Ocean Observation: A Survey" by Yingjian Liu, Meng Qiu, Chao Liu and Zhongwen Guo presents an overview of new challenges and key issues in marine big data. Some representative ocean observing platforms and projects for marine data acquisition are briefly introduced. The paper also reviews some popular big data techniques which may applicable for solving marine big data challenges in the phases of data storage, data computing and analysis. Some applications in

physics, chemistry, geology and biology illustrate the significant uses of marine big data.

The paper "The Soil Moisture Sensor Based on Soil Dielectric Property" by Zhan Huan, Hui Wang, Chen Li and Caiyan Wan presents a soil moisture sensor based on soil dielectric property, which covers circuit theory, hardware simulation, resonance-frequency selection and temperature compensation. The experimental results show that the soil moisture sensor works well and suits for building large-scale agriculture wireless sensor networks.

The paper "Counter-Strike: Accurate and Robust Identification of Low-level Radiation Sources with Crowdsensing Networks" by Chaocan Xiang, Panlong Yang and Shucheng Xiao proposes an accurate and robust identification method called Counter-Strike, which can accurately identify radiation sources based on inaccurate crowdsensing measurements. This method uses truthful probability of sources for robust identification, addressing the problem of existing methods that their identification accuracy tightly depends on identification threshold. Moreover, this paper proposes an iterative truthful-source identification algorithm, which gradually improves the identification accuracy by alternately iterating between sensor efficiency estimation and truthful probability estimation, successfully solving the problem of unknown sensor efficiency. The extensive simulations and theoretical analysis prove the high accuracy and robustness of the proposed method.

The paper "Semantic Trajectories Based Social Relationships Discovery Using WiFi Monitors" by Fengzi Wang, Xinning Zhu and Jiansong Miao presents a semantic trajectory-based way to uncovering social relationships through observing WiFi data collected by WiFi monitors. An semantic trajectory similarity estimation is proposed to measure the similarity among mobile users. Then, the trajectory similarity measurement is used to exploit underlying social networks exist in the university as well as infer specific type of social relationships between a pair of mobile users by further studying their matching trajectory points.

The paper "SentiStory: Multi-grained Sentiment Analysis and Event Summarization with Crowdsourced Social Media Data" by Yi Ouyang et al. presents a multi-grained sentiment analysis and event summarization system called SentiStory. It characterizes events from two perspectives: coarse-grained and fine-grained sentiment analysis. In coarse-grained analysis, SentiStory discovers microblogs which are important in sentiment; while in fine-grained analysis, it detects significant change moments of sentiments in the event and identifies which microblog causes the changes. Experiments based on the real event datasets collected from Sina Weibo demonstrate the effectiveness of our approach.

The paper “Evaluation of Missing Value Imputation Methods for Wireless Soil Datasets” by Jia Shao, Wei Meng and Guodong Sun investigates eight typical methods to infill the missing values of the time-series soil dataset collected by wireless sensing systems. With a real-world soil dataset and practical value-missing settings, extensive evaluations of these eight widely used methods are conducted, and the performances are analyzed and compared comprehensively. The dominant factors affecting the infilling performance and relatively potential policies

employed in these methods are pointed out. Furthermore, a few promising heuristics for designing more effective approaches of infilling continuously missing values in wireless soil dataset are discussed.

Hereby, we would like to take this opportunity to thank all authors for their contributions. We thank all reviewers for their time, effort and expertise for reviewing the papers. Finally, we would also like to express our honor to serve as the guest editors of this issue.