IJCTA, 9(17) 2016, pp. 8651-8657

Data Aggregation using Compressive Sensing for Improved Network Lifetime in Large Scale Wireless Sensor Networks

Deepa Puneeth*, Ruthwik R and Muralidhar Kulkarni*****

ABSTRACT

In large scale Wireless Sensor Networks(WSN's) the amount of data generated is enormous. The data has to be processed efficiently before it reaches the Base Station (BS) by using an efficient routing algorithm as well as data aggregation methods. The nodes in WSN's are randomly deployed, the data emerging from these nodes are highly correlated either spatially or temporally. The data aggregation scheme should employ simple encoding since the sensor nodes are battery operated. The proposed method discusses about a data aggregation scheme using Compressive Sensing(CS) technique which makes use of correlation among the sensor nodes. Our primary focus is to increase the lifetime of the overall network. The underlying protocols used are Low-energy adaptive clustering hierarchy (LEACH) and Multi-threshold adaptive range clustering (M-TRAC). We have computed several network parameters for different network configuration. The reconstruction algorithm is sufficiently robust against noise. The reconstruction of the data is done using greedy method and L1 norm regularization. The implementation of the algorithm is done using the real data-set from Intel Lab. Simulation results validate the data aggregation scheme guarantees data accuracy and doubles the network lifetime.

Keywords: Data aggregation, compressed sensing, norm regularization, greedy method

I. INTRODUCTION

WSN's are composed of tiny battery operated sensor nodes which are responsible for sensing physical parameter, and has the ability to compute and communicate the sensed data for a real time application. The sensor nodes have limited power source as well as the replacement of batteries for sensor nodes is virtually impossible for most applications since the nodes are often deployed in large numbers in harsh and inaccessible environments [1]. In order to extend the life time of the network the battery lifetime has to be extended. Minimizing energy consumption is the major requirement in the design of a WSN's. Careful management of energy consumption of each of the sensor node is an important parameter which is to be considered while designing the network. Sensor nodes in WSN's consume energy during sensing, processing and transmission. But these nodes consume more energy for data transmission/reception than the energy for processing [2]. By using appropriate data aggregation techniques we can reduce the redundant data circulating in the network. The data obtained from nodes placed on same geographical area may be highly correlated. First level of data aggregation is achieved by employing an efficient routing algorithm [3]. The sensor nodes are grouped together into clusters, with each cluster consisting of a Cluster Head (CH) and the member nodes. Further every CH gathers the sensed data and then transmits it to the Base station (BS) either directly or through multi-hop communication. The Cluster Head (CH) generally does the aggregation

* Department of Electronics and Communication Engineering, National Institute of Technology, Surathkal, Karnataka, India, *E-mail: deepapuneethk@gmail.com*

** Department of Electronics and Communication Engineering, R V College of Engineering, Mysore Road, Bangalore, India, *E-mail-ruthwikrruth@gmail.com*

*** Department of Electronics and Communication Engineering, National Institute of Technology, Surathkal, Karnataka, India, *E-mail: mkuldce@gmail.com*

process. Thus CH exploits the inter-correlation among the nodes. Further, the data can be aggregated by selecting suitable data compression techniques.

Normally, the data is acquired in full length, and then the redundant data is eliminated using compressing schemes. But if we are able to sample the data by including only the portion which does not get discarded, the processing and transmission burdens would reduce significantly. Data aggregation techniques are tightly coupled with how data is gathered at the sensor nodes as well as how packets are routed through the network and have a significant impact on energy consumption and overall network efficiency [4]. Data compression techniques can be employed to conserve energy in WSN's, thus improving the network lifetime. Data compression schemes reduce data size before transmitting in the wireless medium which is one way to minimize expensive data transmissions. The savings due to compression directly translate into lifetime extension for the network nodes. Both the local single node that compresses the data as well as the intermediate routing nodes benefit from handling less data [5]. The usefulness of data compression algorithms comes into picture if the execution of compression algorithms does not require an amount of energy greater than the one saved in reducing transmission. Data aggregation techniques are extremely useful in target tracking applications since movements of targets will enable multiple detectors to generate the same data which contains large amount of redundant data [6]. An important task associated with data aggregation technique is to reduce the message complexity. This results in reduction of bandwidth usage of the network and the energy exhaustion of the sensor nodes by taking advantage of the correlations of the sensor nodes with itself and with other sensors. The proposed method of data compression using the compressive sensing technique uses simple encoding algorithm which does not burden the sensor nodes. It uses non-adaptive linear projections to compress the data. By using an optimization technique the data is recovered at the Base Station (BS). The complexity is less at the acquisition stage, as the nodes are battery operated but more at the receiver.

II. DATA COMPRESSION PARADIGM

Consider a one dimensional discrete time signal 'x' [x_1, x_2, x_3, ... x_N] which is of finite length and real valued, $x \in R^N$. When we consider large scale WSN's the data vector may be very large. If we could generate a compressed version of x, and at the receiver recover x, within a reasonable accuracy, it offers attractive benefits to the overall network. There can be several strategies to achieve this. One obvious choice would be distributed source coding, but it needs correlations among the nodes to be known priorly. Many of the times prior knowledge of the precise correlations of the data is unavailable. This motivates us to use a in-network processing and a compression technique.

2.1. Compressed Sensing (CS) Acquisition Model

Compressed sensing is an approach to simultaneous sensing and data compression that assures distributed compression in WSN's. We can recover certain signals from fewer samples than required conventional methods if the signal is sparse. Suppose the signal is 'k', sparse (i.e it has 'k' nonzero entries) where 'k' is smaller than the signal length . Usually natural signals have concise representations when we express using a proper basis. Sparsity is a measure as to how effectively we can acquire signals non-adaptively. Signals which are not strictly sparse but might posses few large and many small coefficients are termed as compressible signals. Under certain assumptions, it is possible to recover signals when the number of available measurements are smaller than the signal's original length.

$$y = Ax \tag{1}$$

where matrix 'A' is $m \times n$ sensing matrix and $y \in R^m$. If the signal is not sparse in the domain which is acquired, the signal can be expressed using a proper basis ζ.

$$x = \zeta.\, s \tag{2}$$

where ζ is the representation basis. 's' is the signal acquired.

The coherence between the sensing basis and representation basis should be as low as possible. According to the available literature, A can be Guassian or Bernouli, Fourier but usually random matrices are largely incoherent with any fixed basis ζ. Random measurements can be used for the -sparse signals in any basis as long as 'A' obeys the condition below [7].

$$x = k \log \frac{n}{s} \tag{3}$$

When the reconstruction is done using L1-norm regularization also called Basis Pursuit (BP) [8] the k -sparse signals can be exactly recovered by ensuring

$$m \geq 4k \tag{4}$$

An important criteria for robust compressed sensing is that measurement matrix must preserve the important features in the signal of interest. Otherwise we will not be able to get back the original information from an under determined system. This criteria can be verified by checking the Restricted Isometry Property (RIP) of reconstruction matrix ψ [9].

RIP is defined on isometry constant δ_k of a matrix, which is the smallest number such that holds for all -sparse vectors 'x'.

$$(1-\delta_k)\|x\|_{l_2}^2 \leq \|\psi x\|_{l_2}^2 \leq (1+\delta_k)\|x\|_{l_2}^2 \tag{5}$$

We can say that a matrix obeys the RIP of order if is not too close to one. All subsets of columns from matrix must be nearly orthogonal as well the sparse signal is not in null space of the matrix which is used as a sensing matrix.

2.2. CS Reconstruction Model

Recovery process must take the 'm' measurements in the vector 'y' and the measurement matrix 'A' as well as the representation basis ζ and recover the 'N' length signal 'x' or its sparse vector.

$$y = \Psi . s \tag{6}$$

where $\Psi = A.\zeta$

The above system is an under determined system, thus we will obtain Infinite number of solutions. Reconstruction algorithms try to solve (6), either using convex relaxation method or by using greedy algorithms. In Convex Relaxation, optimization is done through linear programming [3] to recover the data. Greedy method tries to solve the reconstruction in an iterative manner. The algorithm selects columns of in a greedy manner. At every step, the columns of that correlates with y is selected. At the same time least square error is minimized.

III. SYSTEM MODEL

In order to estimate the transmission energy cost, we have incorporated a standard transmission model [10]. In this model, the energy per bit for transmission over a wireless link is a function of the distance between a transmitter and a receiver. Let E_{Tx}(N,d) and E_{Rx}(N) be the energy consumed for transmitting or receiving a 'N' bit message over a distance 'd', respectively.

$$E_{Tx}(N,d) = E_{T-elec} \times N + \varepsilon_{amp} \times N \times d^2 \tag{7}$$

$$E_{Rx}(N) = E_{R-elec} \times N \tag{8}$$

$E_{T-elec} E_{R-elec}$ are the energy consumption for transmitting and receiving one bit message, and ε_{amp} is the transmission amplifier. Initial simulations were conducted by considering the parameters given in the Table 1.

Table 1
Simulation Parameters

Parameters	*Typical values*
Network area	100 m × 100 m
Number of sensor nodes	100
Position of the sink node	(50,100)
Initial energy of node	0.5 J
E_{T-elec}	50 nJ/bit
E_{R-elec}	50 nJ/bit
ε amp	100 pJ/bit
Size of the data packet	128 bytes

We have used Intel Lab data set in order to validate the results. The underlying protocols used are LEACH [4] and M-TRAC [1] which uses variable transmission ranges. Each CH uses random Gaussian matrix as the measurement matrix to compress the sensed data. Then the compressed data is transmitted to the BS. With the addition of data aggregation using compressive sensing scheme, throughput and network lifetime has been improved.

We have used different compression ratios to compress the data and the corresponding error in the recovered data has been obtained using the following relation.

$$\varepsilon = \frac{\|x(n) - \hat{x}(n)\|_2}{\|x(n)\|_2} \tag{9}$$

where $x(n)$ is raw data $\hat{x}(n)$ is the recovered data. $\|x(n)\|_2 = \sqrt{\sum_{k=1}^{k=n} |x(n)|}$

IV. RESULTS AND ANALYSIS

In this paper, three important performance indices: network lifetime, throughput and reconstruction error are used for analysis of the network. The lifetime of the network using LEACH with and without Compressed Sensing (CS) technique is compared and the plot is as shown in Figure 1. The plot verifies that the lifetime of the network is significantly increased with the use of CS technique. The underlying protocol used for comparison is LEACH. To validate the compression algorithm, we have considered CS for LEACH, MTRAC by considering 100 nodes random deployment in 100 m ×100 m area. In every simulation round, the data from the nodes are transmitted to the CH, and the CH transmits the compressed data to the BS. Then the data is reconstructed using L1 norm regularization and greedy methods. From Figures 1 we can infer that using CS, network lifetime is improved.

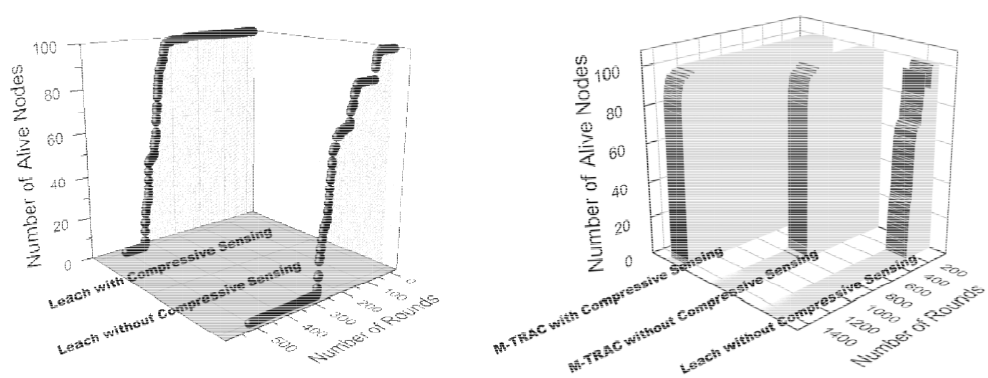

Figure 1: Network life improvement in LEACH, M-TRAC using Compressed Sensing

Figure 2 gives the summary of all dead nodes. LEACH and M-TRAC performance improves with the incorporation of CS. The data from sensor nodes has been reduced significantly which results in keeping the nodes alive for longer duration.

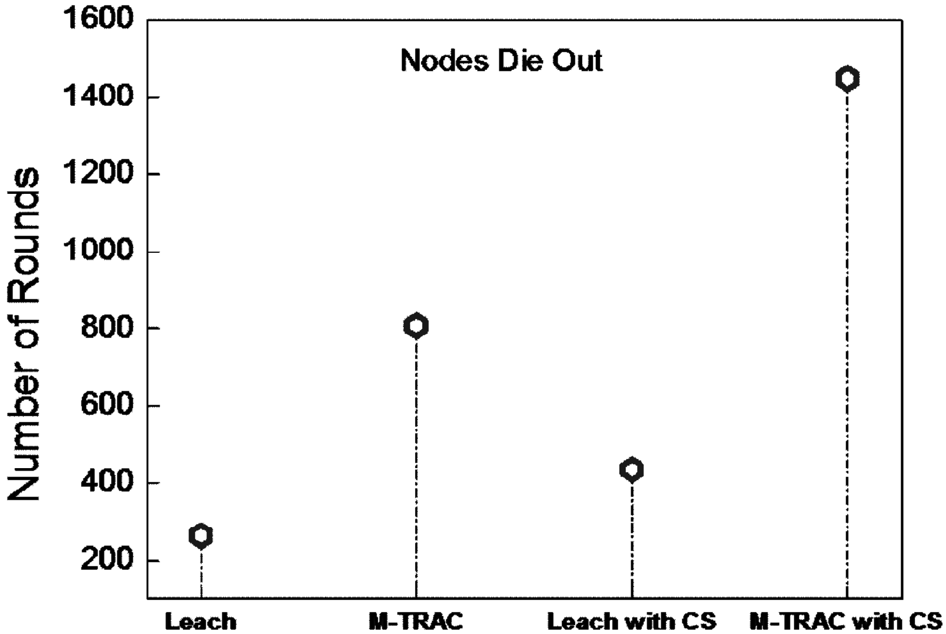

Figure 2: Summary of dead nodes in LEACH and M-TRAC

From Figure 3 we can conclude that with the inclusion of CS as a data aggregation scheme, we can reduce the amount of redundant data along with network lifetime improvement. Reduction of the redundant data results in increased throughput and increased lifetime of the network. Data has been compressed using several compression ratio's and reconstructed at the BS. Corresponding reconstruction error can be calculated using equation [9] which is shown in Figure 4(a). Simulations have been carried out using temperature and humidity parameters of the Intel dataset. Intra-node correlations have been considered. The values for 3

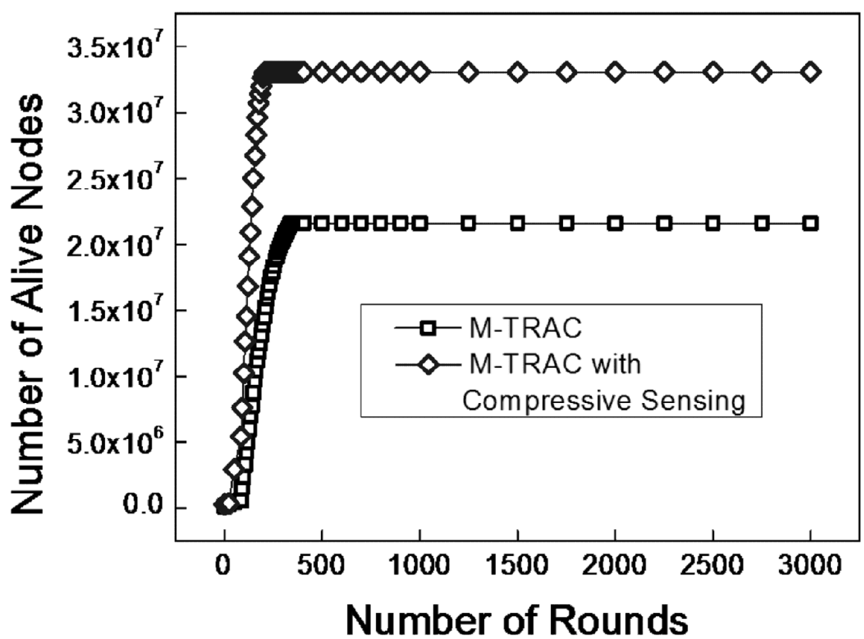

Figure 3: Number of Packets in Base station

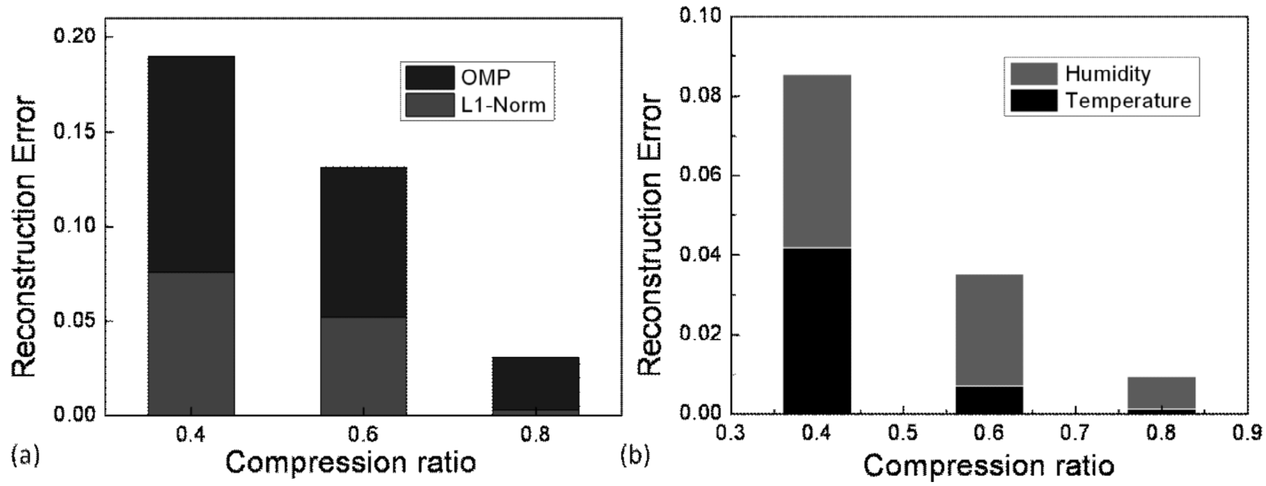

Figure 4: (a) Comparison of Reconstruction error using L1 Regularization and OMP (b) Reconstruction error for temperature and humidity values using L1-norm regularization

days have been considered, collected around 3000 samples. The reconstruction has been done by considering different compression ratios. The real data are not usually sparse, but compressible. In this simulation, we have transformed the data into an appropriate sparsifying basis, then the data has been compressed.

Figure 5 shows the original and reconstructed values for different compression ratios for temperate and humidity values respectively. Depending on the correlation among the nodes and the amount of sparse data, there will be variation in the reconstruction error. Figure 4 (b) shows the reconstruction errors for different compression ratio's for temperature and humidity values.

IV. CONCLUSION

The proposed algorithm deals with data compression and reconstruction based on compressed sensing, which uses correlation between and within nodes. The traffic cost of the network has been reduced which reflects on the network lifetime. The performance metrics namely the lifetime of the network, the throughput and the reconstruction error was analyzed and the results have been validated. It was observed that the network lifetime is doubled in both LEACH and MTRAC. M-TRAC algorithm proved to be a good algorithm over LEACH in terms of the lifetime of the network. Usage of CS in M-TRAC and obtaining the enhancement in both lifetime and throughput further validates the statement that CS is a universal sampling method and

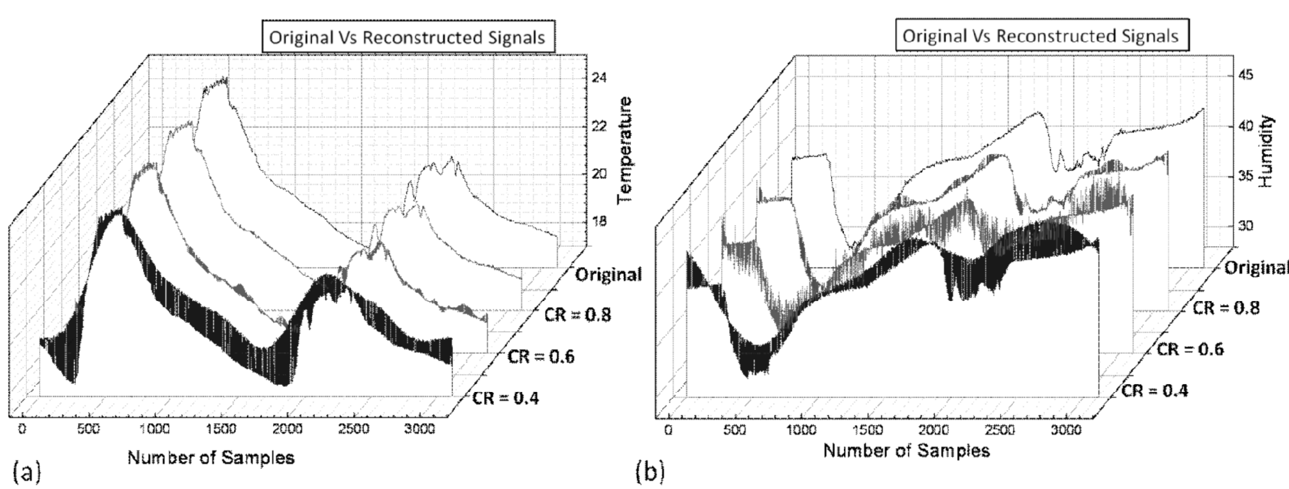

Figure 5: Reconstruction of temperature and humidity values using L1-norm regularization

can be applied to any of the networks. Reconstruction is done by using L1 norm regularization and greedy methods. The outcome of our analysis is that applying CS may not bring the improvement in all cases, but could be applicable where correlation among node exists. The temperature values are highly correlated, thus the reconstruction error is comparably less. Analysis of correlations among nodes must be done before applying CS methods. Future work would be to develop CS methods for multi-hop networks.

REFERENCES

[1] Shivaprakasha, K. S., Muralidhar Kulkarni, and Nishant Joshi. "Improved network survivability using multi-threshold adaptive range clustering (M-TRAC) algorithm for energy balancing in wireless sensor networks." Journal of High Speed Networks 19, no. 2 (2013): 99-113. Springer, pp. 207-215, 2004.

[2] Candès, Emmanuel J., and Benjamin Recht. "Exact matrix completion via convex optimization." *Foundations of Computational mathematics* 9, no. 6 (2009): 717-772. 2005.

[3] P Muruganathan, Siva D., Daniel CF Ma, Rolly I. Bhasin, and Abraham O. Fapojuwo. "A centralized energy-efficient routing protocol for wireless sensor networks." *IEEE Communications Magazine* 43, no. 3 (2005): S8-13.

[4] Akkaya, Kemal, and Mohamed Younis. "A survey on routing protocols for wireless sensor networks." *Ad hoc networks* 3, no. 3 (2005): 325-349.

[5] Sadler, Christopher M., and Margaret Martonosi. "Data compression algorithms for energy-constrained devices in delay tolerant networks." In *Proceedings of the 4th international conference on Embedded networked sensor systems*, pp. 265-278. ACM, 2006.

[6] Woo-Sung Jung, Keun-Woo Lim, Young-Bae Ko, and Sang-Joon Park.Efficient clustering-based data aggregation techniques for wireless sensor networks. Wireless Networks, 17(5):1387–1400, 2011.

[7] Candès, Emmanuel J., Justin Romberg, and Terence Tao. "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information." *IEEE Transactions on information theory* 52, no. 2 (2006): 489-509.

[8] Candes, Emmanuel J., Michael B. Wakin, and Stephen P. Boyd. "Enhancing sparsity by reweighted –! 1 minimization." *Journal of Fourier analysis and applications* 14, no. 5-6 (2008): 877-905.

[9] Candes, Emmanuel J. "The restricted isometry property and its implications for compressed sensing." *Comptes Rendus Mathematique* 346, no. 9 (2008): 589-592.

[10] Heinzelman, Wendi B., Anantha P. Chandrakasan, and Hari Balakrishnan. *"An application-specific protocol architecture* for wireless microsensor networks." IEEE Transactions on wireless communications 1, no. 4 (2002): 660-670.