Uncertainty in Wireless Sensor Networks

Wang-Chien Lee
Department of Computer Science and Engineering
Pennsylvania State University
University Park, PA 16802
wlee@cse.psu.edu

Advances in sensor technology, wireless networking and in mobile devices are revolutionizing the ways that information from the physical world is collected and used. Wireless sensor networks deploying these technologies have had significant impacts on applications in a wide variety of fields including the military, science, industry, commerce, transportation and healthcare. Factors such as constraints of the sensing devices, the varying and potentially hostile environments where sensors are deployed, changing network topologies, unreliable communications, irregular radio patterns, etc. point to *uncertainty* as an inherent operational aspect of sensor networks.

As examples, consider the following three aspects of uncertainty in sensor networks:

1. **Communication uncertainty**. It has been observed that wireless links in sensor networks do not follow an ideal spherical pattern, and thus routing protocols developed based on that assumption will not produce optimal performance. In addition, mobile sensor networks in particular can exhibit intermittent connection patterns (e.g., people or soldiers as sensors). Quantifying the communication uncertainty (e.g., the availability, quality and connection patterns of communication links) will support algorithms for better routing decisions.
2. **Sensing uncertainty**. Sensing coverage is an essential determinant of the sensing quality of a network, yet environmental interference, noise, sensor types and other factors can contribute to uncertainty in coverage. Recent studies have argued that probabilistic models are needed to capture sensor behavior. We argue that statistics and models developed from statistics will better capture that behavior. Quantifying sensing uncertainty (e.g., uncertainty in sensor range) will facilitate effective sensor (re-)deployment strategies for mobile as well as static sensor networks.
3. **Data uncertainty**. Sensor readings, and the derived data collected and reported by sensor nodes, inherently come with some degree of imprecision. Thus, we argue that sensor readings and processing done on sensor data should be augmented with indications of data uncertainty (e.g., in the form of *confidence*). For example, an enquiry for the sensor nodes with the highest temperature reading may return a set of nodes associated with probabilities rather than one single node. Foundational work addressing uncertainty in databases strongly suggests that quantifying data uncertainty will provide better quality results and decision-making in networked sensor systems.

While most existing techniques apply heuristics to cope with uncertainty, we argue that uncertainty should be properly quantified and treated as a first-class citizen in the design of the various layers of networked sensor systems. When the types of uncertainty in sensor networks are captured (empirically or statistically) and presented in quantified forms, more efficient and effective layers of network and data management functionality can be developed.

To address the above-mentioned three aspects of uncertainty inherently residing in the operations and applications of wireless sensor networks, many existing issues and solutions need to be re-visited and further exploited. In the following, we discuss a number of challenges:

- **Uncertainty modeling in link quality.** An on-line, self-adapted link quality estimation mechanism [1] within sensor nodes is essential for making routing decisions and improving network performance. To address this challenging task, machine learning techniques can be exploited to capture the temporal/spatial patterns exhibited in links as well as the correlations among nearby sensor nodes.
- **Uncertainty modeling in network connectivity.** Sensor nodes can be mobile and the operating environment of wireless sensor networks can be dynamic, thus resulting in intermittent connections. Many network communication protocols and applications are built upon global static knowledge of network connectivity derived from collected traces [2,3]. A challenge is how to capture the dynamically evolving relationship between nodes and the intermittent connectivity among sensor nodes.
- **Probabilistic routing.** With different model of uncertain link/connection model, routing and data dissemination algorithms need to be re-designed to deal with intermittent connections in wireless sensor networks. Most of existing works are based on the idea of carry-and-forward, resulting in contention of limited resource such as node storage and wireless bandwidth. To overcome these challenge, probabilistic routing algorithms based on uncertainty models of link quality and connectivity are desirable since it's difficult for conventional deterministic routing algorithms to achieve good performance. Additionally schemes for scheduling fairness and congestion control under intermittent connections need to be revisited.
- **Sensing Coverage.** A typical assumption in studies of sensing coverage problems, based on deterministic disk model, is that every sensor node may definitely detect an event happened within its sensing range. However, recent studies argue that probabilistic sensing models capture the sensing behavior more realistically than the deterministic disk model [4,5]. Under probabilistic sensing models, the sensing coverage problem needs to be revisited to better define the quality of probabilistic coverage and develop new sensor node deployment schemes.
- **Data Uncertainty.** Due to the dynamics of physical environments and possible hardware defeats, raw readings collected in sensor nodes are inherently inaccurate and imprecise. In other words, raw readings can only reflect approximate measurements of the monitored environments and thus are considered as uncertain [6,7]. Probability models can be employed to capture the data uncertainty. Various research effort in the database community has started to treat uncertain data as a first-class citizen in database in order to better capture the data uncertainty [8,9].
- **In-Network Probabilistic query processing.** A series of research work has been developed for probabilistic query processing. However, most of them assume centralized uncertain databases, i.e., all the uncertain data are maintained in a single host. This assumption is not practical for WSNs since the energy-scarce sensor nodes may need to convey their data frequently. Thus, a demand for in-network uncertain data management arises. The challenges for in-network processing of uncertain data in wireless sensor networks includes acquisition

of global knowledge about the distributed datasets and query processing algorithms which usually involves complex interplay among uncertain data attributes and probability [10].

Reference

[1] Y. Xu and W. Lee, “Exploring Spatial Correlation for Link Quality Estimation in Wireless Sensor Networks, IEEE International Conference on Pervasive Computing and Communications (PerCom 06), Pisa, Italy, pp. 200-211, March 2006.

[2] T. Karagiannis et al., “Power law and exponential decay of inter contact times between mobile devices”, Proceedings of the 13th annual ACM international conference on Mobile computing and networking, pp.183-194, 2007.

[3] S. Akoush and A. Sameh, “Mobile User Movement Prediction Using Bayesian Learning for Neural Networks”, Proceedings of the 2007 International Conference on Wireless Communications and Mobile Computing (ACM IWCMC), pp.191-196, 2007.

[4] M. Hefeeda and H. Ahmadi, “A Probabilistic Coverage Protocol for Wireless Sensor Networks”, in Proceedings of IEEE International Conference on Network Protocols (ICNP), pp. 41-50, Oct. 2007.

[5] Y. Zou and K. Chakrabarty, “Sensor deployment and target localization in distributed sensor networks”, ACM Transactions on Embedded Computing Systems, vol. 3, no. 1, pp. 61–91, Feb. 2004.

[6] E. Elnahrawy and B. Nath, “Poster abstract: online data cleaning in wireless sensor networks”, Proceedings of the 1st On Embedded Networked Sensor Systems (SenSys), 2003.

[7] A. Deshpande, C. Guestrin, S. R. Madden, J. M. Hellerstein, and W. Hong, “Model-driven Data Acquisition in Sensor Networks”, Proceedings of the 30th International Conference on Very Large Data Base (VLDB), 2004.

[8] R. Cheng, D. V. Kalashnikov, and S. Prabhakar, “Evaluating probabilistic queries over imprecise data”, Proceedings of the 2003 ACM SIGMOD International Conference on. Management of Data (SIGMOD'03), pages 551–562, New York, NY, USA, 2003.

[9] N. Dalvi and D. Suciu, “Efficient Query Evaluation on Probabilistic Databases”, Proceedings of the 30th International Conference on Very Large Data Base (VLDB), 2004.

[10] M. Ye, X. Liu, W. Lee, and D. Lee, “Probabilistic Top-k query processing in distributed sensor networks”, Proceedings of the 26th International Conference on Data Engineering (ICDE), 2010.