



RESEARCH ARTICLE

Energy-Efficient Data Gathering in Wireless Sensor Network Using Compressive Sensing

Yang Song¹, Tianyi Liu²

¹School of Software, Beijing University of Technology, Beijing, China

¹ wings_yang@emails.bjut.edu.cn

² liutianyi@ emails.bjut.edu.cn

Abstract— *Wireless Sensor Network (WSN) is wildly used for a range of applications, one of the most important issues is to improve network lifetime of the sensor node powered by battery. Inspired by Compressive Sensing theory, we proposed an energy-balanced scheme of data gathering denoted by Changeable Probability Compressive Sensing (CPCS). In the proposed approach, we use Compressive Sensing to reduce the transmission costs. Moreover, we design a method for sensors to dynamically adjust their probability involved in every CS measurement according to the received average in-network residual energy. With extensive experiments over real sensory data, we demonstrate that the proposed method can effectively prolong the network lifetime.*

Keywords— *“Wireless Sensor Network”, “Compressive Sensing”, “Data collection”, “Energy-balanced”, “residual power”*

I. INTRODUCTION

Wireless Sensor Network (WSN) has been approved to be an efficient tool to acquire an interest physical phenomenon for long time. Unfortunately, the large amounts of data transmission shorten the lifetime of WSN. Using spatial-temporal properties in sensory data, Compressive Sensing (CS) based data gathering is a promising approach which can compress original signal and recover it in sink. [1] present the first complete scheme to apply CS for data gathering in WSN. [2], [3] and [4] used the method of random sparse projections based on CS to greatly reduce the transmission cost.

Not only reducing energy consumption, the energy balance should also be considered. Energy imbalance is widely existing in WSN, although CS has efficiently prolonged network lifetime, to some extent, this phenomenon is still happening. Some of sensors run out their energy quickly will lead to damage of network structure. To solve this problem [5] and [6] tried to combine CS with the routing protocol to balance the energy costs of different transmission paths.

To address the above challenges, we proposed an energy-balanced method of CS based data collection for WSN. In our scheme, we use sparse random projections based on CS to reduce the transmission costs. Moreover, using the fault diagnosis mechanism exists in the WSN applications [7], we present a design that every sensor can adjust the probability of data transmission involved in the next round of measurement according to received average residual power. The proposed method, based on CS and fault diagnosis mechanism referred as Changeable Probability Compressive Sensing (CPCS). Profit from the scheme above, the simulation experiment proves that this method can balance the energy costs of every sensor in the network and prolong the lifetime of whole network.

II. SYSTEM MODEL

Consider a grid network, which consists of sensors uniformly distributed on a two-dimensional plane, with R sensors in x and y directions. The sensors are separated by distance d in each direction. Defining the coverage area A of a network as the total area covered by the sensors, in the grid network $A = N d^2$. The sink is deployed in $(1,1)$. The network is deployed to monitor a physical phenomenon x over a long period of time. The map of the process over the entire sensor network is denoted by X_{2D}

$$X_{2D} = \begin{bmatrix} x_{11} & \cdots & x_{1C} \\ \vdots & \vdots & \vdots \\ \cdots & x_{rC} & \cdots \\ \vdots & \vdots & \vdots \\ x_{R1} & \cdots & x_{RC} \end{bmatrix} \quad (1)$$

In the process of networking, the sensors involved in CS measurement send their data by specific route protocol. In order to quickly discover the failure of the network, the node periodically sends its health information to the sink node. The sink node receives the information from this information and stores them in order to use it when necessary.

III. PAGE STYLE

Assuming that the sink has obtained sensory data expressed as formula (1). In order to ensure that the spatial correlations between adjacent nodes and try to obtain the perfect reconstruction, we use formula (2) to transform X_{2D} to $R * C$ vectors X

$$X = Vec(x) = [x_{11} \cdots x_{R1} \ x_{R2} \cdots x_{21} \cdots x_{1C} \cdots x_{RC}]^T \quad (2)$$

Usually most of the natural signals X can be expressed in a specific domain, if X is a k -sparse signal, that is it has only k nonzero elements or $(N - k)$ smallest elements. The signal can be extracted for s by:

$$y = \Phi X = \Phi \psi s \quad (3)$$

Where Φ is $M \times N$ ($M \ll N$) sparse random projection matrix shown as follow

$$\Phi_{ij} = \sqrt{z} \begin{cases} +1 & \text{with prob. } 1/2z \\ 0 & \text{otherwise} \\ -1 & \text{with prob. } 1 - 1/2z \end{cases} \quad (4)$$

ψ is $N * N$ dimensional sparse basis. Reference [4] had proved matrices Φ and Ψ satisfied the restricted isometric property (RIP). At the beginning of CS in-network data gathering, sink broadcast the basic sampling rate $z = 1/p_{base}$. Sensors generated their Φ_{ij} by a pseudo-random number generator. Usually Φ_{ij} is different from each node and the data routing algorithm is beyond the research scope of this paper, we just assume that sink can receive each sensor's data, and then try to recover the original signal. It has been proved that the signal s can be exactly recovered by solving the following minimum l_1 -norm optimization problem with very high probability. The recovery process can be expressed as

$$\hat{s} = argmin ||s||_1 \text{ s.t. } y = \Phi \psi s \quad (5)$$

Usually, in the existing WSN applications [7], sensors send out their healthy status periodically for network administrator to detect the failure of the network in time. Every healthy status includes parents' RSSI, residual power, package drops and so on. We use this function to improve the CS measurement process. Using the health information of the nodes in the network, the sink maintains a health status table of all the sensors in the network. After each compressed sensing reconstruction, sink feedbacks the average residual energy of all sensors. The main processes of our approach are shown in Fig.1.

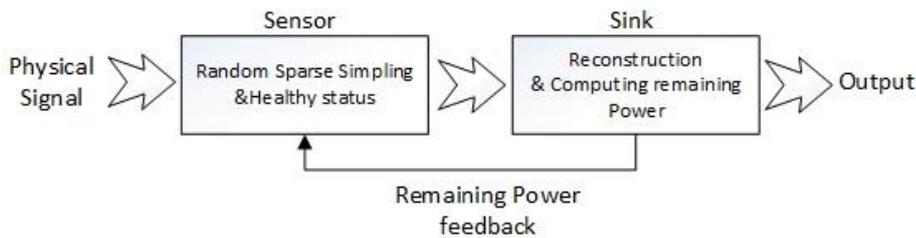


Fig 1: The proposed CS approach

Assuming that in the n th compressed sensing process, the residual energy of the node j in the sink's healthy status database is expressed as $power_j(n)$. Then after data reconstruction, sink calculates the residual average energy of the nodes after the current round of sampling:

$$power_{avg}(n) = (\sum_{j=1}^N power_j(n))/N \quad (6)$$

When the calculation is finished, sink broadcasts the $power_{avg}(n)$ to all sensors for sensors adjusting the probability of participating in the next CS measurement based on the average energy of the current network. The feedback process can be described by the algorithm 1.

Algorithm 1: The method of average residual energy feedback

Input: Sensor's residual power $power(n)$

Output: , Basic sampling rate p_{base} .

1. finish recover original signal;
 2. for $i=1$ to N
 3. search $power_i(n)$ form database;
 4. $power_{sum}=power_{sum} + power_i(n)$;
 5. end for
 6. $power_{avg}(n) = power_{sum}/N$; /*Calculate the Average Remained Power*/
 7. feedback $power_{avg}(n)$ to all sensors via broadcast;
-

In order to take full advantage use of sensor's energy, we proposed an energy-balanced scheme for sensors to adjust their probability of participated in next measurement. At the beginning of the measurement i , all sensors have obtained same distributed random seeds and sampling rate $1/z = p_{base}$. Sensor j generates a random number ε_{ij} and decides whether participates in this measurement. This procedure can be demonstrated as

$$\Phi_{ij} = \sqrt{z} \begin{cases} +1 & \text{if } \varepsilon_{ij} < 1/2z \\ 0 & \text{if } \varepsilon_{ij} > 1 - 1/2z \\ -1 & \text{otherwise} \end{cases} \quad (7)$$

If $\Phi_{ij} = 0$, then sensor j is not involved in the current round of measurement. Furthermore, we improve the process of generating ε_{ij} and take sensor's residual power in to consider. The proposed scheme can be described as follow:

$$\varepsilon_{ij} = rand() + \frac{power_j(i) - power_{avg}(i)}{power_j(i)} \times \alpha \quad (8)$$

where α is influence factors of residual energy. The probability of this node participated in next measurement is determined by 3 factors: the base sample rate p_{base} , the residual energy of nodes $power_j(i)$ and the influence factor α .

Particularly, all nodes will send their health information to sink during the stage of network formation, sink calculates the average residual energy $power_{avg}(0)$ of the initial stage. At the beginning of CS, sink broadcast $power_{avg}(0)$ and p_{base} to all sensors, so sensors can calculate their initial sampling probability through the $power_{avg}$ in the first round. The process of sampling probability control can be described by Algorithm 2.

Algorithm 2: The algorithm of j th node's sampling probability control

Input: Random scheduling probability p_{base} and the average residual power $power_{avg}$;

Output: Sensory value u_j

1. $\varepsilon = rand() + \frac{power_{node}(i) - power_{avg}}{power_{node}(i)} \times \alpha$; /*Generate a random probability*/
 2. if $\varepsilon \leq p_{base}/2$ then
 3. $\Phi_{ij} = 1$;
 4. else if $\varepsilon \geq 1 - p_{base}/2$
 5. $\Phi_{ij} = -1$;
 6. get $x_j = (temp, ID, \Phi_{ij})$;
 7. $u_j = \Phi_{ij}x_j$ /*Calculate the measure value*/
 8. send_data(u_j); /*Send data to parent*/
 9. else
 10. sleep ();
 11. end if;
-

IV. SIMULATION EXPERIMENT

A. Experimental Data

To obtain the real environmental data sequence, this paper collected the surface temperature data of a piece of forest land in July 10th at 11 a.m. by deploying the Crossbow environment monitoring wireless sensor network within the Beijing University of Technology. The monitoring network is composed of $4 * 9 = 36$ Micaz sensing nodes, and the nodes are deployed in the area with the uniform spacing of $d = 5$ meters. The temperature of the collection is shown in Figure 2:

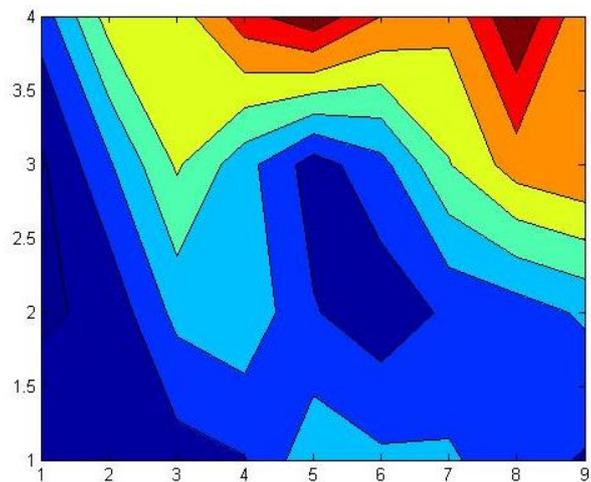


Figure 2. The map of the original temperature data

The experimental results are extended to 1-dimensional temperature data (Fig. 3) by the formula (2). Fig.4 show its corresponding discrete Fourier transform. One can show that signal is almost contained in $K = 3$ ($K \ll N$) Fourier coefficients. Therefore, the experiment will use the basis of Fourier transform basis as sparse basis ψ .

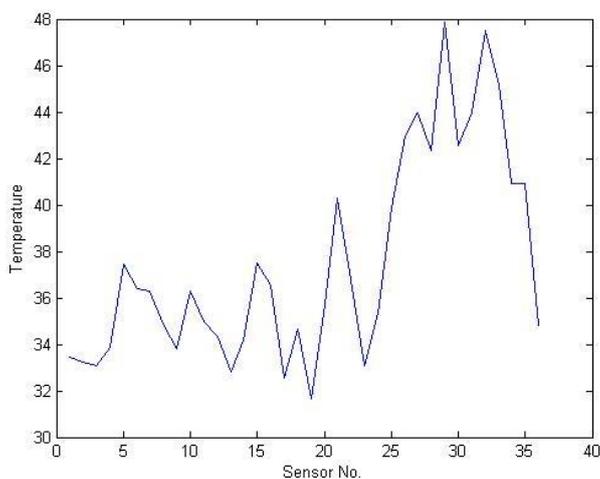


Figure 3. The original temperature data transformed as formula (2)

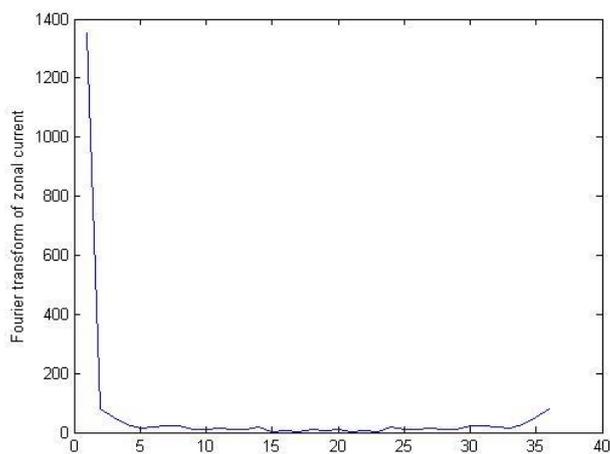


Figure 4. Amplitude of the Fourier transform of the field

B. Measurement matrix and Recovery algorithm

In the construction of the measurement matrix, sink need S to determine M and the nonzero element position of each line. As described in the reference [4], M needs to satisfy the requirement $M \geq k^2 \log^N$. In this experiment, we denote $M = 13$.

In addition, sink gets the sensors' ID which has involved in the process of CS to construct sub sequence as formula (2). If one node participates in the CS process, the corresponding position of sub sequence is nonzero. Therefore, sink can construct observation matrix based on these.

For signal reconstruction, this paper use CVX, a package for specifying and solving convex programs [9].

C. Experiment scheme

In this paper, we select the method of constructing a sparse random projection matrix [4], [5] as a comparison scheme 1 (S-CS), and a traditional clustering routing algorithm (Normal-C) which were proposed by [10].

In the CPCS scheme, in order to show the influence of the residual energy of the nodes on the sampling probability clearly let $\alpha = 1$ in formula (8). There is a problem that higher basic sampling rate may lead more data communication. In order to minimize the energy consumption of the network while maintaining the recovery quality of reconstruction, we choose the smallest value of $p_{base} = 0.05$.

First, Experiments will compare the differences between the S-CS and CPCS in the reconstruction error of the sparse sampling scheme. Then the energy balance performance of CPCS, S-CS and Normal-C are compared through the specific energy model.

D. Performance Analysis

In the experiment, the recovery performance can be calculated by $\epsilon_e = \|\hat{x} - x\|_2 / \|x\|_2$, where x is the actual data and \hat{x} is the recovered data. In the experiment of $1/z = p_{base}$ increases with interval 0.05. In order to reduce the accidental error impacted on the results, each experiment has conducted 100 times to get the average value of reconstruction error. As noted in the Figure 5, accurate reconstruction is possible for a range of values of p_{base} . The experimental results show that the proposed CPCS has no obvious difference compared to the S-CS scheme in the reconstruction error.

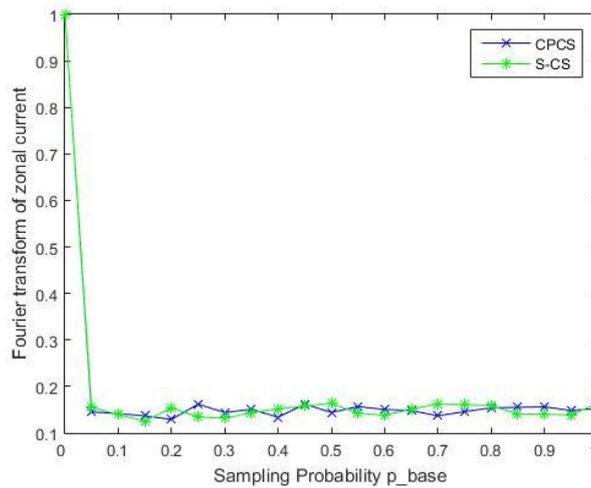


Figure 5. Comparison of recovery errors under different p_{base}

In order to compare the energy consumption of the 3 schemes, this paper selects the bit jump energy model [8], which is given as follows:

$$E_{Tx}(l, d) = E_{elec} \times l + \epsilon_{amp} \times l \times d^2$$

$$E_{Rx}(l) = E_{elec} \times l$$

In the experiment, $E_{elec} = 50nJ/bit$, $\epsilon_{amp} = 100pJ/bit/m^2$, the packet length is 1024bits, and the initial energy of each node is 1J. For simplify the process of the experiment, we neglect the energy consumption of node sampling and data processing in the data collecting. The times of CS measurement is 1.4×10^4 .

In the process of the experimental die node number is shown in Figure 6. Normal-C is affected by the high sampling probability of nodes and rapidly emerge "death" nodes. In contrast, the S-CS and CPCS scheme can effectively prolong the lifetime of the node, and residual energy of the whole network (Figure 7) are almost same. However, compared CPCS to the S-CS scheme, the lifetime of the whole network is longer.

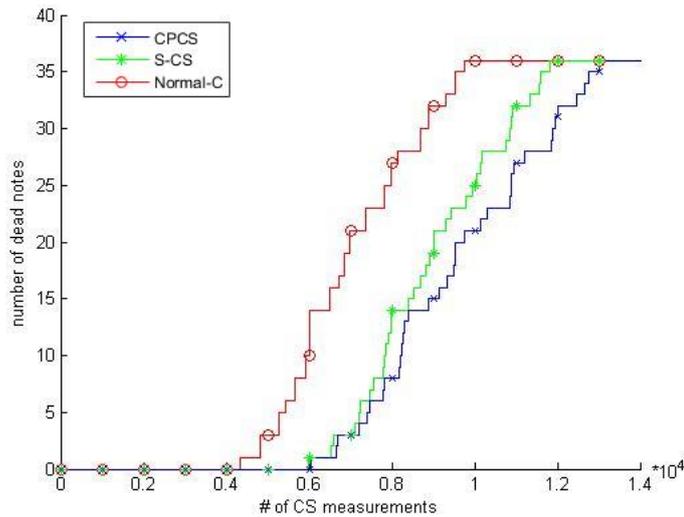


Figure 6. The sum number of sensors with no energy

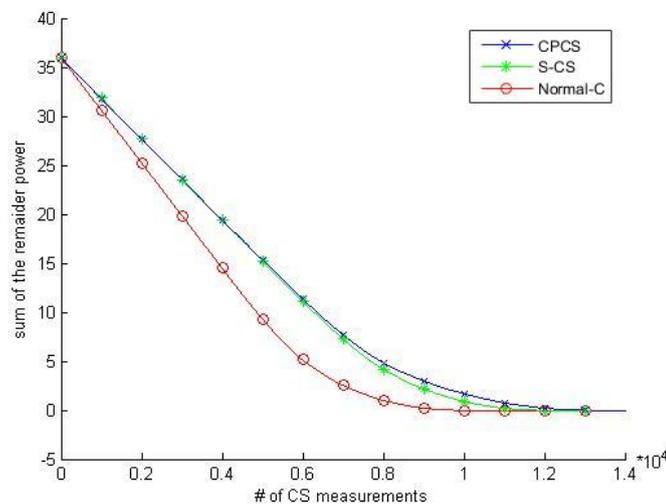


Figure 7. Residual energy of WSN

In summary, through simulation experiments, S-CS and CPCS compared with Normal-C observation scheme has obvious advantages in energy consumption, the wireless sensor network is more suitable for the limited resource. Although the speed of energy consumption is similar, benefited from function of energy-balanced design of CPCS, the time of first sensor running out its energy is delayed compared with S-CS. During the compression, lack of sensor involved in measurement will lead to higher recovery error. Therefore, the proposed CPCS scheme can better adapt to the compressed sensing data collection process of wireless sensor network.

V. CONCLUSION

To improve the lifetime of WSN, we proposed a scheme using Compressive Sensing. The experiments have proved that our scheme has positive effect on prolong lifetime. In addition, the CPCS method emphasizes on the adjustment of the measurement probability, and does not define how to transmit data to the sink, so it can cooperate with different network topology and data routing protocols.

REFERENCES

- [1] Luo C, Wu F, Sun J. Compressive data gathering for large-scale wireless sensor networks. presented at the Proceedings of the 15th annual international conference on Mobile computing and networking. Beijing, China, 2009.
- [2] Wang W, Garofalakis M and Ramchandran K. Distributed sparse random projections for refrmable approximation. In Information Processing in Semor Networks (IPSN 2007), 2007.

- [3] Lee S, Patten S, Sathiamoorthy M. Spatially-Localized Compressed Sensing and Routing in Multi-Hop Sensor Networks. Proceedings of the 3rd International Conference on GeoSensor Networks, Oxford, UK, 2009.
- [4] Quen G, Masiero R, Munaretto D, et al. On the interplay between routing and signal representation for compressive sensing in wireless sensor networks. In Information Theory and Applications Workshop(ITA'09), 2009, pp:206—215.
- [5] Tang L, Zhou Z, Shi L et.al. Energy Balance Based WSN Compressive Sensing Algorithm. Journal of Chinese Computer Systems. 2011, 33(8), pp:1919-1923.
- [6] Yao Y J, Cao Q and Vasilakos A V. EDAL: An energy-efficient, delay-aware, and lifetime-balancing data collection protocol for wireless sensor networks. Proc. IEEE 10th Int. Conf. Mobile Ad-Hoc Sensor Netw. 2013. vol:23(3), pp. 182-190,
- [7] WSN Training: Introduction to XMesh.2006. Crossbow Technology, Inc.
- [8] Heinzelman, W., Chandrakasan, A” and Balakrishnan, H. 2002. An application-specific protocol architecture for wireless microsensor networks. Wireless Communications, IEEE Transactions on, 1(4):660-670.
- [9] M. Grant and S. Boyd, “CVX: Matlab software for disciplined convex programming.” [Online]. Available: <http://cvxr.com/cvx/1>
- [10] Heinzelman, W.,Chandrakasan, A and Balakrishnan, H. An application-specific protocol architecture for wireless microsensor networks. Wireless Communications, IEEE Transactions on. 2002. 1(4):660-670.