International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Index Copernicus Value (2013): 6.14 | Impact Factor (2013): 4.438

The Hybrid Compressive Sensing Data Collection Method in Cluster Structure for Efficient Data Transmission in WSN

Bibi Ameena[1], Prof. Mallanagouda Biradar[2]

[1]PG Student Department of CNE, Appa Institute of Engineering and Technology, Kalaburagi, Karnataka, India
[2]Department of CSE, Appa Institute of Engineering and Technology, Kalaburagi, Karnataka, India

Abstract: ***Wireless sensor network consists of large number of wireless node that are responsible for sensing processing and monitoring environmental .These sensor nodes are battery operated. Clustering is a standard approach for achieving efficient and scalable performance in WSN. Compressive sensing (CS) reduces the amount of data transmission and balances the load of traffic throughout the network. In WSN the total number of transmission for collection of data using pure CS is large. Therefore we are using the hybrid CS method to reduce the number of transmission in sensor network. Though the previous work uses the CS on routing trees method. In this paper, we propose a clustering method that uses hybrid CS for WSN. Clusters are formed by group of sensor nodes where each sensor node in cluster sends their data to cluster head(CH) without using CS, further using CS each CH transmits their data to sink. An analytical model shows the relationship between the number of transmission and cluster size in hybrid CS method. We define a new cost function, with the objective of simultaneously minimizing the intra-cluster distance and optimizing the energy consumption of the network. Obtaining the results from analytical model a centralized clustering algorithm is being proposed. Further a distributed implementation of clustering method is presented. Finally our new approach of energy optimization shows the amount of energy optimized during data transmission with CS and without CS method. Simulation confirms that our method reduces the number of transmissions significantly.***

Keywords: Wireless sensor networks, clustering, compressive sensing, data gathering, energy optimization

1. Introduction

Wireless sensor networking is an emerging technology that has a wide range of potential applications such as environmental monitoring, smart spaces, medical systems and robotic exploration. Improving the performance of WSNS is a recurring issue of wireless networking community.

In many sensor network applications, such as Industrial monitoring, environmental monitoring systems, sensor nodes need to collect data periodically and transmit them to the data sink through multihops. According to experiments, data communication takes majority of energy consumption of sensor nodes. So 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 technique 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. Compressive sensing address the inefficiencies by directly acquiring a compressed signal representation without going through the intermediate stage of acquiring N samples.

The basic of CS works as follows, as shown in Fig.1. Suppose that the system consists of one sink node and N sensor nodes for collecting data from the field. Let p denote a vector of original data collected from sensors. Vector x has N elements, one for each sensor. p can be represented by Ψs, i.e. p=Ψ s, where Ψ is an N * N transform basis, and s is a vector of coefficients. If there are at most k(k << N) nonzero elements in s, p is called k-sparse in the Ψ domain. When k is small, instead of transmitting N data to the sink, we can send a small number of projections of p to the sink, that is y q=Φp, where Φ is an M * N (M << N) random matrix (called the measurement matrix) and q is a vector of M projections. At the sink node, after collecting q the original data p can be recovered by using l1-normminimization.

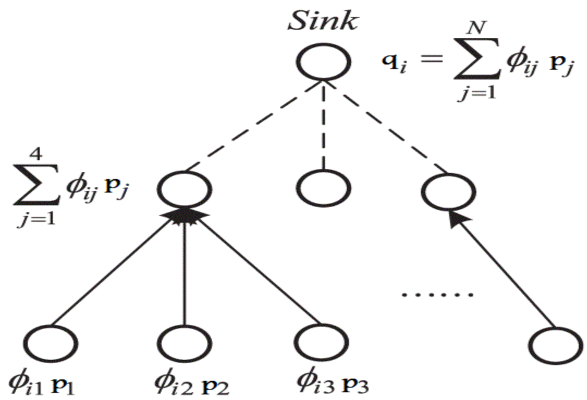

Figure 1: Data collection method using pure CS in tree structure

In this paper we propose a clustering method that uses the hybrid CS for sensor networks. Clustering is a technique which increases the scalability and survivability of nodes, the main aim of clustering is to divide a network into a set of individual and limited nodes that can easily control [5]. 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. In this regard, 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 are introducing a process of energy optimization, as wireless sensor networks (WSNs) are mainly characterized by their limited and non-replenishable energy supply. Hence, the need for energy efficient infrastructure is becoming increasingly more important since it impacts upon the network operational lifetime. Sensor node clustering is one of the techniques that can expand the lifespan of the whole network through data aggregation at the cluster head[16]. We propose a centralized clustering algorithm based on the results obtained from the analytical model. Then we present a distributed implementation of the clustering method. . Finally our new approach of energy optimization shows the amount of energy optimized during data transmission with CS and without CS method.

2. Related Work

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 close to 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. Xiang et al. [10] 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. In addition, the previous works ignored the geographic locations and node distribution of the sensor nodes. While in sensor networks, the information of node distribution can help the design of data gathering method that uses less data transmissions.

An Analysis of a Large Scale Habitat Monitoring Application [2]. Habitat and environmental monitoring is a driving application for wireless sensor networks. An analysis of data from a second generation sensor networks deployed during the summer and autumn of 2003. During a 4 month deployment, these networks, consisting of 150 devices, produced unique datasets for both systems and biological analysis. The focuse on nodal and network performance, with an emphasis on lifetime, reliability, and the static and dynamic aspects of single and multi-hop networks. The results collected to expectations set during the design phase were able to accurately predict lifetime of the single-hop network, but underestimated the impact of multi-hop traffic overhearing and the nuances of power source selection. While initial packet loss data was commensurate with lab experiments, over the duration of the deployment, reliability of the backend infrastructure and the transit network had a dominant impact on overall network performance. Finally, the physical design of the sensor node has been evaluated based on deployment experience and a <i>post mortem</i> analysis. The results shed light on a number of design issues from network deployment, through selection of power sources to optimizations of routing decisions.

An Introduction to Compressive Sampling [3]**.**Conventional approaches to sampling signals or images follow Shannon's theorem the sampling rate must be at least twice the maximum frequency present in the signal (Nyquist rate). In the field of data conversion, standard analog-to-digital converter (ADC) technology implements the usual quantized Shannon representation the signal is uniformly sampled at or above the Nyquist rate. This article surveys the theory of compressive sampling, also known as compressed sensing or CS, a novel sensing/sampling paradigm that goes against the common wisdom in data acquisition. CS theory asserts that one can recover certain signals and images from far fewer samples or measurements than traditional methods use.

Compressive Data Gathering for Large-Scale Wireless Sensor Networks[4].The design is to apply compressive sampling theory to sensor data gathering for large scale wireless sensor networks. The scheme developed in this research is expected to offer fresh frame of mind for research in both compressive sampling applications and large-scale wireless sensor networks. The scenario in which a large number of sensor nodes are densely deployed and sensor readings are spatially correlated. The proposed compressive data gathering is able to reduce global scale communication cost without introducing intensive computation or complicated transmission control. The load balancing characteristic is capable of extending the lifetime of the entire sensor network as well as individual sensors. Furthermore, the proposed scheme can cope with abnormal sensor readings73gracefully and this novel compressive data gathering process has been tested on real sensor data and the results show the efficiency and robustness of the proposed scheme.

Wireless Sensor Network Clustering Using Particles Swarm Optimization for Reducing Energy Consumption [5]. Wireless sensor networks (WSN) is composed of a large number of small nodes with limited functionality. The most important issue in this type of networks is energy constraints. In this area several researches have been done from which clustering is one of the most effective solutions. The goal of clustering is to divide network into sections each of which has a cluster head (CH). The task of cluster heads collection, data aggregation and transmission to the base station is undertaken. In this paper, we introduce a new approach for clustering sensor networks based on Particle Swarm Optimization (PSO) algorithm using the optimal fitness function, which aims to extend network lifetime. The parameters used in this algorithm are residual energy density, the distance from the base station, intra-cluster distance from the cluster head. Simulation results show that the proposed method is more effective compared to protocols such as (LEACH, CHEF, PSO-MV) in terms of network lifetime and energy consumption.

Does Compressed Sensing Improve the Throughput of Wireless Sensor Networks?[6]. The WSNs require independent energy resources and therefore, energy consumption is the most important factor to determine the lifetime of wireless sensors. The CS optimizes energy consumption which is an important factor in WSNs. The CS states that sparse signal of information in WSNs can be exactly reconstructed from a small number of random linear measurements of information in WSNs. The CS provides a new approach to mathematical complexities especially where sparse information is applied. CS tends to recover data vector ∈with N number of information form data vector ∈ with M number of information such that M≪. In fact, CS offers a stable information matrix that does not depend in any way on the information signal, aims to provide a survey of selected topics of CS in WSNs

3. Overview of the Paper

Assumptions:

- The sensor nodes are uniformly and independently distributed in a sensor field.
- All sensor nodes have the same fixed transmission power and transmission rate.
- Each sensor node is aware of its own geographic location, which can be obtained via attached GPS or some other sensor localization techniques. The location information is used in the distributed implementation.

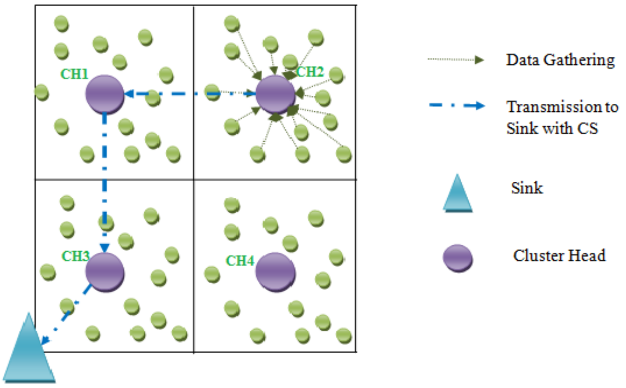

Figure 2: The hybrid CS data collection in cluster structure

In our method, sensor nodes are organized into clusters, and each cluster has a cluster head, represented by the solid square as shown in Fig. 2. 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 Φ) of all nodes within its cluster. In real systems, the measurement coefficient Φij can be generated using a pseudorandom number generator seeded with the identifier of the node υj [5]. 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 Φ can be decomposed into submatrices, one for each cluster. Let Φ^{Hi} denote the submatrix for íth cluster. For íth cluster, let CH_i denote the cluster head and x^{Hi} denote the data vector of the cluster. The CH_i is able to compute the projections of all data x^{Hi} collected from the nodes in its cluster on the submatrix, that is $\Phi^{Hi}\, x^{Hi}$. The CH_i generates M projections from the data within its cluster by using the CS technique. The value of M is determined by the number of nodes N and the sparsity level of the original data. It then forwards them to the sink in M rounds along a backbone tree that connects all CHs to the sink. Taking the sensor nodes in Fig. 2 as an example, all sensors nodes are divided into four clusters. The four cluster heads, CH_1 CH_2 CH_3 CH_4 are connected by a backbone tree to the sink. Data vector x can be decomposed as $[x^{H1}\ x^{H2}\ x^{H3}\ x^{H4}]^T$ and matrix Φ can be written as $[\Phi^{H1}\ \Phi^{H2}\ \Phi^{H3}\ \Phi^{H4}]$.

$$
\begin{aligned}
y &= \Phi x \\
&= \begin{bmatrix} \Phi^{H_1} & \Phi^{H_2} & \Phi^{H_3} & \Phi^{H_4} \end{bmatrix} \begin{pmatrix} x^{H_1} \\ x^{H_2} \\ x^{H_3} \\ x^{H_4} \end{pmatrix} \qquad (1) \\
&= \sum_{i=1}^{4} \Phi^{H_i} x^{H_i}.
\end{aligned}
$$

As shown in (1), the projection of all data in the network on the measurement matrix Φ is the sum of the projections generated from the clusters. Thus in each round, the CH aggregate its own projection and the projections received from its children CHs in the same round and forwards it to the sink following the backbone tree. When the sink receives all M rounds of projections from CHs, the original data for all sensor nodes can be recovered.

Levels of Transmission Used

There are two levels of transmissions in our clustering method using the hybrid CS: intracluster transmissions that do not use the CS technique and intercluster transmissions that use the CS technique. The data size in intercluster transmissions is the same as the data in intracluster transmissions. Thus, reducing the number of transmissions can effectively reduce the energy consumption of sensor nodes. For intracluster transmissions, we simply let sensor nodes transmit their data to the CH following the shortest path routing (in terms of number of hops). For intercluster transmissions, we construct a minimal cost (in terms of number of hops) backbone tree that connects all CHs to the sink and transmit the data projections along this backbone tree.

An important task of our method is to determine the cluster size. As cluster size increases, the number of intracluster transmissions would increase sharply. But when decreasing the cluster size, the number of clusters would increase and the number of intercluster transmissions would increase. Thus, there exists an optimal cluster size that minimizes the total number of data transmissions in the hybrid CS method. Our task is to determine the optimal cluster size and design a distributed clustering method, such that the total number of transmissions is minimized.

4. Modules and Algorithms

4.1 Centralized Clustering Algorithm:

Given the network G = (V, E)
Step1: Select C CHs from the set V of N sensor nodes and divide the sensor nodes into C clusters
Step2: Construct a backbone routing tree that connects all CHs to the sink.

Our algorithm starts from an initial set of CHs, which is randomly selected. At each iteration, the algorithm proceeds following steps:

Step1: Connect sensor nodes to their closest CHs.
Step2: For each cluster, choose a new CH, such that the sum the distances from all nodes in this cluster to the new CH is minimized.
Step3: Repeat the above two steps until there is no more change of the CHs.

This algorithm converges quickly. The simulations show that it takes four or five iterations on average for the algorithm to compute the CHs of clusters.

4.2 Distributed Implementation

This section presents a distributed implementation of the clustering method. We assume that:

1) Every sensor node knows its geographic location. This location information can be obtained via attached GPS or some other sensor localization techniques.
2) The sink knows the area of the whole sensor field, but does not need to know the location information of all sensor nodes. This is a reasonable assumption, since in most applications of the sensor networks; the sink usually knows the area that has sensors deployed for surveillance or environmental monitoring.

In our distributed algorithm, the sink divides the field into C cluster-areas, calculates the geographic central point of each cluster-area, and broadcasts the information to all sensor nodes to elect CHs. The sensor node that is the closest to the center of a cluster-area is selected to be the CH. The CHs then broadcast advertisement messages to sensor nodes to invite sensor nodes to join their respective clusters.
.

4.3 Cluster Head Election

Given the geographic location of the central point of a cluster-area, the sensor node that is the closest to the central point will become the CH. Since the sensor nodes do not know who is the closest to the central point of a cluster area, and we do not know if there is a sensor node falling into the close range of the central point, we let all nodes within the range of Hr from the center be the CH candidates of the cluster, where r is the transmission range of sensors. The value of H is determined such that there is at least one node within H hops from the central point of a cluster. To elect the CH, each candidate broadcasts a CH election message that contains its identifier, its location and the identifier of its cluster. The CH election message is propagated not more than 2H hops. After a timeout, the candidate that has the smallest distance to the center of the cluster among the other candidates becomes the CH of the cluster. In the extreme case that no sensor node falls within H hops from the central point so that there is no CH for this cluster-area, the nodes in this cluster area accept the invitation from neighboring CHs and become members of other clusters. Thus, no node will be left out of the network.

4.4 Sensor Node Clustering

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. The hop count is initialized to be 0. When a sensor node receives an advertisement message, if the hop count of message is smaller than that recorded from the same CH, it updates the information in its record including the node of previous hop and the number of hop to the CH, and further broadcasts the message to its neighbor nodes; otherwise, the message is discarded. After the advertisement of CH is complete, each non-CH node decides which cluster it joins. The decision is based on the number of hops to each CH. The routing from a sensor node to its CH follows the reverse path in forwarding the advertisement message.

4.5 Energy Optimization

In WSN applications, the energy used by a node consists of the energy consumed by computing, receiving, transmitting, listening for messages on the radio channel, sampling data and sleeping. Wireless sensor networks (WSNs) are mainly characterized by their limited and non-replenishable energy supply. Hence, the need for energy efficient infrastructure is becoming increasingly more important since it impacts upon the network operational lifetime. Sensor node clustering is one of the techniques that can expand the lifespan of the whole network through data aggregation at the cluster head. We define a new cost function, with the objective of simultaneously minimizing the intra-cluster distance and optimizing the energy consumption of the network [16].In our protocol we will be maintaining a initial energy of all nodes when simulation process happens the energy level of all nodes will get updated and after the simulation process completes it scans the nodes energy values which will be viewed in pictorial way. We will assume energy level of nodes without CS to some performance N which will be compared with the proposed technique. Then the result in x-graph shows that the proposed technique uses less energy than the existing technique.

4.6 Performance evaluation

All sensor nodes are randomly scattered with a uniform distribution. Randomly select one of the deployed nodes as the source node. The location of the sink is randomly determined.

We evaluate our proposed method with respect to the following metrics: PDR, E2E latency.

No. of transmission: is the number of report messages the sink receives from all the cluster head nodes.

End to end latency: It refers to the time taken for a packet to be transmitted across a network from source to sink node.

These parameter values are recorded in the trace file during the simulation by using record procedure. The recorded details are stored in the trace file. The trace file is executed by using the Xgraph to get graph as the output.

5. Simulation Matrices

We use two metrics to evaluate the performance of the clustering with hybrid CS proposed in this paper:

1) The number of transmissions which is required to collect data from sensors to the sink, shown in fig.3 and
2) The reduction ratio of transmissions (reduction ratio for short) of our method compared with other methods, shown in fig.4. 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 that 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 in [7]. We use another two metrics to evaluate the performance of the clustering with hybrid CS proposed in this paper:

1) E2E Delay with CS and without CS, end to end delay refers to the time taken for a packet to be transmitted across a network from source to sink node, shown in fig.5.
2) Energy optimization, this process shows how much energy is optimized with CS and without CS, shown in fig.6.

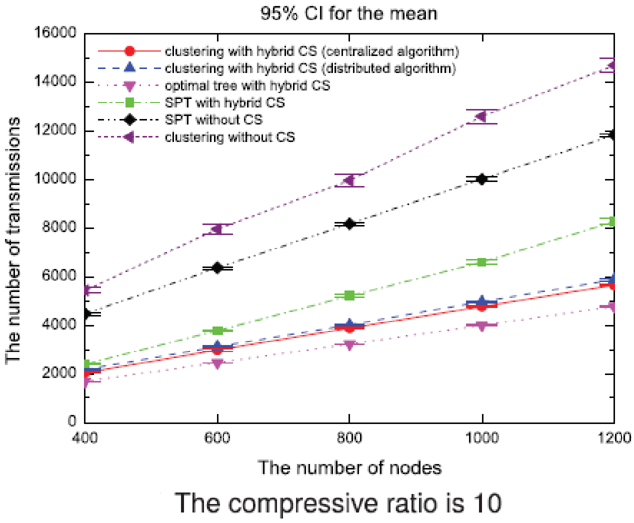

Figure 3: The number of transmissions of data collection methods. The bars around the symbols on the lines represent the 95% confidence interval.

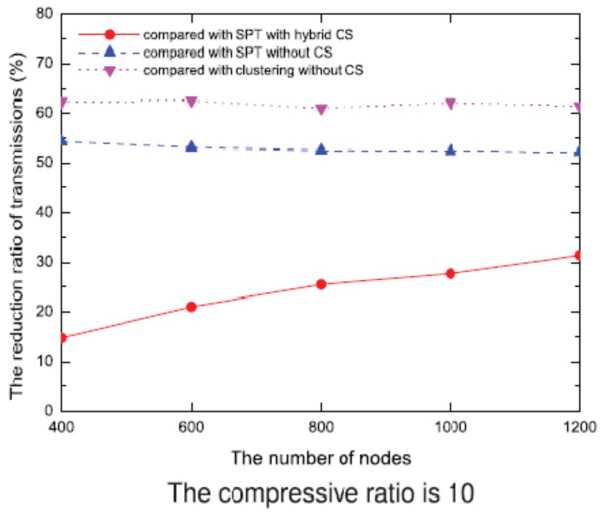

Figure 4: The reduction ratio of transmissions of clustering with the hybrid CS method compared with other methods.

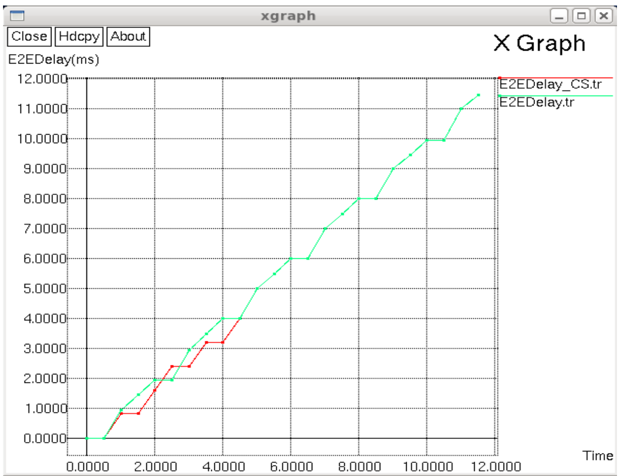

Figure 5: The end to end delay with CS and without CS

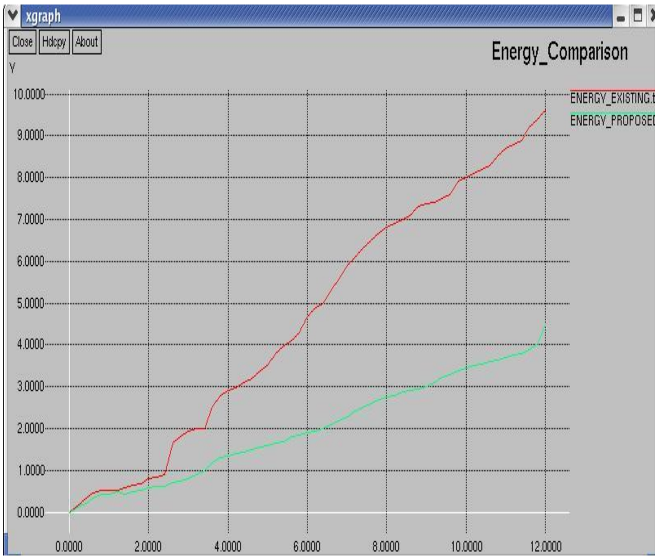

Figure 6: The process of energy optimization with CS and without CS.

6. Conclusions

The hybrid CS is used 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. An analytical model that studies the relationship between the size of clusters and number of transmissions in the hybrid CS method is proposed, to find the optimal size of clusters that can lead to minimum number of transmissions. Then a centralized clustering algorithm is proposed based on the results obtained from the analytical model. Finally a distributed implementation of the clustering method is presented. 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. Meanwhile, our method can reduce the number of transmissions up to 30 percent compared with the data collection method using SPT with the hybrid CS. Extensive process of energy optimization, shows the amount of energy optimized during data transmission with CS and without CS method.

References

[1] Ruitao Xie and Xiaohua Jia, ”Transmission-Efficient Clustering Method for Wireless Sensor Networks Using Compressive Sensing”, IEEE Transaction on parallel and distributed system, vol 25, no. 3, march 2014.

[2] R. Szewczyk, A. Mainwaring, J. Polastre, J. Anderson, and D.Culler, “An Analysis of a Large Scale Habitat Monitoring Application,” Proc. ACM Second Int’l Conf. Embedded Networked Sensor Systems (SenSys ’04), pp. 214-226, Nov. 2004.

[3] E. Candes and M. Wakin, “An Introduction to Compressive Sampling,” IEEE Signal Processing Magazine, vol. 25, no. 2, pp. 21-30, Mar. 2008.

[4] C. Luo, F. Wu, J. Sun, and C.W. Chen, “Compressive Data Gathering for Large-Scale Wireless Sensor Networks,” Proc. ACMMobiCom, pp. 145-156, Sept. 2009.

[5] Amin Rostami and Mohammad Hossin Mottar, ”Wireless Sensor Network Clustering Using Particles Swarm Optimization for Reducing Energy Consumption”, International Journal of Managing Information Technology (IJMIT) Vol.6, No.4, November 2014.

[6] J. Luo, L. Xiang, and C. Rosenberg, “Does Compressed Sensing Improve the Throughput of Wireless Sensor Networks?” Proc.IEEE Int’l Conf. Comm (ICC), pp. 1-6, May 2010.

[7] L. Xiang, J. Luo, and A. Vasilakos, “Compressed Data Aggregation for Energy Efficient Wireless Sensor Networks,” Proc. IEEE Sensor, Mesh, and Ad Hoc Comm. and Networks (SECON ’11), pp. 46-54, June 2011.

[8] F. Fazel, M. Fazel, and M. Stojanovic, “Random Access Compressed Sensing for Energy-Efficient Underwater Sensor Networks,” IEEE J. Selected Areas Comm., vol. 29, no. 8, pp. 1660-1670,Sept. 2011.

[9] J. Wang, S. Tang, B. Yin, and X.-Y. Li, “Data Gathering in WirelessSensor Networks through Intelligent Compressive Sensing,” Proc. IEEE INFOCOM, pp. 603-611, Mar. 2012.

[10] B. Zhang, X. Cheng, N. Zhang, Y. Cui, Y. Li, and Q. Liang, “Sparse Target Counting and Localization in Sensor Networks Based on Compressive Sensing,” Proc. IEEE INFOCOM, pp. 2255-2263, Apr.2011.

[11] E. Candes, J. Romberg, and T. Tao, “Robust Uncertainty Principles: Exact Signal Reconstruction from Highly Incomplete Frequency Information,” IEEE Trans. Information Theory, vol. 52,no. 2, pp. 489-509, Feb. 2006.

[12] J. Tropp and A. Gilbert, “Signal Recovery from Random Measurements via Orthogonal Matching Pursuit,” IEEE Trans. Information Theory, vol. 53, no. 12, pp. 4655-4666, Dec. 2007.

[13] M. Youssef, A. Youssef, and M. Younis, “Overlapping Multihop Clustering for Wireless Sensor Networks,” IEEE Trans. Parallel and Distributed Systems, vol. 20, no. 12, pp. 1844-1856, Dec. 2009.

[14] S. Soro and W.B. Heinzelman, “Cluster Head Election Techniques for Coverage Preservation in Wireless Sensor Networks,” Ad Hoc Networks, vol. 7, no. 5, pp. 955-972, 2009.

[15] G. Sun, J. Chen, W. Guo and K. Liu, "Signal Processing Techniques in Network-Aided Positioning: A Survey of State-of-the-Art Positioning Designs," IEEE Signal Processing Magazine, vol. 22, no. 4, pp. 12-23, July 2005.

[16] N. M. Abdul Latiff, C. C. Tsimenidis, B. S.Sharif, ”Energy Aware Clustering for Wireless Sensor Networks Using Particles Swarm Optimization,” The 18th Annual IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC'07)