Imparting Efficient Clustering Method for Wireless Sensor Networks Using Compressive Sensing

K. Gori Basha*, K. Sendhilnadhan

Department of Computer Science and Engineering, P.S.V. College of Engineering and Technology, Krishnagiri, India
*Email: goribasha@gmail.com

Abstract—Compressive sensing (CS) can reduce the number of data transmissions and balance the traffic load throughout networks. However, the total number of transmissions for data collection by using pure CS is still large. The hybrid method of using CS was proposed to reduce the number of transmissions in sensor networks. In this paper, we propose a clustering method that uses hybrid CS for sensor networks. We first propose an analytical model that studies the relationship between the size of clusters and number of transmissions in the hybrid CS method, aiming at finding the optimal size of clusters that can lead to minimum number of transmissions. Then, we propose a centralized clustering algorithm based on the results obtained from the analytical model. Finally, we present a distributed implementation of the clustering method. Extensive simulations confirm that our method can reduce the number of transmissions significantly.

Keywords— Wireless sensor networks, compressive sensing, data collection, clustering, Digital Signature, Cluster node.

I. INTRODUCTION

In many sensor network applications, such as environment monitoring systems, sensor nodes need to collect data periodically and transmit them to the data sink through multihops. According to field experiments, data communication contributes majority of energy consumption of sensor nodes. It has become an important issue to reduce the amount of data transmissions in sensor networks. The emerging technology of compressive sensing (CS) opens new frontiers for data collection in sensor networks and target localization in sensor networks. The CS method can substantially reduce the amount of data transmissions and balance the traffic load throughout the entire network. In data gathering without using CS, the nodes close to tree leaves relay fewer packets for other nodes, but the nodes close to the sink have to relay much more packets. By using CS in data gathering, every node needs to transmit M packets for a set of N data items. That is, the number of transmissions for collecting data from N nodes is MN, which is still a large number. Hybrid approaches were proposed in [8], [10]. In the hybrid method, the nodes closet the leaf nodes transmit the original data without using the CS method, but the nodes close to the sink transmit data to sink by the CS technique. Applied hybrid CS in the data collection and proposed an aggregation tree with minimum energy consumption. The previous works use the CS method on routing trees. Since the clustering method has many advantages over the tree method, such as fault tolerance and traffic load balancing, we use the CS method on the clustering in sensor networks. The clustering method generally has better traffic load balancing than the tree data gathering method. This is because the number of nodes in clusters can be balanced when we divide clusters. The remainder of this paper is organized as follows an overview of the clustering method by using hybrid CS for data collection. An analytical model for analyzing the relationship between the size of clusters and the number of transmissions, and determining the optimal cluster size. A centralized algorithm for sensor nodes clustering with minimum number of transmissions and a distributed clustering algorithm and its implementations.

II. RELATED WORK

Compressive Sensing (CS) is a method for finding sparse solutions to underdetermined linear systems. It has led to a completely different approach to distributed data compression in WSNs. Compared with traditional DSC, CS-based data compression moves most computation from sensor nodes to the sink, which makes it a good fit for in-network data suppression and compression. Over the past years, a variety of CS-based methods have been devised to solve the data gathering problem in WSNs. summarizes the potential of applying CS to the data gathering problem in multi-hop WSNs. They defined a pair of well-designed measurement matrix and representation basis in order to achieve incoherence and sparsity at the same time. Discussed the energy and latency performance of varied data collection algorithms using compressive sensing. Baronet al. studied joint sparsity models and joint data recovery methods based on CS. As mentioned previously, Luo et al. proposed CS-based CDG to reduce communication cost and prolong network lifespan. Despite that CDG leads to significantly less traffic and longer lifetime than Centralized Exact, there is still much room for improvement.

III. SENSOR NODES CLUSTERING FOR HYBRID COMPRESSIVE SENSING

In our method, sensor nodes are organized into clusters, and each cluster has a cluster head, represented by the solid square. Sensor nodes in each cluster transmit their original data to the CH without using CS. We assume each CH knows the projection vectors in measurement matrix of all nodes within its cluster. In real systems, the measurement coefficient can be generated using a pseudorandom number generator seeded with the identifier of the node. Thus, given the
identifiers of the nodes in the network, the measurement matrix can be easily constructed at CHs or the sink locally. The measurement matrix can be decomposed into sub matrices, one for each cluster.

**Figure 1** SENSOR NODES

### IV. PROPOSED SCHEME OF WORK

The two metrics to evaluate the performance of the clustering with hybrid CS proposed in this paper: the number of transmissions which is required to collect data from sensors to the sink, and the reduction ratio of transmissions (reduction ratio for short) of our method compared with other methods. Four other data collection methods are considered. In the clustering without CS method, the same cluster structure to our method is used, but CS is not used. In the shortest path tree (SPT) without CS, the shortest path tree is used to collect data from sensors to the sink.

We propose a clustering method that uses the hybrid CS for sensor networks. The sensor nodes are organized into clusters. Within a cluster, nodes transmit data to the cluster head (CH) without using CS. A data gathering tree spanning all CHs is constructed to transmit data to the sink by using the CS method. One important issue for the hybrid method is to determine how big a cluster should be. If the cluster size is too big, the number of transmissions required to collect data from sensor nodes within a cluster to the CH will be very high. But if the cluster size is too small, the number of clusters will be large and the data gathering tree for all CHs to transmit their collected data to the sink will be large, which would lead to a large number of transmissions by using the CS method.

**A. Cluster**

In this module we construct CWSN. In which sensor nodes are grouped into clusters, and each cluster has a cluster-head (CH) sensor node, which is elected autonomously. Leaf (non-CH) sensor nodes, join a cluster depending on the receiving signal strength and transmit the sensed data to the BS via CHs to save energy. The CHs perform data fusion, and transmit data to the BS directly with comparatively high energy. In addition, we assume that, all sensor nodes and the BS are time synchronized with symmetric radio channels, nodes are distributed randomly, and their energy is constrained.

**B. Minimum Transmission Clustering**

In this module we construct which expand as Identity-Based digital Signature scheme in this module base station generates a master key and public parameter for private key generator and give that all to sensor node in cluster. Then sensor node generate private key use a private string after that send message along with timestamp and signature. Here signature is generated by signing key. In receiver side message accepted if verification of signature is valid otherwise reject. In our method, within a cluster, each sensor node transmits its data to its designated CH via the shortest path. The routes that sensor nodes use to send their data to the CH form a shortest path tree in each cluster. The total number of intra cluster transmissions is the sum of the distance of all sensor nodes to their CHs. Thus, the clustering problem for minimizing intra cluster transmissions. Data collected from sensor nodes is compressed by the CS method at the CHs. The data projections generated at each CH are forwarded to the sink in M rounds along the backbone tree. At each CH in the backbone tree, it aggregates its own data projection with the projections received from other CHs by using the CS method and forwards the aggregated projection upward toward the sink along the tree.

**C. Centralized Clustering**

We adopt an efficient method that iteratively closes to the near-optimal solution. Our algorithm starts from an initial set of CHs, which is randomly selected. At each iteration, the algorithm proceeds following1. Connect sensor nodes to their closest CHs. Tie-break arbitrarily. 2. For each cluster, choose a new CH, such that the sum of the distances from all nodes in this cluster to the new CH is minimized. 3. Repeat the above two steps until there is no more change of the CHs. This algorithm converges quickly.

**Figure 2** Number of nodes
The simulations show that it takes four or five iterations on average for the algorithm to compute the CHs of clusters. We use a minimum spanning tree (MST)-based method to compute the backbone tree that connects all CHs and the sink. Shortest path in G. After a CH is elected, the CH broadcasts an advertisement message to other sensor nodes in the sensor field, to invite the sensor nodes to join its cluster. An advertisement message carries the information: the identifier and location of the CH, and the number of hop that the message has traveled.

V. PERFORMANCE EVALUATION

We use two metrics to evaluate the performance of the clustering with hybrid CS proposed in this paper: the number of transmissions which is required to collect data from sensors to the sink, and the reduction ratio of transmissions (reduction ratio for short) of our method compared with other methods. Four other data collection methods are considered. In the clustering without CS method, the same cluster structure to our method is used, but CS is not used. In the shortest path tree (SPT) without CS, the shortest path tree is used to collect data from sensors to the sink. In the SPT with hybrid CS, the shortest path tree is used to collect data from sensors to the sink, and CS is used in the nodes who has more than M descendant nodes (including itself).

In the optimal tree with hybrid CS, a tree having minimum transmissions is used. It is computed by the greedy algorithm. It is obvious that the number of transmissions of our method is significantly smaller than that of the clustering method without using CS.

The reason is that data are compressed using the CS method at the CHs in our method. Each node on the backbone tree does M transmissions for the inter cluster data gathering. It is significantly less than the number of transmissions of the method without using CS. The number of transmissions of our method is also visibly smaller than that of SPT with the hybrid CS method. This is because in the cluster structure, sensor nodes transmit data to their cluster head, which is located nearly at the center of the cluster, while in the SPT, sensor nodes transmit data to the nodes near to the sink, which results in more transmissions than our method.

VI. RESULT

We The number of transmissions of our method is slightly larger than that of the optimal tree with the hybrid CS method. However, the cluster structure can be organized in the distributed manner, while the optimal tree with hybrid CS is computed in the centralized manner. In addition, our distributed algorithm is fault tolerant. The greedy algorithm iteratively computes an optimal tree with the input of network topology. The network topology may change due to the node failures or the power outage. Once the network topology changes, the resulting tree may not be energy efficient anymore. While in our distributed algorithm, the sink computes the approximate locations of central points of the cluster-areas based on the geographic area of the sensor field, instead of the network topology. Our algorithm can easily reorganize the cluster structure that has the similar quality in terms of the number of data transmissions when the compressive ratio is 10, our method reduces the number of transmissions by about 60 percent compared with clustering without the CS method. It reduces the number of transmissions by about 50 percent compared with SPT without the CS method. In addition, it reduces the number of transmissions by about 30 percent when the number of nodes is 1,200, compared with SPT with the hybrid CS method. The reduction ratio does not drop as the number of nodes increases. It demonstrates our method is scalable in large-scale networks. the reduction ratio of our method decreases only about 10 percent compared with the case that the compressive ratio is 10.

VII. CONCLUSION AND FUTURE WORK

In this paper, we used hybrid CS to design a clustering-based data collection method, to reduce the data transmissions in wireless sensor networks. The information on locations and distribution of sensor nodes is used to design the data collection method in cluster structure. Sensor nodes are organized into clusters. Within a cluster, data are collected to the cluster heads by shortest path routing; at the cluster head, data are compressed to the projections using the CS technique. The projections are forwarded to the sink following a backbone tree. We first proposed an analytical model that studies the relationship between the size of clusters and number of transmissions in the hybrid CS method, to find the optimal size of clusters that can lead to minimum number of transmissions. Then, we proposed a centralized clustering algorithm based on the results obtained from the analytical model. Finally, we present a distributed implementation of the clustering method.

Extensive simulations confirm that our method can reduce the number of transmissions significantly. When the number of measurements is 10th of the number of nodes in the network. The simulation results show that our method can reduce the number of transmissions by about 60 percent compared with clustering method without using CS. Mean while, our method can reduce the number of transmissions up to 30 percent compared with the data collection method using SPT with the hybrid CS. Even for the non homogenous networks in the irregular sensor field, our method can significantly reduce data transmissions compared with these data collection methods.

REFERENCES


