Special Issue on Wireless Big Data

Yang Yang, Jie Li, Cheng-Xiang Wang, Olav Tirkkonen, and Ming-Tuo Zhou

Big data, which has been following the exponential growth rates in different commercial areas, has profoundly changed the way we live. It has received considerable attention in both academic and industrial communities, in contexts such as mobile communications, distributed computing, e-health, intelligent transportation systems, wireless sensor networks, etc. In the meantime, the Internet of Things (IoT) scenarios considered in the Fifth Generation (5G) wireless communication systems are expected to create many novel applications and services with various requirements [1]. These new directions bring a dramatic increase and change in the amount and types of wireless data, thus driving wireless communications into a new era. Therefore, an in-depth analysis and understanding of wireless big data can greatly facilitate better system design and performance optimization, which will certainly benefit equipment vendors, network operators and service providers.

Specifically, wireless big data enables researchers and engineers to i) create sophisticated classifications, ii) characterize various wireless technologies, iii) tailor different services precisely, and iv) predict both technical trends and customer behaviors. It is essential not only to effectively manage and exploit wireless big data collected from expensive experiments and measurements, but most importantly, to develop efficient practical methods to extract useful information and discover new knowledge from massive wireless big data, thus realizing the full potential of future wireless communications. We are very grateful that many researchers from both industry and academia have contributed their latest research results on wireless big data and the corresponding data-driven wireless technologies. After a rigorous review process, seven papers have been selected out of 30 submissions for this special issue. We are now happy to present these seven papers focusing on different application scenarios, system requirements and technical challenges of wireless big data.

The first paper "A Big Data Enabled Channel Model for 5G Wireless Communication Systems" proposes a big data and machine learning enabled wireless channel model framework. It is based on artificial neural networks, including feedforward neural network and radial basis function neural network. The input parameters are transmitter and receiver coordinates, signal propagation distances, and carrier frequency; while the output parameters are channel statistical properties such as received powers, root mean square of delay spreads and angle spreads. Simulation results indicate that machine learning algorithms provide powerful analytical tools for future measurement-based wireless channel modeling.

The second paper "Clustering Analysis in the Wireless Propagation Channel with a Variational Gaussian Mixture Model" also considers wireless channel models. It introduces the Gaussian mixture model (GMM) to implement channel multipath clustering. Specifically, the expectation-maximization algorithm is utilized to search for posterior estimates of the GMM parameters. Then, the variational Bayesian algorithm is adopted to optimize the GMM parameters and to determine the optimal number of Gaussian distributions without cross-validation. By defining the compact index, its close relationships with the GMM clustering mechanism and multipath propagation characteristics are revealed. Experimental results demonstrate the effectiveness of the proposed method.

Besides channel models, wireless content delivery is addressed. The third paper "MWBS: An Efficient Many-to-Many Wireless Big Data Delivery Scheme" proposes an efficient many-to-many wireless big data delivery scheme (MWBS) to provide group-based data dissemination and retrieval with name-integrated forwarding. Specifically, by adopting designated forwarding and cacheable nodes (DFCNs) as the roots, bi-directional trees are securely constructed for every group through the procedures of group initiation, joining, leaving, publication, and multi-level inter-zone routing. By choosing the optimized number of DFCNs and their preferred locations, simulation results show that MWBS can effectively reduce the overall overheads in control packet and state storage.

When wireless big data is shared, the fourth paper "Incremental Deep Computation Model for Wireless Big Data Feature Learning" proposes an incremental deep computation model for learning wireless data features in dynamic IoT

- *Y. Yang is with the Shanghai Institute of Fog Computing Technology (SHIFT), ShanghaiTech University, 393 Middle Huaxia Road, Pudong, Shanghai 201210, China, and also with the Research Center for Network Communication, Peng Cheng Laboratory, 2 Xingke First Street, Nanshan, Shenzhen 518000, China. E-mail: yangyang@shanghaitech.edu.cn.*
- *J. Li is with the Department of Computer Science and Engineering, School of Electronic Information and Electrical Engineering, Shanghai Jiaotong University, Minhang, Shanghai 200240, China. E-mail: lijiecs@sjtu.edu.cn.*
- *C.-X. Wang is with the National Mobile Communications Research Laboratory, School of Information Science and Engineering, Southeast University, Nanjing 210096, Jiangsu, China and also with the Purple Mountain Laboratories, Nanjing 211111, Jiangsu, China. E-mail: chxwang@seu.edu.cn.*
- *O. Tirkkonen is with the Department of Communications and Networking, Aalto University, Maarintie 8P.O. Box 1550002150 Espoo, Finland. E-mail: olav.tirkkonen@aalto.fi.*
- *M.-T. Zhou is with the Shanghai Institute of Microsystem and Information Technology, Chinese Academy of Sciences, 865 Changning Road, Shanghai 200050, China. E-mail: mingtuo.zhou@mail.sim.ac.cn.*

Digital Object Identifier no. 10.1109/TBDATA.2020.2966806

scenarios. Two incremental tensor auto-encoders are developed, together with a parameter-based and a structure-based incremental learning algorithm. When new wireless data is available, the former only updates network parameters, while the latter simultaneously adjusts the structure and updates the parameters. Experimental results show that the proposed model can modify the network in an incremental manner by efficiently learning from new-arrival data while preserving the prior knowledge from the existing data.

As to new applications, the fifth paper "Sleepy: Wireless Channel Data Driven Sleep Monitoring via Commodity WiFi Devices" utilizes wireless channel data from commercial WiFi devices to develop a low-cost compatible sleep monitoring system, namely Sleepy, which can effectively derive the energy feature of wireless channel from the accumulated channel data over a long period. Experiment results in two real environments show that Sleepy can offer very high detection accuracy (DA) and very low false negative rate (FNR). It does not require any calibration or target-dependent training for achieving these good performance.

Sleepy drivers are quite dangerous on the road. The sixth paper "WiFind: Driver Fatigue Detection with Fine-Grained Wi-Fi Signal Features" utilizes commodity WiFi signals to develop a driver fatigue detection system, namely WiFind, which can recognize the body features of a driver in multiple modes and effectively detect the fatigue symptoms without relying on any visual image or video. Experimental results show that WiFind can achieve the recognition accuracy of 89.6 percent in a single driver scenario. It does not require wearing any special devices, so is simple and easy to deploy in real driving environments.

Last but not least, the seventh paper "Differential Privacy Preserving of Training Model in Wireless Big Data with Edge Computing" proposes a machine learning strategy with differential privacy to protect the privacy in training datasets when smart edges analyze wireless big data. By adding Laplace mechanisms, designing Output Perturbation algorithm and Objective Perturbation algorithm, and developing differential privacy preserving methods, effective privacy protection can be guaranteed in training datasets and in the correlated datasets. Experimental results with four popular datasets show that the proposed methods can protect the privacy of training datasets and guarantee the accuracy on benchmark datasets.

We believe the readers of this special issue will find these papers quite informative and inspiring. Finally, we would like to take this opportunity to thank all the authors for submitting their latest work to this special issue, and thank the Editor-in-Chief and all the anonymous reviewers for their valuable comments and suggestions during the review process. Also, we are very grateful to the TBD team for their fantastic support and help, which made this special issue successful.

References

[1] Y. Yang, "Multi-tier computing networks for intelligent IoT," *Nature Electron.*, vol. 2, pp. 4–5, Jan. 2019.

Yang Yang (Fellow, IEEE) is currently a professor with ShanghaiTech University, China. He is the executive dean of the School of Creativity and Art, and the director of the Shanghai Institute of Fog Computing Technology (SHIFT). He is also with the Research Center for Network Communication, Peng Cheng Laboratory, Shenzhen, China. His current research interests include fog computing networks, service-oriented collaborative intelligence, wireless sensor networks, IoT applications, and advanced testbeds and experiments. He has published more than 200 papers in these areas.

Jie Li (Senior Member, IEEE) is currently a professor with the Department of Computer Science and Engineering, Shanghai Jiaotong University, China. His current research interests include big data, AI, blockchain, edge computing, OS, modeling and performance evaluation of information systems. He is the founding chair of the IEEE Communications Society Technical Committee on Big Data and the co-chair of the IEEE Big Data Community. He is a member of the Engineering Academy of Japan and a senior member of ACM.

Cheng-Xiang Wang (Fellow, IEEE) is currently a professor with Southeast University and Purple Mountain Laboratories, Nanjing, China. His current research interests include wireless channel measurements/modeling and B5G wireless communication networks. He has co-authored three books, one book chapter, and more than 370 papers. He is a fellow of the IET, an IEEE Communications Society distinguished lecturer, in 2019 and 2020, and a highly-cited researcher recognized by Clarivate Analytics, in 2017-2019.

Olav Tirkkonen (Member, IEEE) is currently an associate professor in communication theory with the Department of Communications and Networking, Aalto University, Finland. His current research interests include coding theory, multi-antenna techniques, as well as cognitive management of 5G and beyond cellular networks. He has published more than 200 papers, and co-authored the book "Multi-Antenna Transceiver Techniques for 3G and Beyond".

Ming-Tuo Zhou (Senior Member, IEEE) is currently a professor with the Shanghai Institute of Microsystem and Information Technology, Chinese Academy of Sciences, China. His current research interests include fog/edge computing, wireless sensor networks, industrial Internet of Things, cyber-physical systems, 5G, advanced channel modeling, and big data/machine learning applications in wireless communications. He has co-authored more than 80 papers, and co-edited two technical books.