Computer Networks 141 (2018) 145–156

Contents lists available at ScienceDirect

Computer Networks

journal homepage: www.elsevier.com/locate/comnet

A sparsity feedback-based data gathering algorithm for Wireless Sensor Networks

Cuicui Lv[a,*], Qiang Wang[b], Wenjie Yan[c], Jia Li[d]

[a] School of Computer and Control Engineering, YanTai University, YanTai, China
[b] Department of Control Science and Engineering, Harbin Institute of Technology, Harbin, China
[c] School of Computer Science and Engineering, Hebei University of Technology, Tianjin, China
[d] The 54th Research Institute of China Electronics Technology Group Corporation, Shijiazhuang, China

A R T I C L E I N F O

Article history:
Received 1 September 2017
Revised 28 May 2018
Accepted 29 May 2018
Available online 30 May 2018

Keywords:
Wireless Sensor Networks
Compressive Sensing
Compressive data gathering
Sparsity feedback
Energy balance

A B S T R A C T

As a means of detecting abnormal events in Wireless Sensor Networks (WSNs), this paper presents a Compressive Sensing (CS)-based algorithm, called Minimum Spanning Tree and Mobile Agent-based Greedy Shortest Path (MST-MA-GSP). The algorithm first of all uses a sparsity feedback mechanism to accurately estimate the sparsity k of the sensor measurements. It then uses Monte Carlo experiments to determine the minimum number of required measurements $M_{\min}$. According to the value of $M_{\min}$, the algorithm adaptively adjusts the number of measurements M in order to maximize its recovery performance. The experiments show that the proposed algorithm is superior to other compressive data gathering (CDG) algorithms in terms of energy balance, whilst the adaptive $M_{\min}$ mechanism guarantees a reconstruction accuracy of at least 99%. Additionally, the sparse binary matrix used in the MST-MA-GSP algorithm offers better recovery of sparse zero-one data than other CDG-based measurement matrices.

1. Introduction

Wireless Sensor Networks (WSNs) have the potential to be widely applied for the purposes of environmental monitoring [1–4]. However, because the sensor nodes in WSNs are usually powered by batteries, their computing and communication abilities are limited. Performing efficient data gathering has therefore become a pressing issue in WSN research.

Compressive Sensing (CS) is a novel signal sampling theory. The core idea is that a finite-dimensional signal can be recovered from a small set of linear measurements when the signal is sparse in a basis or a dictionary. If we assume a WSN with N sensor nodes, each node can be assigned a different ID, ranging from 1 to N. Data readings x of the N sensor nodes can be written as the column vector $x = [x_1, x_2, \dots, x_i, \dots, x_N]^T$, where x_i is the data reading for sensor node s_i. If x has a sparse representation in a given sparsifying basis $\Psi \in \mathbb{R}^{N\times N}$, it is denoted as

$$x = \Psi\Theta. \tag{1}$$

where $\Theta \in \mathbb{R}^{N\times 1}$ is a coefficient vector. If the number of non-zero elements in the vector Θ equals k, x is k-sparse in the sparsifying basis Ψ. The measurements collected by a sink node are termed $y = [y_1, y_2, \dots, y_i, \dots, y_M]^T$, and

$$y = \Phi x. \tag{2}$$

where $\Phi \in \mathbb{R}^{M\times N} (k < M \ll N)$ is the measurement matrix.

As an example, extreme temperature changes can be considered a sparse event [5]. CS is therefore applicable to the gathering of data about such events. To detect these kinds of sparse events, we set about designing a compressive data gathering algorithm, which adaptively adjusts the number of the required measurements M to ensure exact recovery. To do this, we take into consideration the sparsifying basis Ψ and the measurement matrix Φ. In the detection of sparse events, we need to determine whether the abnormal event has occurred, that is, the detection is to monitor the state of the abnormal events. In this case, x can be regarded as sparse and Ψ as the identity matrix. The measurement matrices can be dense random matrices or sparse random matrices [6]. CS theory indicates that it is sufficient to recover x from the measurements y when Φ satisfies the Restricted Isometry Property (RIP) of the matrices. Dense random matrices, such as independent and identically-distributed Gaussian random matrices, have an overwhelming probability of obeying the RIP, provided that the number of measurements M meets the condition $M \geq ck\log(N/k)$, where c is a positive constant. However, such matrices have high computational complexity and occupy more stor-

* Corresponding author.
E-mail address: lvcuicuixuchuan@163.com (C. Lv).

https://doi.org/10.1016/j.comnet.2018.05.022
1389-1286/

age space. In contrast, sparse random matrices are sparser, which means that they can reduce the computational complexity of encoding. Wang et al. [7] have showed that this category of matrices is effective for the recovery of sensory data, with a performance comparable to optimal k-term approximation.

It can be seen from the above that sparse random matrices already have an advantage over dense random matrices. Sparse binary matrices also belong to sparse random matrices. The recovery accuracy for these has proved to outperform existing sparse random matrices. So, in our own work, we take a sparse binary approach. On the basis of this, we develop a compressive data gathering algorithm, the MST-MA-GSP algorithm, which first estimates the sparsity of sensory data. It then puts bounds on the minimum value $M_{\min}$ to ensure accurate recovery. This algorithm can also be combined with the MA-Greedy algorithm to achieve an effective balance of energy consumption amongst the sensor nodes.

The rest of the paper is organized as follows: In Section 2, we review related work regarding compressive data gathering in WSNs. In Section 3, we present a mobile agent-based network model using a CS framework. In Section 4, we analyze the path-planning problem for mobile agents and propose a sparsity feedback-based data gathering algorithm as a solution. In Section 5, we provide details of the experiments and undertake to demonstrate the performance of the proposed MST-MA-GSP algorithm. Its performance is then compared with other compressive data gathering algorithms. In Section 6, we give our conclusions.

2. Related work

In this section, we introduce related work regarding compressive data gathering in WSNs. In [8], sensor nodes synchronously transmitted sensory data to a sink node using a single hop transmission approach and CS was then applied to obtain energy efficient estimation of the sensory data. However, multi-hop transmission wasn't considered as a possible alternative. Rabbat et al. [9] developed an early application of CS for health monitoring where a multi-hop WSN approach was used. There, an assumption was made that most sensor nodes were running normally and that only a small number were corrupted. The corrupted sensors set a flag $x_i = 1$, and the other functioning sensors set $x_i = 0$. The goal was to identify the failed sensor nodes by using random projections of the flag data x_i. However, the random projections were obtained by using a simple gossiping algorithm, which resulted in a heavy traffic load.

Luo et al. [10] have compared traditional approaches to data gathering and Compressive Data Gathering (CDG) using a chain-type topology. The chain-type topology was then expanded to a tree-type topology, thus providing the first complete CDG design for large-scale WSNs. In this design, random projections were first obtained by the product of Gaussian random coefficients and sensory data, and then each sensor node transmitted M random projections (or weighted sums of random projections) to a sink node. When the sink node received the M random projections (or weighted sums), it utilized reconstruction algorithms to recover the data. Direct CS coding on each sensor node was termed plain-CS aggregation (plain-CS) (see [11]). Xiang et al. [11] put forward a data aggregation technique called Hybrid-CS. Here, the sensor nodes started CS coding only once the outgoing data flows from the sensor nodes exceeded a certain number of measurements M. A Minimum Spanning Tree (MST) and data aggregation were jointly used in order to minimize the total energy consumption. The difficulty with Hybrid-CS is deciding how to divide the sensor nodes into an aggregator set and a forwarder set. In Hybrid-CS, the data were transmitted using a tree-type structure. When the data volumes transferred by each sensor node were very different, it was possible for an energy imbalance to arise among the sensor nodes. To solve this problem, Ebrahimi and Assi [12] have proposed a Minimum Spanning Tree Projection (MSTP) algorithm and an eMSTP algorithm, which aim to minimize the transmission cost by using M forwarding trees. In the MSTP algorithm, M projection nodes first constructed a tree root at their own position. They then added their interest nodes (the visited nodes) to the tree using MST and Breath-First-Search (BFS) techniques. When each projection node received the weighted sums from its interest nodes, it transmitted them to a sink node by the shortest path. Unlike the MSTP algorithm, the eMSTP algorithm chose the sink node as the root of the tree. Experiments demonstrated that the two algorithms outperformed Hybrid-CS with respect to the overall network cost and load balancing. In [13], the same authors further studied the interaction between the forwarding tree construction and link scheduling.

CS in conjunction with routing was exploited in [14], with the sink node receiving random projections of sensory data through geographic routing. The routing matrix was a measurement matrix whose elements were generated by a pseudo random number generator. Each sensor node looked for a next node within the range of nodes that would provide the largest geographical advancement towards the sink node. Lee et al. [15] took into account communication cost and presented a Low Coherence Projection for efficient Routing (LCPR) algorithm to enact the route design. Motivated by the techniques presented in [16], this algorithm iteratively made a greedy choice for the next node by identifying whether the other nodes within the communication range of the current node could minimize the intermediate coherence of Ψ with a partial measurement matrix $\Phi_{partial}$. However, when each path contained only a small number of nodes, the results indicated that the effectiveness of CS might be limited. Chen and Wassell [17] have proposed an energy-efficient signal acquisition approach for monitoring 1-D environmental information. This approach didn't require prior information about the sparsity. The sparsity was obtained instead through data sampled at the highest rate. A Sampling Rate Indicator (SRI) feedback scheme was then introduced to minimize the number of samples under the condition that the reconstruction quality indicator (RQI) fell within a desired range. In a similar approach to Chen and Wassell [17] and Wang et al. [18], Fazel et al. [19] have proposed a Random Access Compressed Sensing (RACS) algorithm for underwater environmental monitoring. To prolong network life, RACS used random sensing as a sampling procedure and simple random access for the channel access phase.

Mamaghanian et al. [20] have assessed the potential of CS for energy-efficient Electrocardiogram (ECG) data acquisition and compression on resource-constrained Wireless Body Sensor Networks (WBSN) platforms. They explored three different methods (quantized Gaussian random sensing, pseudorandom sensing, and sparse binary sensing) to implement the random sensing matrix. The approach using a sparse sensing matrix was superior in terms of execution time. Implementing a Gaussian random matrix was too complex and time consuming. Besides, for the Mixed Signal Processor 430 (MSP430), their designated task wasn't achievable in real time. Based on a sparse binary matrix, Li and Qi [21] have proposed a distributed compressive sparse sampling (DCSS) algorithm. The algorithm selected M encoding nodes from N sensor nodes and sent their sensory data to a fusion center (FC) using the shortest path. In approaches related to WSNs deployed in a unit square, Zheng et al. [22] have suggested a random walk-based non-uniform method, whilst Quan et al. [23] have developed a neighbor-aided compressive sensing (NACS) scheme. For WSNs deployed in square area or circular areas, Nguyen and Teague [24] have put forward a compressive sensing based random walk data collection (CSR) algorithm, where each CS measurement was forwarded to a Base Station (BS) in both one hop fashion (D-