Towards Energy Saving and Load Balancing Data Aggregation for Wireless Sensor Networks

Wei Liang, Jianbo Xu and Yuanyuan Liu
1Key Laboratory of Knowledge Processing and Networked Manufacturing, School of Computer Science and Engineering, Hunan University of Science and Technology, Xiangtan, 411201, China
2Department of Electrical Engineering, Guangdong Songshan Polytechnic College, Shaoguang, 512126, China

Abstract: In this study, we focus on the development of energy efficient and achievable load balancing mechanisms for wireless sensor networks. Due to resource constraint and tremendous amount of sensors, one possible way of achieving maximum lifetime of the network is applying data aggregation on sensor data. However, existing approaches introduce significant computation and control overheads that often not suitable for sensor networks applications. In view of this, we propose an in-network data aggregation scheme by exploiting the inherent spatial correlation of the sensed data. Each sensor multiplies its reading with a random coefficient and sends the product to the next hop to calculate a weighted sum of all the message. Instead of receiving individual sensor readings, the sink will receive all weighted sum and restore the original data. By doing so, each node only performs one addition and one multiplication in order to compute one weighted sum and consume the same amount of energy. Simulation results demonstrate that the proposed scheme is efficient and outperforms existing schemes in terms of energy gain and network lifetime.

Key words: Wireless sensor networks, energy-efficiency, network lifetime, compressive sampling, spatial correlation

INTRODUCTION

Wireless sensor networks are usually composed of hundreds or thousands of inexpensive, low-powered sensing devices with limited memory, computational and communication resources (Akyildiz et al., 2002). Prior researches explored various aspects of Wireless Sensor Networks (WSNs), including system architecture (Gober et al., 2005), routing (Meng et al., 2008) data gathering (Yu et al., 2008), power efficiency (Wei et al., 2007) and so on. Energy is a primary constraint in the deployment of such WSNs, since sensor nodes are typically powered by small batteries, which cannot be generally charged or recharged (Pottie and Kaiser, 2000).

In-network data aggregation has been a promising solution to reduce the number of transmitters or the length of the packets to be transmitted. On the other hand, the number of sensor nodes deployed could be on the order of hundreds or thousands and data transmissions are accomplished through multi-hop routing from individual sensor nodes to the data sink (Yiek and Ghosal, 2008). Successful deployment of such large-scale sensor networks faces two major challenges.

First, reduce transmission costs and thereby save energy resources. Second, the need for energy consumption load balancing and extend the lifetime of the sensor networks.

The need for global communication cost reduction is obvious because such sensor networks typically are composed of tremendous amount of sensor, generating significant data to be delivered to data sink. It is very much desired to take full advantage of the correlations among the sensor data to reduce the cost of communication. Data aggregation as an important aspect of sensor networks has been widely studied to achieve energy saving. However, existing approaches adopt in-network data aggregation, such as entropy coding (Cristescu et al., 2006) or transform coding (Cianciti et al., 2006), to reduce global traffic. However, these approaches introduce significant computation and control overheads that often not suitable for sensor networks applications.

The need for energy consumption load balancing is also clear because of the required multi-hop data transmission for such large-scale sensor networks. Suppose N sensors form a multi-hop route to the sink.

Corresponding Author: Wei Liang, Key Laboratory of Knowledge Processing and Networked Manufacturing, School of Computer Science and Engineering, Hunan University of Science and Technology, Xiangtan, 411201, China Tel: 86-731-58373103 Fax: 86-731-58373103

409
Obviously, the sensor nodes closer to the sink will soon run out of energy and lifetime of sensor network will be significantly shortened.

This study presents a novel approach to apply compressive sampling theory (Candes and Wakin, 2008) to sensor data aggregation for large-scale wireless sensor networks, successfully addressing the two major challenges as outlined above. First, the proposed scheme is able to achieve substantial sensor data aggregation without introducing excessive computation and control overheads. Second, the proposed scheme is also able to disperse the communication costs to all sensor nodes along a given sensor data gathering route. This will result in a natural load balancing and extend the lifetime of the sensor network.

This study makes three main contributions. First, we take both energy saving and load balancing into consideration in deploying a large-scale sensor networks. Second, the proposed scheme achieves substantial sensor data aggregation without introducing excessive computation and control overheads. Third and more importantly, we take advantage of spatial correlations in sensor readings and distribute the traffic uniform on all sensor nodes, which significantly prolong the network lifetime.

RELATED WORK

The fundamental assumption of in-network data aggregation is that sensor nodes have spatial correlations in their readings. Cristescu et al. (2006) propose a joint entropy coding approach, where nodes use relayed data as side information to encode their readings. By doing so, jointly encoded messages cost less bits than independently encoded messages. However, the joint entropy coding relies heavily on how the routes are organized and performed complex computation in sensors. Ciancil et al. (2006) propose to compress piecewise smooth data through distributed wavelet transform. In doing so, even nodes first broadcast their readings. Upon receiving the readings from both sides, odd nodes compute the high pass coefficients h(•). Then, odd nodes transmit h(•) back and even nodes compute the low pass coefficients l(•). After the transform, nodes transmit coefficients to the sink. Although wavelet de-correlation can be performed for multiple levels, it is not suggested to do so in distributed processing because of the large amount of communication overhead.

Distributed source coding (Chou et al., 2003; Cristescu et al., 2004; Hua and Chen, 2008) intend to reduce complexity at sensor nodes and utilize correlation at the sink. They are based on the Slepian-Wolf coding theory (Slepian and Wolf, 1973), which claims that aggregation of correlated readings, when separately encoded, can achieve same efficiency as if they are jointly encoded, provide that messages are jointly decoded. This important conclusion not only eliminates data exchanges, but also decouples routing from aggregation. After encoding sensor readings independently, each node simply sends the compressed message along the shortest path to the sink. However, a prerequisite of Slepian-Wolf coding is that the global correlation structure needs to be known in order to allocate appropriate number of bits to be used by each node. In view of this, Yuen et al. (2008) adopts a localized Slepian-Wolf coding scheme. Based on the assumption that sensors outside immediate neighbourhood have weak correlation in their readings, a node may only consider its data correlation with one-hop neighbours when determining the size of encoded message.

With the emergence of compressive sampling theory (Baraniuk, 2007; Donoho, 2006; Candes and Wakin, 2008), researchers have seen a new avenue of research in the field of in-network data aggregation. Compressive Wireless Sensing (CWS) (Bajwa et al., 2006) appears to be able to reduce the latency of data gathering in a single-hop network by delivering linear projections of sensor readings through synchronized amplitude-modulated analogue transmissions. Due to the difficulties in analogue synchronization, CWS is less practical for large-scale sensor networks. Rabhet et al. (2006) leverages compressive sampling for data persistence, instead of data gathering, in a WSN. In an overview study, Haupt et al. (2008) also speculate the potential of using compressive sampling theory for data aggregation in a multi-hop WSN. However, no real scheme has been reported based on this initial idea.

THE PROPOSED SCHEME

The objective of compressive data gathering is two-fold: compress sensor readings to reduce global data traffic and distribute energy consumption evenly to prolong network lifetime. Similar to distributed source coding, the data correlation pattern shall be utilized on the decoder end. Besides, aggregation and routing are decoupled and therefore can be separately optimized.

System model: In wireless sensor network, a large number of sensor nodes collect application specific information from the environment and this information is transferred to a central base station where it is processed, analyzed and used by the application. Figure 1 shows a typical routing tree in which the sink has two children.
Fig. 1: Tree-based data aggregation

Each of them leads a sub-tree. Data aggregation and data restore are performed on the sub-tree basis.

To facilitate efficient aggregation, when a node chooses a parent node, it sends a subscribe notification to that node; when a node changes parent, it sends an unsubscribe notification to the old parent. After all nodes acquire their readings, leaf nodes initiate the transmission. Data transmissions are accomplished through multi-hop routing from lower level leaf nodes to the upper level data sink.

**Data aggregation:** We suppose N sensors, denoted as $s_1$, $s_2$, ..., $s_N$. Let, $d_i$ denote the readings obtained by node $s_i$. Consider Figure 1 again as an example, leaf node $s_i$ generates a random number $\Phi_{i,i}$, computes $\Phi_{i,i}d_i$ and transmits the value to $s_j$. The index $i$ denotes the i-th weighted sum ranging from 1 to M, similarly, $s_j, s_k,$ and $s_l$ transmits $\Phi_{j,i}d_j$, $\Phi_{k,i}d_k$, and $\Phi_{l,i}d_l$ to $s_j$. Once $s_j$ receives the four values, it computes $\Phi_{j,i}d_j$, adds to the sum of relayed values and transmits:

$$\sum_{j=1}^i \phi_{ij}d_j$$

To $s_{ij}$, $s_{ij}$ also receives product from $s_j$, denotes as:

$$\sum_{j=1}^i \phi_{ij}d_j$$

then $s_{ij}$ computes $\Phi_{ij}d_{ij}$ and

$$\sum_{j=1}^i \phi_{ij}d_j$$

finally the message containing the weighted sum of all readings in a sub-tree is forwarded to the sink.

Assume that there are N nodes in a particular tree and the sink intends to collect M measurements. Then all nodes send the same number of O (M) messages regardless of their hop distance to the sink. The overall message complexity is O (NM). When M<<N, our scheme transmits less messages than the conventional data collection whose worst-case message complexity is O (N²). More importantly, the transmission load is spread out uniformly so that the lifetime of bottleneck sensors and the entire network is greatly extended.

The i-th weighted sum can be represented by:

$$y_i = \sum_{j=1}^i \phi_{ij}d_j$$

(1)

The sink obtains M weighted sums $\{y_i\}$, $i=1, 2, ..., M$. Mathematically, we have:

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_M \end{bmatrix} = \begin{bmatrix} \phi_{11} & \phi_{12} & \cdots & \phi_{1M} \\ \phi_{21} & \phi_{22} & \cdots & \phi_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ \phi_{M1} & \phi_{M2} & \cdots & \phi_{MM} \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_M \end{bmatrix}$$

(2)

In Eq. 2, each column of $[\Phi_i]$ contains the series of random numbers generated at a corresponding node. In order to avoid transmitting this random matrix from sensors to the sink, we adopt a simple strategy: before data transmission, the sink broadcasts a random seed to the entire network. Then each sensor generates its own seed using this global seed and its unique identification. With a pre-installed pseudo random number generator, each sensor is able to generate the corresponding series of coefficients. These coefficients can be reproduced at the sink given that the sink knows the identifications of all sensors.

Since, the random coefficients $\Phi_i$ are irrelevant to sensor readings, we may treat $d_i$ as a vector. The weighted sums $y_i$ become vectors of the same dimension too. When M<<N, solving a set of M linear equations with N unknown variables is an ill-posed problem. However, in most cases, the sensor field follows a certain structure because of the spatial or temporal correlations. Hence, the signal is sparse in a transform domain. Under this assumption, what requirement M should meet to solve them and how these equations can be solved. Fortunately, the compressive sampling theory has a positive answer to these questions.

**Data restore:** According to compressive sampling theory, a K-sparse signal can be reconstructed from a small number of measurements with a probability close to one. The weighted sums obtained in Eq. 2 are a typical type of measurements. Signal sparsity characterizes the correlations within a signal. An N-dimensional signal is
considered as a K-sparse signal if there exists a domain in which this signal can be represented by K \((K<<N)\) non-zero coefficients.

In a densely deployed sensor networks, sensors have spatial correlations in their readings. Let N sensor readings form a vector \(d=\{d_1, d_2, \ldots, d_N\}^T\), then \(d\) is a K-sparse signal in a particular domain \(\Phi\). Denote \(\Psi=\{\psi_1, \psi_2, \ldots, \psi_k\}\) as the representation basis with vectors \(\{\psi_j\}\) as columns and \(x=\begin{bmatrix} x_1 & x_2 & \ldots & x_N \end{bmatrix}^T\) are the corresponding coefficients. Then, \(d\) can be represented in the \(\psi\) domain as:

\[ d = \sum_{j=1}^{K} x_j \psi_j, \quad \text{or} \quad d = \Phi x \]  

(3)

Compressive sampling theory tells that a K-sparse signal can be reconstructed from M measurements if M satisfies the following conditions (Candes et al., 2006):

\[ M \geq c \cdot \mu^2(\Phi, \Psi) \cdot K \cdot \log N \]  

(4)

where, \(c\) is a positive constant, \(\Phi\) is the sampling matrix as defined in Eq. 2 and \(\mu(\Phi, \Psi)\) is the coherence between sampling basis \(\Phi\) and representation basis \(\Psi\). The coherence metric measures the largest correlation between any two elements of \(\Phi\) and \(\Psi\) and is defined as:

\[ \mu(\Phi, \Psi) = \sqrt{N} \cdot \max_{i \neq j} |<\phi_i, \psi_j>| \]  

(5)

From Eq. 5, we can see that the smaller the coherence between \(\Phi\) and \(\Psi\) is, the less measurements are needed to reconstruct the signal. In practice, using random measurement matrix is a convenient choice, since a random basis has been shown to largely incoherent with any fixed basis and \(M=3K\sim4K\) is usually sufficient to satisfy Eq. 4.

With sufficient number of measurements, the sink is able to reconstruct sensor readings through solving an \(l_1\)-minimization problem:

\[ \min ||x||_1 \quad \text{s.t.} \quad y = \Phi d, \quad d = \psi x \]  

(6)

Suppose \(x^*\) is the solution to this convex optimization problem, then the proposed reconstruction of the original signal is \(d^*=\Psi x^*\). It has been shown that the above \(l_1\)-minimization problem can be solved with Linear Programming (LP) techniques (Donoho, 2006).

In Eq. 6, the \(\Psi\) matrix describes the correlation pattern among sensor readings. It utilizes only in data recovery process and is not required to be known to sensors. In this way, most of the computations are shifted from sensors to the sink. Such asymmetry of computation complexity makes it an appealing choice for WSNs.

**Capacity gain:** The previous subsection illustrated how to gather and recover sensor readings acquired in one time instance. This subsection will investigate the benefit of the proposed scheme from the viewpoint of network capacity, i.e., how it can achieve the capacity over baseline transmission. So, let us look at the baseline transmission first, in Fig. 1, nodes \(s_1, s_2, s_3\), and \(s_4\) transmit their readings \(d_1, d_2, d_3\), and \(d_4\) to \(s_{10}\). \(s_1\) transmits both its reading \(d_1\) and the relayed readings to \(s_{10}\). At the end of the route, \(s_{10}\) transmits all readings in its sub-tree to the sink.

**Definition 1:** We shall define that a route \(e\) is achievable in a data gathering sensor network, if there exists a time instance \(t_e\) and duration \(T\) such that \(t_{\text{start}} + T\] the sink receives \(e\) bits of data generated by each of the sensors \(s_{1, 2, 3, \ldots, N}\). Then, network capacity \(C\) is defined as the supremum of the achievable rates, or \(C=\sup \{\lambda_e\}\).

Note that the traffic pattern in our study is many-to-one. We let all sensors generate data at the same rate and assume that sensor readings acquired at the same time instance are K-sparse. Denote \(\lambda_e\) as the network capacity of baseline transmission. It is achieved when every node is allowed to transmit once every \(n_1\) slots and traffic is evenly distributed among \(n_1\) one-hop neighbours of the sink. Then we have:

\[ \frac{W}{n_1 n_2} \]  

(7)

where, \(W\) is the total transmission capacity of each sensor node. Denote \(\lambda_2\) as the network capacity of the proposed scheme. If the same transmission schedule and routing structure are adopted, we have:

\[ \frac{W}{n_1 n_2} \]  

(8)

From Eq. 7 and 8, we can conclude that our scheme can achieve a capacity gain of \(N/M\) over baseline transmission.

**SIMULATION**

The goal of this simulation study is assessment of the achievable energy gain during the compress-with-transmission process. Unless specified otherwise, our simulations were set up with the following parameters: we adopt a grid topology contains \(N\) nodes, the
network capacity gain keeps $N/M = 5$ in different circumstances. The distance between adjacent nodes is 15 m. The sink is in the middle of the network. The sub-trees contain similar number of sensor nodes, though not exactly the same. We run 10 times for each parameter setting and in each test run, ten messages per node are collected at given intervals. We compare our proposed scheme with chain-based gathering scheme (Chen and Wang, 2008) and clustered-based gathering scheme (Yoon and Shahabi, 2005).

The chain-based scheme (Chen and Wang, 2008) can treated as a Baseline Gathering (BG) scheme since it is expected to deploy in large-scale networks and data transmissions are accomplished through multi-hop routing from individual sensor nodes to the data sink. The clustered aggregation (CAG) technique (Yoon and Shahabi, 2005) forms clusters based on sensing values. By grouping sensors with similar readings, CAG only transmits one reading per group to achieve a predefined error threshold.

Figure 2 presents the total energy consumption as a function of the number of sensors nodes for the BG scheme, the CAG scheme and our scheme. Based on Fig. 2, we observed that the BG scheme has highest energy consumption among the three schemes and our scheme has least energy consumption, besides, our scheme gain more benefits as $N$ increases. For example, when the number of sensors is 2000, the BG scheme and the CAG scheme consume 32.2 J and 19.5 J of total energy, respectively, approximately 5 times and 3 times more than our scheme, respectively, which consumes 6.3 J of total energy cost. While, in the BG scheme, the sensor readings are transmitted along a long-range multi-hop route to the data sink. The closer a sensor is to the sink, the more energy consumed. Thus, the sensor nodes closer to the data sink will soon run out of energy and lifetime of sensor network will be significantly shortened. Similarly, in the CAG scheme, cluster-head sensor nodes consumed more energy in performing complex computation and communication tasks, leading to the depletion of its energy quickly, as a result, this imbalance consumption drops down the system lifetime as well as individual sensor node, when the number of sensor nodes increases, the performance worsen. For our scheme, the energy saved from fewer communication costs and control overheads, leading to lower total energy consumption. We can conclude that the benefit of compressive sampling theory has major effects on total energy saving in our scheme and this verifies the effectiveness of our idea.

In order to gain more insight on the effectiveness of the proposed scheme, we have made a small modification to the above network settings, with 1089 nodes arrange in 33 rows by 33 columns, the sink has four sub-trees and data on each sub-tree can be reconstructed from 55 random measurements. We show energy gain as a function of time ($t$) for the BG scheme, the CAG scheme and the proposed scheme in Fig. 3. As showed in Fig. 3, our scheme has an increases energy gain graceful with the run time, while in both BG scheme and CAG scheme, the energy gain drop dramatically with the run time. We attribute this difference to apply different data gathering patterns in the three schemes. In our scheme, we applying compressive sampling theory to achieves communication efficient and load balance, leads to almost 50% energy gain when $t = 50$, while in the BG scheme and the CAG scheme, they are either introduce complex computation in sensors or require a large
amount of data exchanges, resulting in consume significant energy and load imbalance. Thus, when t = 50, the energy gain of BQ and CAG scheme is 0.5 and 7, respectively. Therefore, the results verify the above analysis and well confirm the effectiveness of the proposed scheme again.

CONCLUSION

In this study, we take the advantage of the inherent spatial correlation among the sensed data and developed novel techniques to aggregate sensed data in large-scale wireless sensor networks. The proposed scheme is able to achieve a capacity gain of N/M over baseline scheme, moreover, our scheme can achieve desired energy saving and load balancing. Compared with existing works, our scheme can significant extend the lifetime of the sensor networks as well as individual sensors. In our future work, we will explore different aggregation algorithms and consider other important metrics, such as data accuracy, latency, fault-tolerant and aggregation ratio.

ACKNOWLEDGMENTS

This study is supported by the Scientific Research Fund of Hunan Provincial Education Department under Grant No. 09C403 (2009.6-2012.6, Hunan University of Science and Technology), National Natural Science Foundation of Hunan Province and Xiangtan united Foundation under Grant No. 09JJ9006 (2009.6-2012.6, Hunan University of Science and Technology), Key Program of Scientific Research Fund of Hunan Provincial Education Department under Grant No. 09A0027 (2009.6-2012.6, Hunan University of Science and Technology).

REFERENCES