Data Management of Wireless Sensor Networks

Haiying Shen, Ting Li
Department of Computer Science and Computer Engineering
University of Arkansas, Fayetteville, AR 72701
Email: {hshen, txl005}@uark.edu
Telephone: (479) 575-4382 Fax: (479) 575-5339

***Abstract*—Wireless sensor networks can monitor physical and environmental conditions inexpensively, and the networks have been widely used in military and civilian applications. The sensor nodes in the networks not only have the ability to sense, but also can supply data storage, computation and communicate with each other. However, energy efficiency is the major concern in wireless sensor networks, which usually rely on battery for energy source. Centralized data storage and query processing schemes lead to high energy consumption. This paper proposes a distributed data storage scheme (DDS) that disseminates related event data in a neighborhood in wireless sensor networks. Besides offering data query according to event names, DDS also provides distributed querying. Experimental results show that the DDS scheme yields higher efficiency and effectiveness than the centralized scheme.**

***Index Terms*—wireless sensor network, data management.**

I. Introduction

A wireless sensor network (WSN) is a wireless network consisting of a large number of distributed low-power and inexpensive sensor devices. WSNs can be used for monitoring environment and observing phenomena. Particularly, WSNs have been used in many military and civilian application areas such as military target tracking, habitat monitoring, health monitoring, environmental contaminant detection and industrial process control [1]. The sensor nodes are usually equipped with memory, wireless communication device and microcontroller [1], thus they have the capabilities of computation, storage and communication [2]. Sensor nodes are expected to be battery operated, and it may be deployed in a harsh environment, so sensor nodes may not be able to renew their energy resources. Therefore, energy efficiency is one of the major constraints in accessing the data captured by the sensor network.

Radio communication is a major energy consumer [2]. Pottie and Kaiser conducted a comparison of the cost of computation to communication [2]. They found that the energy cost of transmission 1Kb over a distance of 100 meters is the same as the energy cost of executing 3 million instructions. Numerous techniques have been proposed for extending the lifetime of WSNs, such as Sensor-MAC [7], directed diffusion [8] and hybrid energy-efficient distributed clustering (HEED) [9].

In data management field, the traditional data storage approach is the centralized method, in which sensor nodes send all sensed data to a sink located outside the wireless sensor network for storage [2, 3, 4]. The queries are also forwarded to and processed at the sink. This approach involves unnecessary communication cost, because query is flooded to every node, though some nodes do not have the query data. It may also lead to bottleneck and unbalanced energy consumption, in which the energy consumed by the sensor nodes closer to the sink is much higher than that by the nodes further from the sink [3].

Sensed data also can be stored locally. In local storage schemes, all sensed data are stored locally at the sensor nodes that detected the data. There is no communication cost for storing the sensed data. However, query processing leads to high energy cost. Because data can reside anywhere in the network, a data query must be flooded to all the sensor nodes in the network. Directed Diffusion protocol [10] can be used to save the energy cost when query happens in the WSN. In directed diffusion, all communication is for named data. The sensing task is disseminated throughout the sensor network as an interest for named data. The events matching the interest flow towards the originators of interests along multiple paths. Since different paths have different latencies, a small number of paths with low latencies are chosen for the data forwarding from source node to the query node. Zhang et al. [11] proposed an index-based data dissemination scheme to address the problem. In the scheme, only a small portion of sensed data is queried among a huge amount of sensed data. However, local storage cannot return the query data on time. It needs to find the index node which saves the location information of the storing node, and then get the data from the storing node.

To overcome these weaknesses, distributed data storage approaches are used (e.g. [3, 5]). These techniques differ in the aggregation mechanisms used, but are loosely based on the idea of geographic hashing and structured replication. One such data storage scheme is Data-Centric Storage (DCS) that provides a hash function for mapping from event name to location [4]. DCS hashes the sensed data according to their names. DCS hashes the name to an arbitrary key, and then decides where the data will be stored based on the key. The data having the same name will be saved at the same location of the WSN. The use of geographic hash tables [12] provides a simple way to combine data-centric storage with geographic routing. Every unique event or data attribute name that can be queried for is assigned a unique key, and each data value is stored jointly with the name of the data as a key value pair. A geographic hash function is used to hash each key to a unique geographic location within the sensor network coverage region. The node whose location is closest to this hashed location is the intended storage point for the data.

This paper proposes a distributed data storage scheme (DDS) to accelerate the query speed and reduce the communication consumption. In DDS, each sensor node preprocesses its data using min-wise data processing method [6], and sends its processed data to a certain neighborhood. Because min-wise data processing method can group related data, all related data will be gathered together in a neighborhood after preprocessing. Query is disseminated only to the subsets of the network, without the need to the broadcasted to every sensor node in the network. Besides gathering the related data, DDS can deal with spatial-temporal query. For example, a sensor s_1 senses data d at location (x, y) at time t, which can be presented as 5-dimensional data represented by $<x, y, s_1, d, t>$. Instead of only hashing d and mapping it to a location in the sensor network, the 5-dimensional data is hashed and mapped to a neighborhood. Therefore, a data query can include time t to find events occurred at a specified time. Compared to other distributed data storage schemes, DDS is more advanced in that it offers event data query based on not only data name but also event location and time.

We compared the performance of DDS with centralized data storage scheme. Experiment results show that DDS can dramatically reduce the energy consumption and extend the lifetime of WSN by efficiently disseminating the related data to a neighborhood.

II. Performance of DDS

In a WSN, sensor nodes could be mobile. Node mobility affects whether event data can be correctly saved in the right location. The simulation of this section is conducted on a mobile sensor network with 1000 sensor nodes. The sensor nodes randomly move in the WSN. The moving speed simulated is 0.1 m/s and 1 m/s. Figure 1 illustrates the success rate when the sensor nodes move with different speeds. It demonstrates that DDS achieves more than 75% success rate when sensor nodes are in mobile status. In DDS, sensor field is divided into small grids. A sensor node knows its location and the grid it residents. When a sensor node moves out its grid, it refreshes its database, acquires the event data from the new grid head. Therefore, the sensor node can dynamically obtain the event data should be saved in the grid where the sensor node resides.

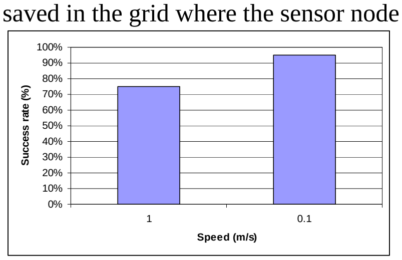

Fig. 1. The success rate of data query with different mobile speed.

DDS efficiently disseminate data in the network such that related event data is mapped into a neighborhood. During data preprocessing phase, DDS hashes all the attributes of a data including name, time and location. Therefore, it can deal with spatial-temporal query. The experimental results show that DDS can efficiently and successfully search related events. In addition, it can dramatically reduce the communication overhead between sensor nodes in event searching.

The future work will be focused on improving the accuracy of hashing data to a neighborhood, and solving the problem of imbalanced data distribution in a WSN.

Acknowledgment

This research was supported in part by U.S. NSF grants CNS-0834592 and CNS-0832109.

References

[1] B. Krishnamachari, "*Networking Wireless Sensors*", Cambridge University, Cambridge, UK, 2005.

[2] G. J. Pottie, W. J. Kaiser, "Wireless Integrated Network Sensors", *Communications of the ACM*, 43(5):51-58, May 2000.

[3] Y. Yao, X. Tang, E. Lim, "In-Network Processing of Nearest Neighbor Queries for Wireless Sensor Networks", *DASFAA*, pp. 35-49, Springer-Verlag LNCS (vol. 3882), 2006.

[4] S. Ratnasamy, D. Estrin, R. Govindan, B. Karp, S. Shenker, L. Yin, F. Yu, "Data-centric Storage in Sensornets", *In 1st Workshop on Sensor Networks and Applications (WSNA)*, Atlanta, GA, September, 2002.

[5] M. Chen, T. Kwon, Y. Yuan, V. Leung, "Mobile Agent Based Wireless Sensor Networks", *Journal of Computers*, vol. 1. NO. 1. 2006.

[6] H. Shen, T. Li, T. Schweiger, "An Efficient Similarity Searching Scheme Based on Locality Sensitive Hashing", *The 3rd International Conference on Digital Telecommunications (ICDT)*, Bucharest, Romania, June 29 – July 5, 2008.

[7] Y. Wei, J. Heidemann, D. Estrin, "An Energy-efficient MAC Protocol for Wireless Sensor Networks", *In Proc. of the 21st International Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM)*, 2002.

[8] C. Intanagonwiwat, R. Govindan, D. Estrin, "Directed Diffusion: A Scalable and Robust Communication Paradigm for sensor networks", *In Proc. of the 6th Annual International Conference on Mobile Computing and Networks (MobiCom)*, 2000.

[9] O. Younis, S. Fahmy, "Distributed Clustering in Ad-hoc Sensor Networks: A Hybrid, Energy-efficient Approach", *In Proc. of the 23rd Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM)*, 2004.

[10] W. R. Heinzelman, A. Chandrakasan, H. Balakrishnan, "Energy-efficient communication protocol for wireless microsensor networks", *In Proc. of the Hawaii International Conference on System Sciences*, January 2000.

[11] W. S. Zhang, G. H. Cao, T. L. Porta, "Data Dissemination with Ring-Based Index for Wireless Sensor Networks", *IEEE Transactions on Mobile Computing*, vol. 6, No. 7, July 2007.

[12] S. Ratnasamy, B. Karp, S. Shenker, D. Estrin, R. Govindan, Y. Li, F. Yu, "Data-centric Storage in Sensornets with GHT, A Geographic Hash Table", *MONET*, 2003.