An Efficient Data Gathering and Reconstruction Method in WSNs Based on Compressive Sensing

Wenjie Yan, Qiang Wang, *Member, IEEE*, Yi Shen, *Member, IEEE*, Yan Wang, Qitao Han

School of Astronautics, Harbin Institute of Technology, Harbin, China

Email: wenjieyanhitharbin@gmail.com, wangqiang@hit.edu.cn, shen@hit.edu.cn, wyabc@hit.edu.cn, hanqitao2010@163.com

Abstract—In this paper, we introduce a very simple deterministic measurement matrix design algorithm(SDMMDA), based on which the data gathering and reconstruction in wireless sensor networks(WSNs) are greatly enhanced. Although SDMMDA is very simple, but the measurement and reconstruction performance is more efficient than the random matrix and the matrix designed by schnass. The basic principle of the proposed algorithm can be stated as follows. First, generating a random redundant matrix Φ . Second, constructing a Gram matrix G , which can be denoted as $\Phi^T * \Phi$. Third, decreasing the absolute value of the off-line entries of the Gram matrix. Finally, mutual coherence of the random measurement matrix can be decreased greatly and the compressive data gathering as well as the signal reconstruction performance are greatly improved simultaneously. Besides that, we adopt backtracking-based adaptive OMP(BAOMP) method to reconstruct the original signal gathered by WSNs. By using BAOMP, We need not to know the signal sparse level K anymore. Extensive simulations and practical experiments of WSNs have shown that reconstruction performance of the compressive data gathered with CS method is improved greatly by using the proposed SDMMDA and BAOMP.

Index Terms—deterministic measurement matrix; mutual coherence; compressive data gathering; wireless sensor networks; BAOMP

I. INTRODUCTION

Traditional data gathering and processing method is to use the multi-hop route to transmit data from one sensor to another. Finally the data will be transmitted to the sink node according to the route. Disadvantage of traditional method lies in the unbalanced energy consumption for each sensor and some redundant data transmissions. The sensor closer to the sink will consume more energy than other sensors. To avoid the redundant data transmissions, some researches introduce data fusion methods to process data in WSNs. More completed routing protocols and much higher computation ability will be needed for each sensor. Sometimes data fusion methods cannot solve unbalanced energy consumption problems. A novel method named compressive sampling theory(CS)[1-3] has received more attentions at present. In this paper, we investigate compressive data gathering and original signal reconstruction in wireless sensor networks(WSNs). By adopting the Compressive Sampling theory, the energy consumption can be balanced and the redundant data transmissions can also be avoided. The data transmitted can be sparse in a certain domain and CS theory can make sure that a K -sparse signal can be reconstructed from a relative small number of measurements

M with a probability even up to one. The main contributions of this paper lie in the following three aspects:

First, we adopt the compressive sampling theory to gathering each sensors' monitoring value in WSNs, instead of using the traditional multi-hop routing transmission scheme for each sensor.

Second, instead of using random measurement matrix to measure each sensor's monitoring value and reconstruct the original signal, we propose an simple but more efficient deterministic measurement matrix design algorithm to achieve the better data gathering and original signal reconstruction performance in WSNs.

Third, extensive simulations and practical experiments of WSNs have shown that if we choose a proper number of measurements M , the sensors' monitoring values can be gathered efficiently as well as the energy consumption can be reduced greatly by using our proposed algorithm. Besides that, the original signal can be reconstructed successfully with a very high probability by using BAOMP.

The paper is organized as follows. In section 2 we recall briefly the related works in this research area. In section 3, we describe the proposed simple deterministic measurement matrix design algorithm in detail. In section 4, extensive simulations and practical experiments of the WSNs are presented. At last, section 5 concludes the paper.

II. RELATED WORKS

Network data processing has received more and more attentions recently. Traditional compression techniques utilize data correlation during the encoding process and the explicit data communication among sensors will be needed. Cristescu in [4] introduce a joint entropy coding algorithm to achieve the compressive data transmission. However, the approach by utilizing data correlation is only unidirectionally. The clustered aggregation(CAG) technique[5] is proposed by S. Yoon. By grouping sensors with similar readings, CAG only transmits on reading per group to achieve a predefined error threshold. The problem of the traditional compression methods relies heavily on the routing topology. Besides that, joint entropy coding techniques and the clustered aggregation method need complex computation and lots of data exchange.

Based on Slepian-Wolf coding theory[6], Distributed source coding techniques[7] want to reduce communication complexity at sensors and utilize data correlations at the sink node. This technique can not only eliminates data exchanges, but

also can decouple routing from compression. However, in order to allocate appropriate number of bits used by each sensor, the global correlation structure needs to be known which is a prerequisite of Slepian-Wolf coding theory. But the requirement is hard to be satisfied, especially in large scale wireless sensor networks.

With the emergence of compressive sampling theory[1-3], a novel in-network data compression method has emerged. Haupt et al. in [8] speculate the potential of using compressive sampling theory for data fusion in a multi-hop wireless sensor networks. Luo *et al.* in [9] adopt compressive sampling theory to sensor data gathering for large scale wireless sensor networks. The proposed compressive data gathering method can not only reduce global scale communication cost without introducing additional computation or complicated transmission control protocols, but also can balance the energy consumption for all the sensors in WSNs. Besides that, in order to better utilize compressive sampling theory to gather and reconstruct the sensing data, Luo also in [10] introduces the novel measurement matrix design method. Her method indeed increases the performance of data gathering and signal reconstruction performance. However, the basic principle of her random measurement matrix design method will result in that the signal reconstruction performance cannot be promised.

Follow Luo's idea on designing the measurement matrix and inspired by schnass[11] who introduces the sensing matrix design method to increase the reconstruction performance under CS framework, we introduce the proposed SDMMDA. By adopting SDMMDA to gathering the sensed data, the sink node and all the sensors do not need the extra data exchange for the random measurement entries, because the measurement entries can be considered as a priori knowledge of the whole WSNs. So the communication load is reduced a lot by using SDMMDA. Besides that, our proposed SDMMDA is based on reduction of mutual coherence of columns of measurement matrix, so the signal reconstruction performance increases greatly at the same time.

III. DESCRIPTION OF SIMPLE DETERMINISTIC MEASUREMENT MATRIX DESIGN ALGORITHM

In this section, we mainly want to introduce the inspiration of deterministic measurement matrix design and the way to design the proposed simple deterministic measurement matrix.

A. Inspiration of Deterministic Measurement Matrix Design

In order to explain work principle of CS, we should recall the definition of *mutual coherence*[12].

Definition 1: Assume there is a measurement matrix Φ , its mutual coherence is defined as the the largest absolute and normalized inner product between two different columns in Φ . Specifically, it can be denoted as follows:

$$\mu\{\Phi\} = \max_{1 \leq i, j \leq n \text{ and } i \neq j} \frac{|\phi_i^T \phi_j|}{\|\phi_i\| \cdot \|\phi_j\|} \quad (1)$$

The *mutual coherence* describes strongest similarity between the different columns of the measurement matrix Φ .

mutual coherence can reflect the weakness of the measurement matrix, because the two closely related columns may confuse any greedy pursuit algorithm and finally lead to the wrong reconstruction result. A different way to understand the *mutual coherence* is by considering the Gram matrix $G = \Phi^T * \Phi$, constructed using the measurement matrix Φ after normalizing each of its columns. The off-diagonal entries in G are the inner products that explained in (1). The *mutual coherence* is the off-diagonal entry in G with the largest absolute magnitude. If the following equation holds, several greedy algorithms like OMP[13] can reconstruct the original signal X exactly.

$$\|X\| < \frac{1}{2} \left(1 + \frac{1}{\mu\{\Phi\}}\right) \quad (2)$$

The equation (2) implies that if the measurement matrix Φ is designed to make the $\mu\{\Phi\}$ as small as possible, it allows a wider support set of the original signal to be reconstructed successfully by adopting several greedy pursuit algorithms. This is also the exact reason why we work on reduction of the *mutual coherence* of the measurement matrix and the inspiration of our proposed measurement matrix design.

Another way to measure the coherence between the different columns of the measurement matrix Φ named cumulative coherence.

Definition 2: For a measurement matrix Φ , its cumulative coherence is defined to measure the maximum total coherence between a fixed atom and a predefined collection of other atoms. Comparing with mutual coherence, we believe that the cumulative coherence should reflect the coherence between the columns of the measurement matrix more true and more precisely. We defined it as follows:

$$\mu_K\{\Phi\} = \max_i \max_{|\Lambda|=K, i \notin \Lambda} \sum_{j \in \Lambda} \frac{|\phi_i^T \phi_j|}{\|\phi_i\| \cdot \|\phi_j\|} \quad (3)$$

Due to the difficulty of computation of cumulative coherence than that of mutual coherence, we mainly focus on the reduction of mutual coherence in measurement matrix and then the simple deterministic measurement matrix design algorithm is proposed.

B. Design of SDMMDA

In this section, we introduce the proposed simple deterministic measurement matrix design algorithm. We assume that the initial signal X is a K -sparse signal, which means that the number of nonzero elements in X is no more than K . So for the equation $y = \Phi X$, we have to minimize $\mu\{\Phi\}$ or $\mu_K\{\Phi\}$ in order to obtain a better reconstruction performance for the greedy pursuit algorithms.

In fact, our algorithm inspired by author schnass in [11] to design a Gram type matrix $G = \Psi' * \Phi$ closest to the ideal Gram matrix. Exactly speaking, the ideal Gram matrix has only ones on the diagonal and all off diagonal entries of absolute value $\mu = \sqrt{(N-M)/M(N-1)}$. Based on above

analysis, our proposed simple deterministic measurement matrix design algorithm is described as follows:

Algorithm 1 Simple Deterministic Measurement Matrix Design Algorithm(SDMMDA)

Objective: minimize $\mu\{\Phi\}$

2: **Input:**

- An measurement matrix $\Phi \in R^{M \times N}$.
- Coherence fixed threshold- μ (welch bound).

Algorithm begin:

6: 1) **Normalization:** Normalize all the columns of the matrix Φ .

2) **Calculate the initial Gram type matrix:** $G = \Phi^T * \Phi$

8: 3) **Set threshold:** we can adopt μ as a fixed threshold in the proposed algorithm.

4) **Obtain the proposed Gram type matrix G by**

10: if $i=j$

$$G(i, j) = 1$$

else

$$G(i, j) = \begin{cases} G(i, j) & |G(i, j)| \leq \mu \\ \mu \cdot \text{sign}(G(i, j)) & |G(i, j)| > \mu \end{cases}$$

12: 5) **Obtain new Φ by solving the formulation:**

$$\Phi^T * \Phi = G.$$

Output: Expected deterministic measurement matrix Φ with low mutual coherence among its columns.

From the above description of the proposed algorithm, we know that the proposed algorithm is very simple and can be implemented easily. Because it is not a iterative algorithm, the proposed algorithm runs very fast. What is more important, SDMMDA is much more application oriented than other deterministic measurement matrix design algorithms.

In order to better understanding the proposed algorithm, we investigate the mutual coherence and cumulative coherence of Gaussian matrix with different measurement matrix design algorithms. For the simulation of mutual coherence, we set the number of the row of measurement matrix M equals to 128 and the number of the column of measurement matrix N ranges from 128 to 512. From Fig.1(a), the mutual coherence of measurement matrix after using the proposed SDMMDA is more closer to the optimal Grassmannian frames, and much lower than the random Gaussian matrix and the sensing matrix adopting schnass' method. For simulation of cumulative coherence, we set the number of the row of measurement matrix M equals to 128, and the column of the measurement matrix N equals to 256. Besides that, we set the cumulative parameter which is named the signal sparse level K ranges from 1 to 20. From Fig.1(b), the cumulative coherence of the measurement matrix after using the SDMMDA is the smallest among all the measurement matrixes. The following extensive simulation results are quite consensus with the coherence analysis above.

Fig. 1. coherence comparison using Gaussian matrix among different measurement matrix design algorithms

IV. EXPERIMENTS AND CORRESPONDING RESULTS ANALYSIS

In this section, we mainly focus on the compressive signal gathered by WSNs reconstruction by using the different measurement matrixes under the CS framework. We have done three kinds of the experiments to verify original signal reconstruction performance of the proposed SDMMDA.

A. Signal Reconstruction Performance of Different Measurement Matrix design methods by Using OMP

In this part, we investigate the signal reconstruction performance of the different measurement matrix design methods by using orthogonal matching pursuit algorithm. Exactly speaking, we have done two kinds of the experiments. In the first experiment, we use measurement matrix with the size of $M = 128$ and $N = 256$, and the signal sparse level K ranges from 1 to 60 with zero-one signal and 1 to 70 for Gaussian signal. From Fig.2 (a) and (b), we know that the reconstruction performance of matrix using the proposed SDMMDA is the best among all the measurement matrix design methods for both Gaussian signal and zero-one signal when the signal sparse level K varies. Obviously, signal reconstruction performance of matrix via schnass' method is better than that of random Gaussian matrix. In the second experiment, we fix the number of the matrix columns $N = 256$ and the signal

(a) signal reconstruction performance of zero-one signal

(b) signal reconstruction performance of Gaussian signal

Fig. 2. signal reconstruction performance versus signal sparse level K with different measurement matrix design algorithms

(a) signal reconstruction performance of zero-one signal

(b) signal reconstruction performance of Gaussian signal

Fig. 3. signal reconstruction performance versus number of measurements M with different measurement matrix design algorithms

sparse level $K = 30$. The number of measurements M varies from 100 to 200 for zero-one signal and from 60 to 160 for Gaussian signal. From Fig.3 (a) and (b) we can clearly see that the reconstruction performance of measurement matrix via SDMMDA is the best of all the measurement matrixes for both zero-one signal and Gaussian signal. Of course, the reconstruction performance of matrix via schnass' method is better than the random Gaussian measurement matrix.

B. Signal Reconstruction Performance of Different Measurement Matrix design methods by Using BAOMP

In this part, we mainly investigate the Signal Reconstruction Performance of Different Measurement Matrix design methods by Using BAOMP[14]. BAOMP incorporates a simple backtracking technique based on the OMP algorithm. BAOMP algorithm does not need signal sparse level K to be known as a priori and its reconstruction performance is much better than several existing reconstruction methods. So it is quite suitable for the signal reconstruction in wireless sensor networks. Besides that, we also have done two kinds of the experiments to verify the signal reconstruction performance of the proposed SDMMDA. In the first experiment, we investigate the signal reconstruction performance with different measurement matrix design methods as the signal sparse level K varies. From Fig.4 (a) and (b) we can observe that signal reconstruction performance of proposed SDMMDA is the best among all the measurement matrix design methods for both zero-one signal and Gaussian signal. In the second experiment, we mainly focus on the signal reconstruction performance by using different measurement matrix design methods as the number of measurements M varies. From Fig.5 (a) and (b) we can also definitely know that matrix with the proposed SDMMDA is of the best signal reconstruction performance among all the measurement matrix design methods.

Besides that, It should also be noted that the signal reconstruction performance of matrix via schnass' method is worse than that of the random Gaussian matrix. This indicates that the sensing matrix via the schnass' matrix design method is not fit for BAOMP, which can also be clearly observed from the analysis of mutual coherence and cumulative coherence in the former part.

C. Practical Signal Reconstruction Performance of SDMMDA by Using BAOMP in WSNs

In this part, we mainly investigate the practical signal gathering and reconstruction performance of SDMMDA by using BAOMP in wireless sensor networks. The benefits of BAOMP and the proposed SDMMDA have been stated clearly. We combine them to the practical compressive data gathering and reconstruction in WSNs. First we adopt 50 sensors to constitute a wireless sensor networks. The network topology are chain topology and tree topology respectively. We use them to detect the sparse temperature events in the proposed wireless sensor networks. As Fig.6 (a) and (b) shows, there are 8 temperature events in the two kinds of WSNs respectively, and the number besides sensor is the sensor's logical address and

(a) signal reconstruction performance of zero-one signal

(b) signal reconstruction performance of Gaussian signal

Fig. 4. signal reconstruction performance versus signal sparse level K with different measurement matrix design algorithms

(a) signal reconstruction performance of zero-one signal

(b) signal reconstruction performance of Gaussian signal

Fig. 5. signal reconstruction performance versus number of measurements M with different measurement matrix design algorithms

(a) Eight Temperature Monitoring Events Under Chain Topology

(b) Eight Temperature Monitoring Events Under Tree Topology

Fig. 6. Eight Temperature Monitoring Events of $WSNs$ Under Different Network Topologies

the symbol of the temperature events. To be easy to transmit data for the sensors, we set a temperature threshold for each sensor. When temperature the sensor monitoring is higher than the threshold, the temperature data is considered as logical one, or the temperature data is considered as logical zero. In order to avoid the negative effects of the packets loss to the compressive data gathering process, we use acknowledgement mechanism to promise each measurement data of each sensor can reach the sink node successfully. The data frame and acknowledgement frame are shown in Fig.7. In detail, the exact data of each measurement we want is shown in the SComAssistant in Fig.6. From SComAssistant of Fig.6, we can clearly see that the 8 temperature events data weighted sum has been transmitted to the sink node successfully.

Finally, we test signal reconstruction performance of the proposed algorithm with the data gathered by WSNs in this experiment. From Fig.8 (a) and (b) we can observe clearly

Fig. 7. data ack mechanism of WSNs

(a) reconstruction performance versus the signal sparse level K

(b) reconstruction performance versus the number of measurements M

Fig. 8. Reconstruction performance of zero-one signal by using the proposed algorithm with Gaussian measurement matrix

that whether we change the signal sparse level K or the number of measurements M, matrix via proposed SDMMDA has a better reconstruction performance than that of random Gaussian matrix under the same parameter settings.

V. CONCLUSIONS

In this paper, we introduce a very simple deterministic measurement matrix design algorithm(SDMMDA) to gathering compressive data in WSNs under the CS framework. Through analyzing mutual coherence and cumulative coherence of different measurement matrix, we find that the coherence of the measurement matrix via proposed SDMMDA is much

smaller than other measurement matrixes. Besides that, we adopt BAOMP algorithm without knowing the signal sparse level K to adaptively reconstruct the original signal gathered by WSNs and obtain a much better reconstruction performance. Extensive simulations and practical experiments of WSNs have shown that the signal reconstruction performance using SDMMDA is much better than those using the random Gaussian measurement matrix and matrix via schnass' method. In the future research work, we will investigate more practical and more challenging scenarios. Under such condition, we will combine the measurement matrix design with routing protocols, network topology and sparse representation of the original signal together.

ACKNOWLEDGMENT

This work was supported by the National Natural Science Foundation of China (Grant No. 61174016 and No. 61171197)

REFERENCES

- [1] D. Donoho, M. Elad, and V. Temlyakov, *Stable recovery of sparse overcomplete representations in the presence of noise*. *IEEE Trans. Inf. Theory*, vol. 52, pp. 6-18, 2006.
- [2] D. Donoho, *Compressed sensing*. *IEEE Trans. Inf. Theory*, vol. 52, pp. 1289-1306, 2006.
- [3] E. Candes, J. Romberg, and T. Tao, *Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information*. *IEEE Trans. Inf. Theory*, vol. 52, pp. 489-509, 2006.
- [4] R. Cristescu, B. Beferull-Lozano, M. Vetterli et al, *Network correlated data gathering with explicit communication: Np-completeness and algorithms*. *IEEE/ACM Trans. on Networking*, vol.14, no.1, pp.41-54, 2006.
- [5] S. Yoon and C. Shahabi, *The clustered aggregation (cag) technique leveraging spatial and temporal correlations in wireless sensor networks*. *ACM Trans. on Sensor Networks*, vol.3, no.1, pp.xxx, 2007.
- [6] D. Slepian and J. K. Wolf, *Noiseless encoding of correlated information sources*. *IEEE Trans. Inf. Theory*, vol.19, no.4, pp.471-480, 1973.
- [7] G. Hua and C. W. Chen, *Correlated data gathering in wireless sensor networks based on distributed source coding*. *Intl. Journal of Sensor Networks*, vol.4, no.1/2, pp.13-22, 2008.
- [8] J. Haupt, W. U. Bajwa, M. Rabbat, et al, *Compressed Sensing for Networked Data*. *IEEE Signal Processing Mag.*, vol. 25, no. 2, pp. 92-101, Mar. 2008.
- [9] Chong Luo, Feng Wu, Jun Sun, et al, *Compressive Data Gathering for Large-Scale Wireless Sensor Networks*. *MobiCom'09*, Beijing, China, pp. 145-156, 2009.
- [10] Chong Luo, Feng Wu, Jun Sun, et al, *Efficient Measurement Generation and Pervasive Sparsity for Compressive Data Gathering*. *IEEE Trans. Wireless Commun.*, vol.9, no.12, pp. 3728-3738, Dec. 2010.
- [11] K. Schnass and P. Vandergheynst, *Dictionary preconditioning for greedy algorithms*. *IEEE Trans. Signal Process.*, Vol.56, no.5, pp. 1994-2002, May. 2008.
- [12] D. L. Donoho and M. Elad, *Optimally sparse representation in general(non-orthogonal) dictionaries via l^1 minimization*. *Proc. Nat. Acad. Sci.*, Vol.100, pp. 2197-2202, 2002.
- [13] J.A. Troop and A.C. Gibert, *Signal recovery methods from random measurements via orthogonal matching pursuit*. *IEEE Trans. Inf. Theory*, Vol.53, no.12, pp. 4655-4666, 2007.
- [14] H. Huang and A. Makur, *Backtracking-Based Matching Pursuit Method for Sparse Signal Reconstruction*. *IEEE signal processing letters*, Vol.18, no.7, pp. 391-394, 2011.