
Editorial

Weisong Shi*

Department of Computer Science,
Wayne State University,
5143 Cass Ave,
Detroit, MI 48202, USA
Email: weisong@wayne.edu
*Corresponding author

Xueyan Tang

School of Computer Engineering,
Nanyang Technological University,
50 Nanyang Avenue, Singapore
Email: asxytang@ntu.edu.sg

Biographical notes: Weisong Shi is an Associate Professor of Computer Science at Wayne State University. He received his BS degree from Xidian University in 1995, and PhD degree from the Chinese Academy of Sciences in 2000, both in Computer Engineering. His current research focuses on mobile computing, distributed systems and high performance computing.

Xueyan Tang is currently an Associate Professor in the School of Computer Engineering at Nanyang Technological University, Singapore. He received his BEng degree in Computer Science and Engineering from Shanghai Jiao Tong University, Shanghai, China in 1998, and his PhD degree in Computer Science from the Hong Kong University of Science and Technology in 2003.

As new fabrication and integration technologies continue to reduce the cost and size of wireless sensors, the observation and control of our physical world will expand dramatically with the temporally and spatially dense monitoring afforded by wireless sensor network technology (Estrin et al., 1999; Estrin et al., 2002). Several applications such as ZebraNet (Sadler et al., 2004), Counter-sniper system (Simon et al., 2004), environment sampling (Batalin et al., 2004), target tracking (Shrivastava et al., 2006) and structure monitoring (Xu et al., 2004) have been launched, showing the promising future of wide range of applications of networked sensor systems. Their success is nonetheless determined by whether the sensor networks can provide a high-quality stream of data over a long period. The inherent feature of unattended and untethered deployment of networked sensors in a potentially malicious environment, however, imposes challenges to the underlying systems. These challenges are further complicated by the fact that sensor systems are usually seriously energy constrained. Most previous efforts have focused on devising techniques to save the sensor node energy and thus to extend the lifetime of the whole sensor network. However, with more deployments of real sensor systems, in which the main function is to collect interesting data and to share with peers, data quality is becoming a more important issue in the design of sensor systems (Sha and Shi, 2008). Consistency, accuracy, reliability and survivability concerns have to be addressed in sensor data collection, storage and processing.

This special issue collects five papers that represent recent progress in data quality management in wireless sensor networks, covering a broad spectrum of interesting results, including data collection, outlier data detection, fault tolerance and enabling technologies.

Many sensor applications often require collecting raw sensed values from many sensor nodes to one centralised server. Sensor data collection typically comes with various quality requirements, e.g. the level of precision requested for temperature values, the time constraints for getting the data, or the percentage of data that is needed. Han, Hakkarinen, Boonma and Suzuki in their paper 'Quality-aware sensor data collection' propose a quality-aware sensing framework that characterises the quality needs of sensor applications and classifies different sensor data collection problems. The authors list several interesting research directions and provide a holistic approach to sensor data collection with diverse quality requirements.

Outlier detection plays an important role in wireless sensor networks applications in order to ensure high data quality and reliable detection of interesting and critical events. Unlike traditional outlier detection techniques that are performed offline, constrained resources available in wireless sensor networks necessitate identifying outliers in an online manner. This means that outliers in distributed streaming data should be detected in (near) real time with a high accuracy while keeping the resource consumption of the wireless sensor network to a

minimum. In the paper ‘Ensuring high sensor data quality through use of online outlier detection techniques’, Zhang, Meratnia and Havinga propose outlier detection techniques based on one-class quarter-sphere support vector machine to meet the constraints and requirements of wireless sensor networks.

Understanding anomaly of sensing data is the key for wireless sensor networks. Bigrigg, Matthews and Garrett in their paper ‘Fault perturbations in building sensor network data streams’ investigate the causes of temperature sensor data perturbations (two temperature data readings taken within seconds of each other that differ by several degrees) in the context of a building monitoring application, by leveraging Critter data sensors, which are an analog computer-attached sensor device developed at Carnegie Mellon University. They argue that the technologies used in off-the-shelf integrated sensor devices hide the details of raw readings, limiting the effects of data outliers. The novelty of Critter is to provide raw instantaneous readings including outlier data that may be considered anomalies or perturbations. They found that temperature sensor data perturbations are actually the effects of user activity within buildings. By capturing the raw data without automatic processing, they are able to show a correlation between time of the work day, and the frequency of data perturbation.

In the paper ‘A systematic probabilistic approach to energy-efficient and robust data collections in wireless sensor networks’, Zhao and Liang propose a systematic approach, based on probabilistic graphical model, to infer missing observations in wireless sensor networks for sustaining environmental monitoring. By leveraging a pairwise Markov Random Field to model the spatial correlations in a sensor network, they are able to address two critical challenges: (a) energy-efficient data gathering through planned communication disruptions resulting from energy-saving sleep cycles, and (b) sensor-node failure tolerance in harsh environments. They demonstrate the effectiveness of their approach using real-world sensed soil moisture data and vegetation data on 32×32 grids. Empirical results show that their approach can achieve high rates of accurate estimation for unobserved nodes with minimal training data in the model construction, even when the unobserved nodes consist of more than 50% of the total sensor nodes.

Using multiple paths from source to destination is a popular technique to achieve the tradeoff between reliability and efficiency in volatile wireless networks. In the paper ‘Efficient multipath in wireless networks using network coding over braided meshes’, Toledo and Wang argue that several

challenges, such as high maintenance overhead and low energy efficiency, make efficient multipath transmission difficult to implement. They instead propose the use of network coding over a multipath braided mesh topology that exploits both the low cost mesh-topology construction, such as those obtained by diffusion algorithms, and the capacity achieving capability of linear network coding. Simulation results show that their approach is able to achieve the best energy efficiency of existing methods and it easily adapts to the changing conditions of the network and can be used to adjust reliability on demand.

In summary, we are pleased to present this selection of five articles in this special issue on data quality management in WSNs. We trust that you will enjoy reading this special issue.

References

- Batalin, M.A., Rahimi, M., Yu, Y., Liu, D., Kansal, A., Sukhatme, G.S., Kaiser, W.J., Hansen, M., Pottie, G.J., Srivastava, M. and Estrin, D. (2004) ‘Call and responses: experiments in sampling the environment’, *Proceedings of the ACM SenSys '04*, 3–5 November, Baltimore, USA.
- Estrin, D., Culler, D., Pister, K. and Sukhatme, G. (2002) ‘Connecting the physical world with pervasive networks’, *IEEE Pervasive Computing*, Vol. 1, No. 1, pp.59–69.
- Estrin, D., Govindan, R., Heidemann, J. and Kumar, S. (1999) ‘Next century challenges: scalable coordination in sensor networks’, *Proceedings of the 5th Annual ACM/IEEE International Conference on Mobile Computing and Networking (MobiCom'99)*, August 1999.
- Sadler, C., Zhang, P., Martonosi, M. and Lyon, S. (2004) ‘Hardware design experiences in zebrant’, *Proceedings of the ACM SenSys '04*, 3–5 November, Baltimore, USA.
- Sha, K. and Shi, W. (2008) ‘Consistency-driven data quality management in wireless sensor networks’, *Journal of Parallel and Distributed Computing*, Vol. 68, No. 9, pp.1207–1221.
- Shrivastava, N., Mudumbai, R., Madhow, U. and Suri, S. (2006) ‘Target tracking with binary proximity sensors: fundamental limits, minimal descriptions, and algorithms’, *Proceedings of the ACM SenSys '06*, 1–3 November, Colorado, USA.
- Simon, G., Ledeczi, A. and Maroti, M. (2004) ‘Sensor network-based countersniper system’, *Proceedings of the ACM SenSys '04*, 3–5 November, Baltimore, USA.
- Xu, N., Rangwala, S., Chintalapudi, K., Ganesan, D., Broad, A., Govindan, R. and Estrin, D. (2004) ‘A wireless sensor network for structural monitoring’, *Proceedings of the ACM SenSys '04*, 3–5 November, Baltimore, USA.