Efficient Measurement Method for Spatiotemporal Compressive Data Gathering in Wireless Sensor Networks

Xiao Xue[1], Song Xiao[2] and Lei Quan[3]
[1] State Key Laboratory of Integrated Services Networks, Xidian University
Xi'an, China
[e-mail: xxue@stu.xidian.edu.cn]
[2] State Key Laboratory of Integrated Services Networks, Xidian University
Xi'an, China
[e-mail: xiaosong@mail.xidian.edu.cn]
[3] School of Aerospace Science and Technology,
Xidian University, Xi'an China
[e-mail: gloomy2110@hotmail.com]
*Corresponding author: Song Xiao

Abstract

By means of compressive sensing (CS) technique, this paper considers the collection of sensor data with spatiotemporal correlations in wireless sensor networks (WSNs). In energy-constrained WSNs, one-dimensional CS methods need a lot of data transmissions since they are less applicable in fully exploiting the spatiotemporal correlations, while the Kronecker CS (KCS) methods suffer performance degradations when the signal dimension increases. In this paper, an appropriate sensing matrix as well as an efficient sensing method is proposed to further reduce the data transmissions without the loss of the recovery performance. Different matrices for the temporal signal of each sensor node are separately designed. The corresponding energy-efficient data gathering method is presented, which only transmitting a subset of sensor readings to recover data of the entire WSN. Theoretical analysis indicates that the sensing structure could have the relatively small mutual coherence according to the selection of matrix. Compared with the existing spatiotemporal CS (CS-ST) method, the simulation results show that the proposed efficient measurement method could reduce data transmissions by about 25% with the similar recovery performance. In addition, compared with the conventional KCS method, for 95% successful recovery, the proposed sensing structure could improve the recovery performance by about 20%.

Keywords: Spatiotemporal correlations, Data gathering, Compressive sensing, Wireless sensor networks, Kronecker product.

1. Introduction

Wireless sensor networks(WSNs) have been considered as a powerful tool for environment, health and military monitoring [1][2]. A general WSN consists of one sink node(or base station) and a large number of sensor nodes deployed over a target area of interest and equipped with self-organizing wireless devices that could sense a wide variety of information, including humidity, illumination, temperature and so on [1]. Since the sensor nodes are often battery-powered, the energy of them is highly limited and it is impractical to charge or replace all the batteries simultaneously, which lead to extreme energy constraints for many WSNs platforms [3][4]. To reduce the energy consumption, traditional approaches reduced the data transmissions of the entire network by in-network data compression, including entropy coding or transform coding. However, the coding procedure brings the computation and control overheads to the senor nodes.

Recently, the theory of Compressive sensing (CS) [5][6][7][8] states that a compressible signal can be recovered through highly incomplete information which is obtained by linear measurements. And since the preliminary works for data collection in [9][10], a number of works like [11-20] have indicated that the sampling and transmission costs can be reduced through CS linearly combining the content of packets as they are routed to the sink node. We found that the classical Compressive Data Gathering (CDG) [11][12] requires many sensor nodes to participate in generating each measurement and the sparse measurement matrix in [14] needs more measurements, so the transmission cost of the network remains relatively high. To overcome the problem, a sparsest random scheduling for CDG is presented in [17]. We noticed that the schemes above only considered the spatial correlations. In practical WSNs with densely deployed nodes, data over a certain consecutive time encompass both spatial and temporal (spatiotemporal) correlations [4]. CS could also be applied to exploit the temporal correlations to enhance the precision of data gathering for networks without stringent delay constraint.

Several works have been proposed using different methods to exploit data correlations both temporally and spatially. For example, in [21] the authors mainly made use of diffusion wavelets to characterize the spatiotemporal correlations well, and investigates minimum-energy problem of NP-completeness. As a measurement vector is obtained by solving an optimization problem, the computational and communication overhead of the adaptive algorithm proposed in [22] is relative high. For the online recovery of large data in WSN, G. Quer *et al.* [23] is focused on Principal Component Analysis (PCA) to capture the spatiotemporal characteristics of real signals. Afterwards, a sequential CS (Seq-Pro-CS) scheme [24] utilizing sliding window is proposed to reconstruct the data with the spatiotemporal correlations. The paper [25] introduced low-rank constraints to improve the recovery accuracy of signal. And the recovery performance depends on the historical matrix. The papers [24][25] above mainly focused on treating the data as continuous-time streams. For a mobile sink node, the authors [26][26] have provided an energy-efficient data storage model to reduce the number of transmissions by exploiting the spatiotemporal correlations. This model introduces a post-processing procedure to optimize the measurement collection, which brings additional calculations. The spatiotemporal (CS-ST) method in [28] conveyed the temporal measurements of each sensor node across the links without any aggregation at intermediate nodes. Compared with the method of the spatial (or temporal) dependence in [28],

the CS-ST method reduces the data traffic. However, the CS-ST method requires every node to send measurements to the sink node, and the in-network transmissions are still high.

Actually, the optimization procedure of the measurement collection could be eliminated by carefully designing the structure of the sensing matrix. Moreover, with a well-designed matrix structure, the in- network transmissions could be reduced. In this paper, by designing the sensing matrix carefully, an efficient data gathering method is proposed. In the proposed method, only a subset of all sensor nodes needs to send measurements during data collection. Therefore, the in-network transmissions could be reduced significantly. Moreover, the proposed method does not introduce any pre-processing or post-processing procedure, and therefore does not require extra computational costs. In addition, the proposed method does not exhibit performance degradation as that in the conventional Kronecker CS (KCS) [29] [30][31] data gathering method.

Specifically, the proposed efficient measurement method for data gathering mainly has two phases. During the first phase, the temporal measurements are generated independently based on both different CS measurement matrices and data of different sensor nodes over several consecutive time slots. Then in the second phase, a part of temporal linear measurements is sent to the sink node in the shortest path. Data of all sensor nodes could be recovered by CS theory. By different measurement matrices operating on each portion of the multidimensional signal, the method formed a new measurement structure. According to the properties of the Kronecker product, we proved that the proposed measurement structure can be unified into an equivalent sensing matrix which can be solved by the existing minimization program of one-dimensional (1-D) vector.

This paper has three main contributions:

· An efficient measurement method for gathering data over a certain time is designed, in which a new structure is formed in order to improve the recovery performance of conventional KCS and reduce data transmissions of the CS-ST method.

· A theoretical analysis is carried out and indicates that the mutual coherence of the improved measurement structure and sparsifying basis is relatively small according to the selection of the measurement matrix, which ensures the sensor data could be reconstructed with high probability from a small number of CS measurements.

· The validation of our efficient measurement method through simulations with different network sizes. Compared with the conventional CS-ST method, the simulation results show that the proposed efficient measurement method could reduce data transmissions and achieve the same data reconstructed performance. In addition, the method outperforms general KCS method in reconstruction quality.

The rest of the paper is organized as follows. Section 2 recapitulates CS background and system model in WSN. Section 3 proposes the spatiotemporal data gathering method and studies the recovery performance of the improved KCS structure. Simulation results are given in Section 4 and Section 5 provides conclusion and future work.

2. Compressive Sensing Background and System model for WSN

In this section, we first briefly introduce the basic fundamentals of Compressive Sensing theory, and then describe the system model for WSN.

2.1 Fundamentals of Compressive Sensing

According to CS theory, K-sparse signal $X \in R^N$ can be faithfully recovered from $M(M < N)$ linear measurements (projections) $y \in R^M$ by l_1 minimization

$$y = \Phi X = \Phi \Psi f \tag{1}$$

where Φ represents a $M \times N$ ($M < N$)measurement matrix, and Ψ is transformation basis of size $N \times N$. CS technology seeks to successfully reconstruct the signal with as few measurements as possible. To ensure exact reconstruction, one conclusion is presented in [32], which indicates that the minimal number M relies on the mutual coherence μ of an orthogonal matrix $N \times N$ Φ_0 and another orthogonal matrix $N \times N$ Ψ_0

$$M \geq CK\mu^2(\Phi_0, \Psi_0)\log(N/\delta) \tag{2}$$

where C is a small positive constant, μ is defined as the following (3):

$$\mu(\Phi_0, \Psi_0) = \max_{1 \leq k, j \leq N} \left| \left\langle \phi_k, \psi_j \right\rangle \right| \qquad \Phi_0^* \Phi_0 = NI, \Psi_0^* \Psi_0 = NI, \quad 1 \leq \mu \leq \sqrt{N} \tag{3}$$

where the rows of Φ_0 are the ϕ_k, $k = 1, 2 \cdots N$ and the columns of Ψ_0 are the ψ_j, $j = 1, 2 \cdots N$, Φ_0^* is the conjugate transpose matrix of Φ_0, $I \in R^{N \times N}$ is the identity matrix.

2.2 System Model for WSN

We consider that N nodes in a WSN are randomly distributed over a unit square area of size $A = 1 \times 1$. In order to guarantee the connectivity of the entire network, all nodes are assumed to have identical transmission range of $r^2 > A \ln(N)/\pi N$ and any two nodes could communicate with each other if their Euclidian distance is less than r [26][33]. Sensor data of the whole network over a certain time are assumed to exhibit both temporal and spatial correlations.

Let $X \in R^{T \times N}$ denote the sensor readings of a WSN with N nodes over T time slots, and the (t,i)-th entry $X_{t,i}$, $1 \leq t \leq T, 1 \leq i \leq N$ represents the sensor reading of the $i-th$ node at the $t-th$ time slot. Meanwhile, the column vector $\overline{X_{:,i}} \in R^{T \times 1}$ denotes the readings of the senor node i across T time slots, and the row vector $\overline{X_{t,:}} \in R^{1 \times N}$ denotes the readings of all nodes at the time slot t. Taking advantage of the spatiotemporal correlations among the data block $X \in R^{T \times N}$, the design goal is to achieve the faithful recovery of data $X \in R^{T \times N}$ through randomly gathering a subset of the temporal measurements of sensor readings.

3. Efficient Measurement Method for Data Gathering

In this section, we describe in detail how a separable sensing operator can be applied to the collection of data with the spatiotemporal correlations in WSNs, which formed the proposed efficient measurement method for data gathering as shown in **Fig. 1**. It begins with the CS measurements of the selected sensor readings in temporal domain and then transmits the measurements to the sink node. Based on above data transmitting process, an improved KCS structure is formed for sensor readings recovery at the sink node. To illustrate the data recovery performance of the proposed new structure, its mutual coherence is analysed in theory.
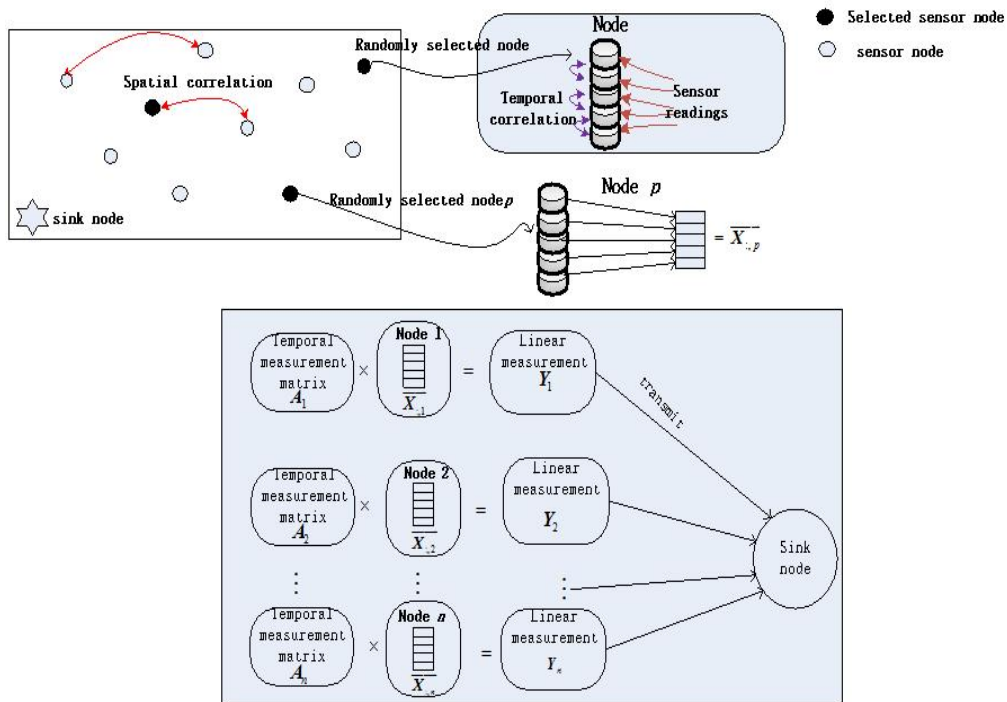

Fig. 1. The framework of the efficient measurement method for data gathering in a WSN

3.1 Spatiotemporal Data Gathering Method through Efficient Measurement

Since each sensor is supposed to be a tiny battery powered sensing unit with a finite amount of energy, we should reduce the number of data transmitting as much as possible to save the energy consumption. Note that the sensor readings generally exhibit both temporal and spatial correlations in the real applications of densely deployed WSNs [24][34]. To reduce the data transmissions in the network, a separable CS operator could be used to transmit the temporal measurements of partial sensor nodes by exploiting the spatiotemporal correlations. Next we will describe the operation of each stage in detail. The communication in the whole network is assumed to be synchronized and slotted.

At the beginning of the efficient measurement method, i.e., the initialization stage, given a compressible sensor reading block $X \in R^{T \times N}$, each sensor node is equipped with a temporal measurement matrix $A_i \in R^{m \times T}$ $i = 1, 2 \cdots N$ .To effectively capture the information of the

temporal data, the structures of A_i are designed to be different. During each sensing period T, n ($n < N$) different sensor nodes are selected randomly and uniformly from all N sensor nodes.

Based on the nature of the linear combinations in CS, it is noticed that the larger the number of the sensor nodes involved in the generation of one CS measurement, the higher the transmission cost. We choose a subset of nodes in order to reduce the number of nodes participating in the generation of one measurement, thus reducing the transmission cost of the network. The destination of uniformly selecting nodes at random is to capture enough spatial information and keep each sensor node participating evenly. The subset that formed by n selected nodes will change when the period T changes. Due to the random selection, the probability of each node being selected is equal, denoted as p_T. For the network after several periods T_T, from the perspective of probability, the average number of times that a node is selected is $T_T p_T$. (The relevant instruction is provided in appendix A.) When the sensor nodes selection is completed, the identification numbers (IDs) of the selected nodes are denoted by a set $NS = \{sl_1, sl_2, \cdots sl_n\}$, $sl_p \in \{1, 2 \cdots N\}$, $p = 1, 2 \cdots n$.

Then, the data forwarding stage is started, when a sensor node receives the instruction to collect data, this node obtains $m < T$ linear measurements according to the following expression:

$$Y_p = A_{sl_p} \overline{X_{:,sl_p}} \quad sl_p \in NS, Y_p \in R^{m \times 1}, \text{ where } \overline{X_{:,sl_p}} \in R^{T \times 1} \text{ is the original reading of the sensor node}$$

sl_p.

According to the CS theory, each sensor node compresses its T - length signal over T time slots into m linear measurements. So instead of transmitting original T - length signal, m measurements could be transmitted to the sink node ($m < T$) which could reduce transmission cost. Each node $sl_p \in NS$ ($p = 1, 2 \cdots n$) could form its initial transmission packet. Specifically, the packet of the $sl_p - th$ node, denoted by the structure PA_{sl_p}, has two independent components. The first component in PA_{sl_p} is the ID index sl_i of the node: $PA_{sl_p}.c1 = [sl_p]$, and the second one is data $\overline{X_{:,sl_p}}$ multiplied by the coefficient matrix: $PA_{sl_p}.c2 = [A_{sl_p} \overline{X_{:,sl_p}} = Y_p]$. These packets are transmitted to the sink node through conventional routing protocol (e.g. the shortest path).

Finally, when the multi-hop forwarding process of all the packets is finished, we can extract the corresponding information of different components to form the CS decoding information at the sink node. The sink node is responsible for the sensor data recovery as shown in **Algorithm 1** in section 3.2. Considering that CS encoding procedure takes advantage of the spatial-temporal correlations, the original reading data $X \in R^{T \times N}$ could be recovered with high precision through less data transmissions.

According to the gathering method above, we will elaborate on the formation of CS measurement matrix which is important in decoding algorithm. To reduce the transmitting overhead, data packets do not need transmitting the temporal matrices A_i. The sink node could obtain A_i through adopting the following method [35]: each node generates its temporal measurement matrix A_i using the random seeds instructions from the sink node, then according to the random seeds, the sink could generate all temporal measurement matrix A_i. On the other hand, the sink node could obtain the linear measurements

$\overline{Y} = \begin{bmatrix} Y_1^* & Y_2^* & Y_3^* & \cdots & Y_n^* \end{bmatrix}^* \in R^{nm}$ by the second components of the packets of n nodes as the following representation:

$$\overline{Y} = \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{pmatrix} = \begin{pmatrix} \beta_{11} A_1 \overline{X_{:,1}} + \beta_{12} A_2 \overline{X_{:,2}} + \cdots + \beta_{1N} A_N \overline{X_{:,N}} \\ \beta_{21} A_1 \overline{X_{:,1}} + \beta_{22} A_2 \overline{X_{:,2}} + \cdots + \beta_{2N} A_N \overline{X_{:,N}} \\ \vdots \\ \beta_{n1} A_1 \overline{X_{:,1}} + \beta_{n2} A_2 \overline{X_{:,2}} + \cdots + \beta_{nN} A_N \overline{X_{:,N}} \end{pmatrix} \tag{4}$$

Where Y^* denotes the conjugate transpose matrix of Y.

More concretely, the expression (4) above could be taken as:

$$\overline{Y} = \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{pmatrix} = \left(\begin{pmatrix} \beta_{11} & \beta_{12} & \cdots & \beta_{1N} \\ \beta_{21} & \beta_{22} & \cdots & \beta_{2N} \\ \vdots & \vdots & \vdots & \vdots \\ \beta_{n1} & \beta_{n2} & \cdots & \beta_{nN} \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} \right) \left(\begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & A_N \end{pmatrix} \begin{pmatrix} \overline{X_{:,1}} \\ \overline{X_{:,2}} \\ \vdots \\ \overline{X_{:,N}} \end{pmatrix} \right)$$

$$= (\beta \otimes I) A_{\text{BLOK}} x \tag{5}$$

Where $I \in R^{m \times m}$ is identity matrix, $A_i \in R^{m \times T}, i = 1, 2 \cdots N$,

$A_{\text{BLOK}} = diag[A_1 \quad A_2 \quad \cdots \quad A_N] \in R^{mN \times TN}$ is a block-diagonal temporal measurement matrix, and

$$x = vec(X) := \begin{bmatrix} \overline{X_{:,1}} \\ \overline{X_{:,2}} \\ \vdots \\ \overline{X_{:,N}} \end{bmatrix} \in R^{TN}$$

So based on the property of Kronecker product, we formed this new structure $(\beta \otimes I) A_{\text{BLOK}}$ in (5) to efficiently measure the spatiotemporal sensory data. The structure of A_{BLOK} means that different measurement matrices for the temporal signal of each sensor node are separately designed, which is different from the structure of KCS that based on the same measurement matrix operating on each node. It is noticed that the spatial measurement matrix $\beta \in R^{n \times N}$ is formulated by the transmission protocol, denoted by

$$\beta = \begin{pmatrix} \beta_{11} & \beta_{12} & \cdots & \beta_{1N} \\ \beta_{21} & \beta_{22} & \cdots & \beta_{2N} \\ \vdots & \vdots & \vdots & \vdots \\ \beta_{n1} & \beta_{n2} & \cdots & \beta_{nN} \end{pmatrix} \tag{6}$$

which is crucial to the data gathering method based CS in WSN.

According to the operation of the first stage in the proposed method, the sparse sampling(routing) matrix $\beta \in R^{n \times N}$ has only one element equal to 1 per row and at most one

element equal to 1 per column, and all the other elements are equal to zero. Thus, the elements in $\overline{Y}$ are a random subset of those in $\{A_i \overline{X_{:,i}}, i = 1, 2 \cdots N\}$. In other words, if the subset is denoted by $NS, |NS| = n$, the measurements could be obtained from the second component $\{PA_{sl_p}.c2| sl_p \in NS\}$. That indicated the location of non-zero entries of β is determined by the sensor nodes which are selected to sample measurement. In particular, if β is $N \times N$ identity matrix, $(\beta \otimes I)A_{\text{BLOK}}$ is degraded into the structure of CS-ST [28] that requires transmitting the temporal measurements of all nodes.

3.2 Formation of the Equivalent Sensing Matrix by Improved Kronecker Structure

Assume that the sensor reading $X \in R^{T \times N}$ is sparse and the corresponding sparsifying basis is two-dimensional (2-D) separable [29]. Thus, X can be written as: $X = \Psi_x \chi (\Psi_y)^*$. where $\Psi_x \in R^{T \times T}$ and $\Psi_y \in R^{N \times N}$ are two sparsifying basis corresponding to temporal domain and spatial domain, respectively, the matrix $\chi \in R^{N \times N}$ is a sparse matrix with sparsity S. Depending on a standard property of the Kronecker product, we have

$$x = vec(X) = vec\left(\Psi_x \chi (\Psi_y)^*\right) = (\Psi_y \otimes \Psi_x) vec(\chi) = \Psi^{\text{se}} f \tag{7}$$

Where $\Psi^{\text{se}} = \Psi_y \otimes \Psi_x$ is a sparsifying basis that is 2-D separable, and $vec(\chi) = f$. Substituting (7) into (5), we can obtain the equivalent sensing matrix Q:

$$\overline{Y} = (\beta \otimes I)A_{\text{BLOK}} (\Psi_y \otimes \Psi_x) vec(\chi) = \Phi \Psi^{\text{se}} f = Qf \tag{8}$$

where $Q = \Phi \Psi^{\text{se}}$, $\Phi = (\beta \otimes I)A_{\text{BLOK}}$ is a spatiotemporal measurement operator. By means of the CS technology, the sink node could recovery the coefficient f according to the expression (8) and then recovery the entire data X as shown in **Algorithm 1**.

Algorithm 1: Sensor data recovery at the sink node

Input: n nodes are randomly selected as the spatial sampling nodes, different temporal measurement matrices $A_i, i = 1, 2, \cdots N$, received linear measurements (Y_p) of sensor data over temporal domain, sparsifying basis Ψ^{se} according to (7), $\Psi^{\text{se}} = \Psi_{dct} \otimes \Psi_{dct}$;

Output: recovery sensory data $\overline{X}$, the mean square error(MSE);

1 $\beta \leftarrow 0$;   /* *Initializing the spatial measurement matrix*/

2 $NS \leftarrow \{sl_p | if \ Y_p \ is \ received\}$ /* *Record the index of received linear measurements of sensor data* */

$sl_p \in \{1, 2 \cdots N\}, p = 1, 2 \cdots n$;

3 for $p = 1$ to $p = n$ do

     for $j = 1$ to $j = N$ do

         if $j = sl_p$, $\beta(p, j) \leftarrow 1$;   /* *Assign spatial measurement matrix*/

```
      else
         continue;
      end if;
   end for;
 end for;
```

4 $\boldsymbol{\Phi} \leftarrow (\boldsymbol{\beta} \otimes \boldsymbol{I})A_{\mathbf{BLOK}}$ according to formula (5);   /* *Generate spatial-temporal measurement matrix*/

5 $\boldsymbol{Q} \leftarrow \boldsymbol{\Phi}\boldsymbol{\Psi}^{\mathbf{se}}$;   /* *Generate equivalent sensing measurement matrix*/

6 $\boldsymbol{f} = CS\text{-}recovery(\boldsymbol{Q}, \boldsymbol{Y}_p)$;   /* *Use* the *GPSR algorithm to reconstruct the sparse coefficient of the signal*/

7 $\overline{\boldsymbol{X}} = \boldsymbol{\Psi}^{\mathbf{se}-1}\boldsymbol{f}$;   /* *Recovery the entire sensory data*/

8 $MSE = \dfrac{\left\| vec(\boldsymbol{X}) - vec(\overline{\boldsymbol{X}}) \right\|_2^2}{\left\| vec(\boldsymbol{X}) \right\|_2^2}$ /* *Recovery the entire sensory data*/

3.3 Recovery performance Analysis of the Equivalent Sensing Matrix

According to the conclusion in [32], smaller mutual coherence will achieve accurate recovery of the original data X through fewer CS measurements. The following theorem indicates that according to the selection of the matrices A_i, the improved structure could reduce the mutual coherence compared to conventional KCS structure, so the new method formed by the improved structure could enhance the recovery performance of the signal.

Remark1: Suppose $\boldsymbol{\beta}^{\mathbf{N}} \in R^{N \times N}$ is an orthogonal matrix by shuffling the columns of $N \times N$ identity matrix, and $A_i^{T} = \begin{pmatrix} a_{11}^i & a_{12}^i & \cdots & a_{1T}^i \\ a_{21}^i & a_{22}^i & \cdots & a_{2T}^i \\ \vdots & \vdots & \vdots & \vdots \\ a_{T1}^i & a_{T2}^i & \cdots & a_{TT}^i \end{pmatrix} \in R^{T \times T}$ $i = 1, 2 \cdots N$. Then we can obtain

that $\boldsymbol{\beta} = \boldsymbol{R}_1\boldsymbol{\beta}^{\mathbf{N}}$, $A_i = \boldsymbol{R}_2 A_i^T$, where $\boldsymbol{R}_1$ samples n rows uniformly at random from matrix $\boldsymbol{\beta}^{\mathbf{N}}$ and $\boldsymbol{R}_2$ samples m rows uniformly at random from A_i^T [29].

Let $A_{\mathbf{BLOK}}^{\mathbf{NT}} \in R^{NT \times NT}$ be a block diagonal matrix given by the following structure, in which A_i^T is the above expression:

$$A_{\mathbf{BLOK}}^{\mathbf{NT}} = \begin{pmatrix} A_1^T & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & A_2^T & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & A_N^T \end{pmatrix}$$

Then the matrix $\boldsymbol{\Phi}^{\mathbf{N}} = (\boldsymbol{\beta}^{\mathbf{N}} \otimes \boldsymbol{I}_{\mathbf{T}})A_{\mathbf{BLOK}}^{\mathbf{NT}} \in R^{TN \times TN}$ is also an orthogonal matrix, in which $\boldsymbol{I}_{\mathbf{T}}$ is the $T \times T$ identity matrix.

Theorem 1：   Let $\boldsymbol{\Phi}^{\mathbf{N}} = (\boldsymbol{\beta}^{\mathbf{N}} \otimes \boldsymbol{I}_{\mathbf{T}})A_{\mathbf{BLOK}}^{\mathbf{NT}} \in R^{NT \times NT}$ denotes an orthogonal matrix, $\boldsymbol{\Psi}^{\mathbf{se}} = \boldsymbol{\Psi}_{\mathbf{y}} \otimes \boldsymbol{\Psi}_{\mathbf{x}} \in R^{NT \times NT}$ is a separable orthogonal base on R^{NT}, in which $\boldsymbol{\Psi}_{\mathbf{x}} \in R^{T \times T}$ and $\boldsymbol{\Psi}_{\mathbf{y}} \in R^{N \times N}$ are the base matrices. Then the mutual coherence $\mu(\boldsymbol{\Phi}^{\mathbf{N}}, \boldsymbol{\Psi}^{\mathbf{se}})$ obeys the following inequality:

$$\mu(\boldsymbol{\Phi}^{\mathbf{N}}, \boldsymbol{\Psi}^{\mathbf{se}}) = \mu((\boldsymbol{\beta}^{\mathbf{N}} \otimes \boldsymbol{I}_{\mathbf{T}})A_{\mathbf{BLOK}}^{\mathbf{NT}}, \boldsymbol{\Psi}_{\mathbf{y}} \otimes \boldsymbol{\Psi}_{\mathbf{x}}) \leq \mu_{\max}\mu\left(\boldsymbol{\beta}^{\mathbf{N}}, \boldsymbol{\Psi}_{\mathbf{y}}\right) \tag{9}$$

where $\mu_{\max} = \max\limits_{i=1,2\cdots N}\left\{\mu(A_i^T,\boldsymbol{\Psi}_x)\right\}$.

Proof: Using the definition of μ in (3), the quantity of interest is $\mu(\boldsymbol{\Phi}^N,\boldsymbol{\Psi}^{se})=\max\limits_{1\le p,q\le NT}\left|\left\langle\boldsymbol{\phi}_p,\psi_q^{se}\right\rangle\right|$,

where $\boldsymbol{\phi}_p$ are the rows of $\boldsymbol{\Phi}^N$ and ψ_q^{se} are the columns of $\boldsymbol{\Psi}^{se}$.

In further, $\left\langle\boldsymbol{\phi}_p,\psi_q^{se}\right\rangle=\left\langle\left[(\boldsymbol{\beta}^N\otimes\boldsymbol{I}_T)A_{\text{BLOK}}^{NT}\right]_p,\psi_q^{se}\right\rangle,\ p,q=1,2\cdots NT$.

When the row index p of $\boldsymbol{\Phi}^N$ satisfies $p\in\left[(k-1)T+1,kT\right]\quad k=1,2,\cdots N$, the p-th row of $\boldsymbol{\Phi}^N$ can be represented as the following formula (10):

$$\left[(\boldsymbol{\beta}^N\otimes\boldsymbol{I}_T)A_{\text{BLOK}}^{NT}\right]_p=(\beta_{k1}\mathbf{a}_l^1\ \ \beta_{k2}\mathbf{a}_l^2\ \ \cdots\beta_{kN}\mathbf{a}_l^N)\tag{10}$$

where $A_i^T=\left(\mathbf{a}_1^{i*},\mathbf{a}_2^{i*},\cdots\mathbf{a}_T^{i*}\right)^*$, $\mathbf{a}_l^i\in R^T$ is the l-th row vector of the matrix A_i^T, $i=1,2\cdots N,\ l=1,2\cdots T$.

Suppose that $\boldsymbol{\Psi}_y=\left(\psi_1^y,\psi_2^y\cdots\psi_N^y\right)$, $\psi_r^y\in R^N$ ($r=1,2\cdots N$)denotes the r-th column vector of the matrix $\boldsymbol{\Psi}_y$, $\boldsymbol{\Psi}_x=\left(\psi_1^x,\psi_2^x\cdots\psi_T^x\right)$, $\psi_s^x\in R^T$ ($s=1,2\cdots T$) denotes the s-th column vector of the matrix $\boldsymbol{\Psi}_x$.

Because $\boldsymbol{\Psi}^{se}=\boldsymbol{\Psi}_y\otimes\boldsymbol{\Psi}_x$, then the following formula (11) could be obtained:

$$\boldsymbol{\Psi}^{se}=\left(\psi_1^{se},\psi_2^{se}\cdots\psi_{NT}^{se}\right)\qquad\psi_q^{se}=\psi_r^y\otimes\psi_s^x\tag{11}$$

where ψ_q^{se} ($q=1,2\cdots NT$) denotes the q-th column vector of the matrix $\boldsymbol{\Psi}^{se}$.

According to above formulas (10) and (11), then we have:

$$\begin{aligned}\left\langle\left[(\boldsymbol{\beta}^N\otimes\boldsymbol{I}_T)A_{\text{BLOK}}^{NT}\right]_p,\psi_q^{se}\right\rangle&=\left\langle(\beta_{k1}\mathbf{a}_l^1\ \ \beta_{k2}\mathbf{a}_l^2\ \ \cdots\beta_{kN}\mathbf{a}_l^N\ ),\psi_r^y\otimes\psi_s^x\right\rangle\\&=\left\langle(\beta_{k1}\mathbf{a}_l^1\ \ \beta_{k2}\mathbf{a}_l^2\ \ \cdots\beta_{kN}\mathbf{a}_l^N\ ),vec(\psi_{1r}^y\psi_s^x,\psi_{2r}^y\psi_s^x,\cdots\psi_{Nr}^y\psi_s^x)\right\rangle\\&=\left(\beta_{k1}\psi_{1r}^y\right)\left\langle\mathbf{a}_l^1,\psi_s^x\right\rangle+\left(\beta_{k2}\psi_{2r}^y\right)\left\langle\mathbf{a}_l^2,\psi_s^x\right\rangle+\cdots+\left(\beta_{kN}\psi_{Nr}^y\right)\left\langle\mathbf{a}_l^N,\psi_s^x\right\rangle\end{aligned}\tag{12}$$

Where ψ_{or}^y represents the o-th row and r-th column entry of matrix $\boldsymbol{\Psi}_y$, $o=1,2\cdots N,\ r=1,2\cdots N$.

Consequently, we can conclude that

$$\begin{aligned}\mu(\boldsymbol{\Phi}^N,\boldsymbol{\Psi}^{se})&=\max_{1\le p,q\le NT}\left|\left\langle\boldsymbol{\phi}_p,\psi_q^{se}\right\rangle\right|=\max_{1\le p,q\le NT}\left|\left\langle\left[(\boldsymbol{\beta}^N\otimes\boldsymbol{I}_T)A_{\text{BLOK}}^{NT}\right]_p,\psi_q^{se}\right\rangle\right|\\&=\max_{1\le k,r\le N}\left\{\max_{1\le s,l\le T}\left|\left(\beta_{k1}\psi_{1r}^y\right)\left\langle\mathbf{a}_l^1,\psi_s^x\right\rangle+\left(\beta_{k2}\psi_{2r}^y\right)\left\langle\mathbf{a}_l^2,\psi_s^x\right\rangle+\cdots+\left(\beta_{kN}\psi_{Nr}^y\right)\left\langle\mathbf{a}_l^N,\psi_s^x\right\rangle\right|\right\}\\&\le\max_{1\le k,r\le N}\left\{\left|\left(\beta_{k1}\psi_{1r}^y\right)\mu\left(A_1^T,\boldsymbol{\Psi}_x\right)+\left(\beta_{k2}\psi_{2r}^y\right)\mu\left(A_2^T,\boldsymbol{\Psi}_x\right)+\cdots+\left(\beta_{kN}\psi_{Nr}^y\right)\mu\left(A_N^T,\boldsymbol{\Psi}_x\right)\right|\right\}\end{aligned}\tag{13}$$

where $\mu\left(A_i^T,\boldsymbol{\Psi}_x\right)=\max\limits_{1\le s,l\le T}\left|\left\langle\mathbf{a}_l^i,\psi_s^x\right\rangle\right|,i=1,2\cdots N$,that is,

$$\max_{1\le s, l\le T}\left|\left\langle \mathbf{a}_l^1, \boldsymbol{\psi}_s^x\right\rangle\right|=\mu\left(A_1^T, \boldsymbol{\Psi}_x\right), \max_{1\le s, l\le T}\left|\left\langle \mathbf{a}_l^2, \boldsymbol{\psi}_s^x\right\rangle\right|=\mu\left(A_2^T, \boldsymbol{\Psi}_x\right), \cdots, \max_{1\le s, l\le T}\left|\left\langle \mathbf{a}_l^N, \boldsymbol{\psi}_s^x\right\rangle\right|=\mu\left(A_N^T, \boldsymbol{\Psi}_x\right) \quad (14)$$

Let $\mu_{\max}=\max_{i=1,2\cdots N}\left\{\mu(A_i^T, \boldsymbol{\Psi}_x)\right\}$. According to the expression (13), $\mu(\boldsymbol{\Phi}^N, \boldsymbol{\Psi}^s)$ can be bounded as:

$$
\begin{aligned}
\mu(\boldsymbol{\Phi}^N, \boldsymbol{\Psi}^{se})&= \max_{1\le p, q\le N^2}\left|\left\langle \left[(\boldsymbol{\beta}^N\otimes \boldsymbol{I}_T)A_{\mathbf{BLOK}}{}^{NT}\right]_p, \psi_q^{se}\right\rangle\right| \\
&\le \max_{1\le k, r\le N}\left\{\left|\left(\beta_{k1}\psi_{1r}^y\right)\mu\left(A_1^T, \boldsymbol{\Psi}_x\right)+\left(\beta_{k2}\psi_{2r}^y\right)\mu\left(A_2^T, \boldsymbol{\Psi}_x\right)+\cdots+\left(\beta_{kN}\psi_{Nr}^y\right)\mu\left(A_N^T, \boldsymbol{\Psi}_x\right)\right|\right\} \\
&\le \max_{1\le k, r\le N}\left\{\mu_{\max}\left|\left(\beta_{k1}\psi_{1r}^y\right)+\left(\beta_{k2}\psi_{2r}^y\right)+\cdots+\left(\beta_{kN}\psi_{Nr}^y\right)\right|\right\} \\
&= \mu_{\max}\mu\left(\boldsymbol{\beta}^N, \boldsymbol{\Psi}_y\right)
\end{aligned}
\quad (15)
$$

where $\mu\left(\boldsymbol{\beta}^N, \boldsymbol{\Psi}_y\right)=\max_{1\le k, r\le N}\left|\left\langle \boldsymbol{\beta}_k, \psi_r^y\right\rangle\right|$, $\boldsymbol{\beta}_k$ is the *k-th* row of the matrix $\boldsymbol{\beta}^N$.

So the mutual coherence $\mu(\boldsymbol{\Phi}^N, \boldsymbol{\Psi}^{se})$ obeys

$$\mu(\boldsymbol{\Phi}^N, \boldsymbol{\Psi}^{se}) = \mu((\boldsymbol{\beta}^N\otimes \boldsymbol{I}_T)A_{\mathbf{BLOK}}{}^{NT}, \boldsymbol{\Psi}_y\otimes \boldsymbol{\Psi}_x)\le \mu_{\max}\mu\left(\boldsymbol{\beta}^N, \boldsymbol{\Psi}_y\right)$$

$$\#$$

Based on the expressions (2) and (9), we can conclude that the number of the measurements $M=mn$ satisfies $mn\ge CS\left(\max_{i=1,2\cdots N}\left\{\mu(A_i^T, \boldsymbol{\Psi}_x)\right\}\mu\left(\boldsymbol{\beta}^N, \boldsymbol{\Psi}_y\right)\right)^2\log(NT/\delta)$. This means that the number of measurements required for faithful recovery satisfies $mn=\mathrm{O}(S\left(\max_{i=1,2\cdots N}\left\{\mu(A_i^T, \boldsymbol{\Psi}_x)\right\}\mu\left(\boldsymbol{\beta}^N, \boldsymbol{\Psi}_y\right)\right)^2\log NT)$. If we select the appropriate matrices A_i to make this formula established: $\max_{i=1,2\cdots N}\left\{\mu(A_i^T, \boldsymbol{\Psi}_x)\right\}\mu\left(\boldsymbol{\beta}^N, \boldsymbol{\Psi}_y\right)=\mathrm{O}(1)$, then $mn=\mathrm{O}(S\log NT)$ and the proposed structure $\boldsymbol{\Phi}^N=(\boldsymbol{\beta}^N\otimes \boldsymbol{I}_T)A_{\mathbf{BLOK}}{}^{NT}\in R^{TN\times TN}$ will achieve the recovery performance comparable to that of random matrices.

Remark 2: Let

$$A_1^T = A_2^T = \cdots = A_N^T = A^T \tag{16}$$

When the (16) is satisfied, it is interesting to find that the structure $\boldsymbol{\beta}^N\otimes A^T$ of conventional KCS is a special case of the structure $(\boldsymbol{\beta}^N\otimes \boldsymbol{I}_T)A_{\mathbf{BLOK}}{}^{NT}$, and the following conclusion is consistent with the expression in [29]:

$$\mu(\boldsymbol{\Phi}^N, \boldsymbol{\Psi}^{se})=\mu((\boldsymbol{\beta}^N\otimes \boldsymbol{I}_T)A_{\mathbf{BLOK}}{}^{NT}, \boldsymbol{\Psi}_y\otimes \boldsymbol{\Psi}_x)=(\mu(\boldsymbol{\beta}^N\otimes A^T, \boldsymbol{\Psi}_y\otimes \boldsymbol{\Psi}_x)=\mu(A^T, \boldsymbol{\Psi}_x)\mu\left(\boldsymbol{\beta}^N, \boldsymbol{\Psi}_y\right) \quad (17)$$

where $\mu_{\max}=\mu(A^T, \boldsymbol{\Psi}_x)$. In other words, the new structure (5) extends the structure of conventional KCS, we could select appropriate matrices A_i to reduce $\mu\left(A_i^T, \boldsymbol{\Psi}_x\right)$ and thus to reduce the mutual coherence $\mu(\boldsymbol{\Phi}^N, \boldsymbol{\Psi}^{se})$.

4. Simulations and Analysis

In this section, we will evaluate the performance of our proposed method by virtue of the Matlab simulation.

4.1 Simulation Environment

Two different metrics are used to measure the performance of our proposed method. For energy efficiency of the sensor network data gathering, the total number of transmitted packets of each sensing period is used to measure the lever of energy consumption in our simulations. Here, we use the typical hypothesis that both sensing energy and energy consumed for computation can be neglected in WSNs, and so the energy consumption is dominated by radio communication in transmitting and receiving data packets [22]. For the reconstruction quality of the sensor data, the mean square error(MSE) is applied to measure the recovery accuracy of different methods and defined as following formulation: $MSE = \left\| vec(X) - vec(\overline{X}) \right\|_2^2 \Big/ \left\| vec(X) \right\|_2^2$,

where $\overline{X}$ represents the recovery data, X represents the sensor data and $\left\| \cdot \right\|_2$ is the L_2 norm of a vector.

We consider a WSN with N sensors randomly deployed in a unit square, and the number of monitoring time slots for each sensing period is T . We assume that a shorted path tree has already been built in the network. In order to fairly compare the recovery results with the typical methods in [28], Ψ_x and Ψ_y are chosen to be the discrete cosine transformation (DCT) matrices with the sizes of $T \times T$ and $N \times N$, respectively. The source sparse signal X is artificially generated based on Ψ_x , Ψ_y and the coefficient matrix χ of random S -sparsity. The sampling rate of the signal X is chosen to be 0.25.

4.2 Reconstruction Performance

In this section, we carry out simulations to evaluate the performance of our method. The evaluation process mainly includes two aspects: energy consumption of data collection in WSN and the recovery performance of the spatiotemporal data for different dimensions of the sensor networks. The gradient projection algorithm (GPSR) [36] is employed to reconstruct the sensor data from CS measurements. To validate the recovery performance of the proposed approach, we utilize the performance of the following different measurement methods for compared empirically.

1. Global measurements (dense) for signals of any dimension multiplex all the values of the signal together. According to CS theory, the global random Gaussian matrix (GRG method) measuring $vec(X)$ is universally incoherent with most fixed transforms (with overwhelming probability). So that GRG method is used as a baseline, which represents the optimal performance.

2. The CS-ST method in [28] exploits the spatiotemporal correlations, in which each sensor node separately compresses its length T reading data into t linear measurements $t < T$, then the measurements of each node are conveyed to the sink node. For fair comparison in simulations, the sampling rate of the signal is also 0.25, so the value of t is $0.25T$. The reading data of each node is measured by distinct $t \times T$ Gaussian matrix, in which the elements are independent identically distributed

(i.i.d.) random variables, so the measurement matrix is a block-diagonal matrix formed by N Gaussian matrices.

3. The classical kronecker measurement matrix obtained from the kronecker product of two dense Gaussian matrices (KGG method), which represent the optimal performance of conventional KCS model. The signal is sampled at the same sampling rate 0.5 on each dimension.

In the proposed efficient measurement method, the measurements of a subset of sensor nodes are sent instead of transmitting the measurements of each node. Therefore, the selected nodes n would send the temporal measurements to the sink node, in which the selected nodes n are a fraction of WSN nodes N. When the sampling rate 0.25 of the signal X is fixed ($mn = 0.25TN$), the value of n varied with different m. We should select the appropriate value of n to transmit data as small as possible without the loss of the recovery performance. For the case $N = 64$ and $T = 32$ in WSN, we simulate the recovery performance of the proposed method with different proportions of sensor nodes N to select the appropriate value of n.

By repeating the trails 500 times in each simulation and calculating the fraction of the successful recovery, the empirical probability of the successful recovery is plotted as a function of the signal sparsity S in **Fig. 2**. The x-axis of the performance curves represents the values of the signal sparsity S vary from 75 to 145. The y-axis, instead, shows the probability of the signal successful recovery, where the successful recovery is achieved if the mean square error of the estimated sensor readings is less than 10^{-3} ($MSE < 0.001$). By using distinct i.i.d. Gaussian matrix for A_i^T ($A_i = R_2 A_i^T$) in the proposed method, **Fig. 2** shows that the recovery performance varies depending on the different values of n. For the values of $n = [0.89N]$ and $n = [0.74N]$ ([•] denotes the rounding operation), the recovery performance is almost the same as GRG method. And for the values of $n = [0.8N]$ and $n = [0.61N]$, the recovery performance declined compared with GRG method. The reason is that the sparsities of the temporal and spatial signal are different, and the selected measurements m and n have a certain matching requirement. So in the following simulations, we select the $n = [0.74N]$ with the optimal recovery performance.
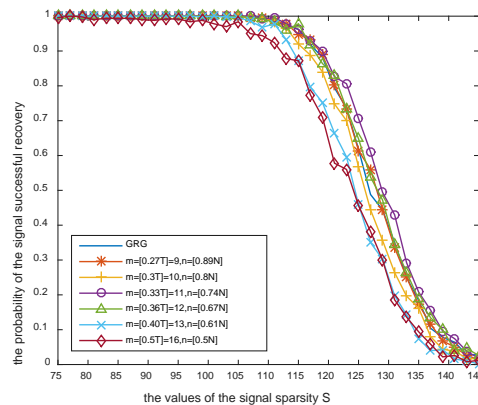

Fig. 2. When the sampling rate 0.25 of the signal X is fixed ($mn = 0.25TN$), the successful recovery probabilities of X with different proportions of N in different sparsities S.

A. Energy consumption of data collection in WSN

In terms of energy efficiency, for the case of $N = 64$ and $T = 32$, the total number of wireless transmission packets is counted by KGG, CS-ST and the efficient measurement method, respectively. Data transmission of dense GRG is too large to compare. For fair comparing with CS-ST and KGG methods, our proposed efficient measurement method is also used distinct i.i.d. Gaussian matrix for A_i^T $\forall i \in N$. Meanwhile we select $m = 11$ and $n = \lceil 0.74N \rceil \approx 47$ in simulation. In **Fig. 3**, the curves of the number of packets are plotted and the sensing period is increasing from 100 to 5000. It is clearly shown that the efficient measurement method transmitted fewer data packets than CS-ST and KGG, which indicates that the proposed measurement method collects data in a more energy efficient manner for WSN. Compared with CS-ST, the number of wireless transmissions is reduced by almost 25%. This is because the efficient measurement method makes full use of both the spatial and temporal correlation during data gathering, which doesn't need each node to send the measurements to the sink node.
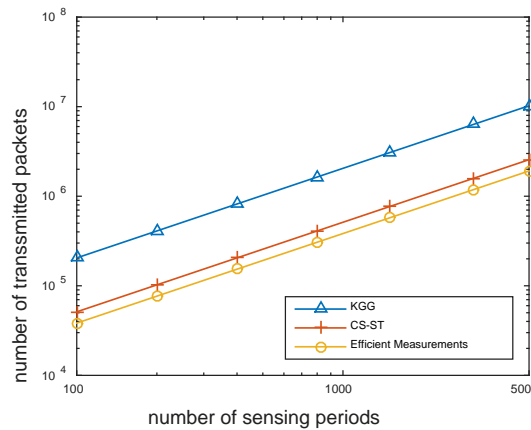

Fig. 3. The total number of wireless transmission packets by KGG, CS-ST and the proposed efficient measurement method for the case of $N = 64$ $T = 32$ and A_i^T is distinct i.i.d. Gaussian matrix.

B. The recovery performance of the spatiotemporal sensory data for different dimensions of the sensor networks

The advantage of the proposed efficient measurement method is that collecting the measurements of a subset of nodes could achieve the same data reconstructed performance as collecting the measurements of each node. In **Fig. 4**, the average successful reconstructed probabilities of GRG, CS-ST, the proposed efficient measurement method and KGG are shown for case $N = 64$ and $T = 32$. For fair comparison, we also use distinct i.i.d. Gaussian matrix for A_i^T in our proposed efficient measurement method in subfigure 1**(a) of Fig. 4**. In order to illustrate the validity of our proposed structure, in subfigure 2**(b) of Fig. 4**, we also consider distinct sparse matrix [23][24] for A_i: all entries are zero, except for a single one in each row and at most a single one in each column. The simulation environment is the same as in **Fig. 3**.
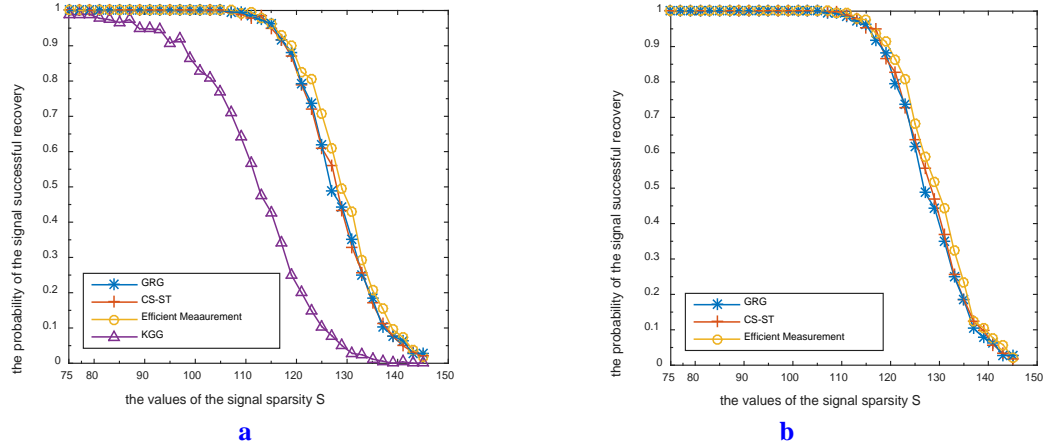
Fig. 4. The recovery performance of sensor data for case $N = 64, T = 32$ in different sparsities S.
a Subfigure 1 The measurement methods are GRG, CS-ST, the proposed efficient measurement method
and KGG. A_i^T ($A_i = R_2 A_i^T$) is distinct i.i.d. Gaussian matrix.
b Subfigure 2 The measurement methods are GRG, CS-ST and the proposed efficient measurement
method. A_i is distinct sparse matrix.

In subfigure 1**(a)** of **Fig. 4**, it is clear shown that the proposed efficient measurement method has slightly better recovery performance than CS-ST and GRG. Meanwhile, the reconstructed performance of the efficient measurement method outperforms that of KGG by about 40% on sparsity of near 110. That is because the operations of distinct measurement matrix on each sampling node data reduce the mutual coherence of the measurement matrix and sparsifying basis. In particular, with the same sampling rate 0.25, the sparsity values required for 95% successful recovery are 117, 116, 117 and 95 for GRG, CS-ST, the efficient measurement method and KGG, respectively. In subfigure 2**(b) of Fig. 4**, the same trend exists for the efficient measurement method with distinct sparse matrix, which has the same recovery performance as CS-ST and GRG method.

To compare with the different ratio of T/N, similar to **Fig. 4**, the recovery performance curves of GRG, CS-ST, the proposed efficient measurement method and KGG are given for the other case $N = 100$ and $T = 12$ in **Fig. 5**. The values of the signal sparsity S vary from 40 to 100. For our proposed efficient measurement method in subfigure 3**(c) of Fig. 5**, we also use distinct i.i.d. Gaussian matrix for A_i^T ($A_i = R_2 A_i^T$). And in subfigure 4**(d) of Fig. 5**, we consider distinct sparse matrix for A_i. The sampling rate is also 0.25, so the parameters of A_i is $m = 4$ and the number of collected sensor nodes is $n = 75$. It is clear in **Fig. 5** that the trend of successful recovery for different methods is almost the same as that in **Fig. 4**. Specifically, the proposed efficient measurement method exhibits almost the same performance as that of GRG and CS-ST. The simulations in **Fig. 4** and **Fig. 5** indicate that the proposed efficient measurement method could be applied to the different ratio of T/N.
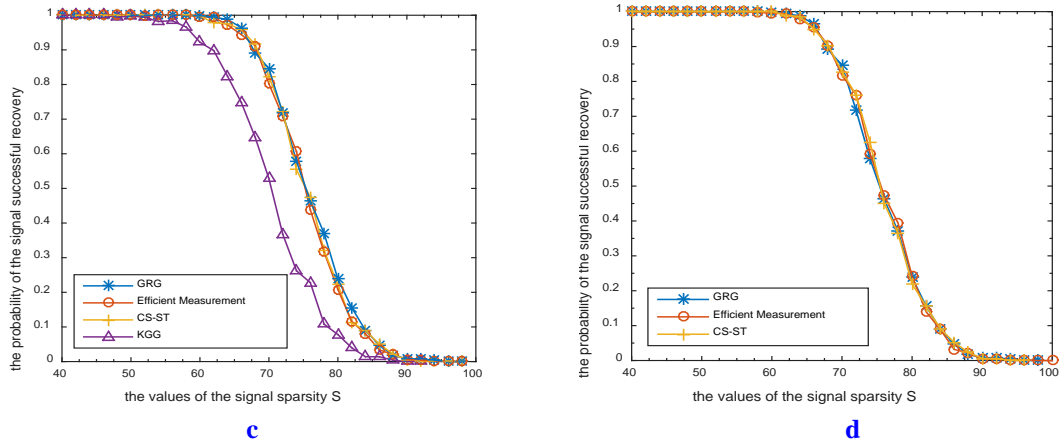
| c | d |

Fig. 5. The recovery performance of sensor data for case $N = 100, T = 12$ in different sparsities S.
c Subfigure 3 The measurement methods are GRG, CS-ST, the proposed efficient measurement method and KGG. A_i^T ($A_i = R_2 A_i^T$) is distinct i.i.d. Gaussian matrix.
d Subfigure 4 The measurement methods are GRG, CS-ST and the proposed efficient measurement method. A_i is distinct sparse matrix.

In the previous simulations, the number of network nodes is small. To see how the efficient measurement method affects the recovery performance of data in large-scale WSN, we simulated another scenario with $N = 500$ and $T = 6$. The values of the signal sparsity S vary from 95 to 205. We compare the average successful reconstructed probabilities of the proposed efficient measurement method with that of GRG in **Fig. 6**, where A_i^T is distinct i.i.d. Gaussian matrix in subfigure 5(**e**) of **Fig. 6** and A_i is distinct sparse matrix in subfigure 6(**f**) of **Fig. 6**. The sampling rate is also 0.25, so the parameters of A_i is $m = 2$ and the number of collected sensor nodes is $n = 375$. It is shown in **Fig. 6** that the efficient measurement method has almost the similar recovery performance as GRG. In particular, with the same sampling rate 0.25, the sparsity values required for 95% successful recovery are 170 and 172 for GRG and the efficient measurement method, respectively. As can be seen in these simulations above, the trend of the probability of success for the networks with different dimensions is similar.
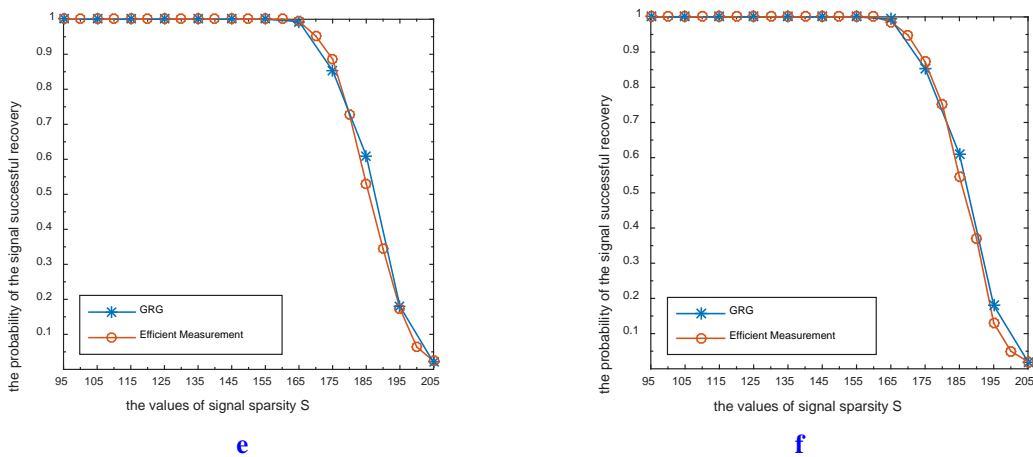

| e | f |

Fig. 6. The recovery performance of sensor data for case $N = 500, T = 6$ in different sparsities S.

e Subfigure 5 The measurement methods are GRG and the proposed efficient measurement method.
A_i^T ($A_i = R_2 A_i^T$) is distinct i.i.d. Gaussian matrix.

f Subfigure 6 The measurement methods are GRG and the proposed efficient measurement method.
A_i is distinct sparse matrix.

Compared with the conventional methods, the above simulation results indicate that the proposed efficient measurement method could reduce the energy consumption of the entire network in the process of the data collection without the loss of the reconstruct performance, when data show the spatial and temporal correlations. By virtue of the powerful decoding ability of the sink node and the sink node could accept the signal processing delay in non-real-time WSN application scenario, the efficient measurement method will be promising in practical implementation. The advantage of the proposed method is to remove the redundant data, so the method will be expectedly extended to the large-scale WSN.

5. Conclusion and Future Work

This paper has presented a framework for approximate data gathering which exploits both spatial and temporal correlations among the sensor readings in WSNs in order to achieve potentially higher compression ratio and improve the energy efficiency. In particular, instead of conveying the CS linear measurements of every sensor node to the sink node, only a subset of all sensor nodes according to our designed method is randomly selected to transmit the measurements to the sink node. Thereby the framework could reduce the number of wireless transmissions without sacrificing the data recovery performance. Numerical simulations suggest that the proposed measurement method is more energy efficient for data gathering in multi-hop WSNs.

Meanwhile, the two-dimensional separable operator formed by the proposed data gathering method improved the recovery performance of the conventional KCS method. In the future work, we can jointly design a better method considering both routing and SNR to further improve the performance of data collection in WSN. The properties of the proposed two-dimensional separable operator will be studied in more detail, which could be applied to the recovery of the other two-dimensional signal.

Appendix A

The relevant mathematical model:
During each sensing period T, n ($n < N$) different sensor nodes are selected randomly and uniformly from all N sensor nodes. That means the subset that formed by n selected nodes will change when the period T changes, so after some periods, each node could be balanced participation from the perspective of probability. The probability statistics model of the random selection is established as follows:

(I): During a sensing period T, n different sensor nodes are selected randomly and uniformly from all N nodes. In an experiment, the event that the node i is selected can be denoted as E_i, $i = 1, 2 \cdots N$, and we assume that:

$$Y_i = \begin{cases} 1 & node \ i \ is \ selected, \ the \ event \ E_i \ occurred \\ 0 & node \ i \ is \ not \ selected, \ the \ event \ E_i \ is \ not \ occurred \end{cases}$$

Then Y_i is a random variable that obeys two-points distribution. There are C_N^n kinds of the subsets that formed by n different nodes. Because of the random selection, the probability of each node i being selected is equal:

$$p(Y_1 = 1) = p(Y_2 = 1) = p(Y_3 = 1) = \cdots = p(Y_N = 1) = p_T .$$

(II): For different sensing periods, the selection of one node i could be seen as independent repeated experiments. In the T_T independent repeated experiments, the number of occurrences of the event E_i, is a random variable X_i, which obeys the binomial distribution $X_i \sim B(T_T, p_T)$. We have the probability: $P(X_i = k) = C_{T_T}^k p_T^k (1-p_T)^{T_T - k}$ $k = 0,1,2 \cdots T_T$, $T_T \geq 1, 0 \leq p_T \leq 1$

Then mathematical expectation satisfies $E(X_i) = T_T p_T$ according to the properties of the binomial distribution. So based on Bernoulli's law of large numbers, the average number of occurrences of the event E_i is $T_T p_T$. Because the status of all N nodes is equal, so according to the rotation symmetry principle, we have $E(X_1) = E(X_2) = \cdots = E(X_N)$.

So from the perspective of probability, after multiple sensing periods, the average number of times that each node is selected is equal.

References

[1]   Akyildiz I F, Su W, Sankarasubramaniam Y, et al., "Wireless sensor networks: a survey," *Computer Networks*, vol.38, no.4, pp. 393-422, March, 2002. Article (CrossRef Link).

[2]   Mainwaring A., Polastre J.,Szewczyk R., et al.,"Wireless sensor networks for habitat monitoring," in *Proc. of WSNA'02 pro. Of the 1st ACM International Workshop on Wireless Sensor Networks and Applications,* pp. 88-97, September 28-28, 2002. Article (CrossRef Link).

[3]   Z. Xiong, A. Liveris, and S. Cheng, "Distributed source coding for sensor networks", *IEEE Signal Processing Magazine,* vol. 21, no. 5, pp. 80–94, September, 2004. Article (CrossRef Link).

[4]   G. Anastasi, M. Conti, M. Di Francesco, and A. Passarella, "Energy conservation in wireless sensor networks: A survey," *Ad Hoc Networks,* vol. 7, no. 3, pp. 537–568, May, 2009. Article (CrossRef Link).

[5]   D. L. Donoho, "Compressed sensing," *IEEE Transactions Information Theory*, vol. 52, no. 4, pp. 1289–1306, April, 2006. Article (CrossRef Link).

[6]   E. Candes, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Transactions Information Theory*, vol. 52, no. 2, pp. 489–509, February, 2006. Article (CrossRef Link).

[7]   E. Candes and T. Tao, "Near-optimal signal recovery from random projections: Universal encoding strategies?," *IEEE Transactions Information Theory*, vol. 52, no. 12, pp. 5406–5425, December, 2006. Article (CrossRef Link).

[8]   E. J. Candes and M. B. Wakin, "An introduction to compressive sampling," *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 21–30, March, 2008. Article (CrossRef Link).

[9]   Haupt J., Nowak R., "Signal reconstruction from noisy random projections," *IEEE Transactions on Information Theory*, vol. 52,no. 9, pp. 4036-4048, September, 2006. Article (CrossRef Link).

[10]   Bajwa W U., Haupt J D., Sayeed A M, et al., "Joint source–channel communication for distributed estimation in sensor networks," *IEEE Transactions on Information Theory*, vol. 53**,** no. 10, pp. 3629-3653, October, 2007. Article (CrossRef Link).

[11] Luo C., Wu F., Sun J., et al., "Compressive data gathering for large-scale wireless sensor networks," in *Proc. of 15th annual International Conf. on Mobile Computing and Networking*, pp. 145-156, September 20-25, 2009. Article (CrossRef Link).

[12] Luo C., Wu F., Sun J., et al., "Efficient measurement generation and pervasive sparsity for compressive data gathering," *IEEE Transactions on Wireless Communications*, vol. 9, no. 12, pp. 3728-3738, December, 2010. Article (CrossRef Link).

[13] Wang J, Tang S, Yin B, et al., "Data gathering in wireless sensor networks through intelligent compressive sensing," in *Proc.* of *IEEE INFOCOM*, pp. 603-611, March 25-30, 2012. Article (CrossRef Link).

[14] Wang W, Garofalakis M, Ramchandran K.W., "Distributed sparse random projections for refinable approximation," in *Proc. of the 6th International conference on Information processing in sensor networks*, pp. 331-339, April 25-27, 2007. Article (CrossRef Link).

[15] Lee S., Pattem S., Sathiamoorthy M., et al., "Compressed sensing and routing in multi-hop networks," *Usc Ceng Technical Report*, 2010. Article (CrossRef Link).

[16] M. Sartipi and R. Fletcher, "Energy-efficient data acquisition in wireless sensor networks using compressed sensing," in *Proc. of the 2011 Data Compression Conference,* pp. 223–232, March 29–31, 2011. Article (CrossRef Link).

[17] Wu X, Xiong Y, Yang P, et al., "Sparsest Random Scheduling for Compressive Data Gathering in Wireless Sensor Networks," *IEEE Transactions on Wireless Communications*, vol. 13, no. 10, pp. 5867-5877, October, 2014. Article (CrossRef Link).

[18] Xiang L, Luo J, Vasilakos A., "Compressed data aggregation for energy efficient wireless sensor networks," in *Proc. 8th Annual IEEE Communication Society Conf. on Sensor, Mesh and Ad Hoc Communications and Networks*, pp. 46-54, June, 27-30, 2011. Article (CrossRef Link).

[19] Chen W, Wassell I J., "Optimized node selection for compressive sleeping wireless sensor networks," *IEEE Transaction on Vehicular Technology*, vol. 65, no. 2, pp. 827-836, February, 2015. Article (CrossRef Link).

[20] Chen W, Wassell I J., "Compressive sleeping wireless sensor networks with active node selection," in *Proc. of IEEE Global Communications Conference*, pp. 7-12, December 8-12, 2014. Article (CrossRef Link).

[21] Xiang L, Luo J, Rosenberg C., "Compressed data aggregation: Energy-efficient and high-fidelity data collection," *IEEE/ACM Transactions on Networking.*, vol. 21, no. 6, pp. 1722-1735, December, 2013. Article (CrossRef Link).

[22] Chou C T, Rana R, Hu W., "Energy efficient information collection in wireless sensor networks using adaptive compressive sensing," in *Proc. of IEEE 34th Conf. on Local Computer Networks,* pp. 443-450, October 20-23, 2009. Article (CrossRef Link).

[23] G. Quer, R. Masiero, G. Pillonetto, et al., "Sensing, compression, recovery for WSNs: Sparse signal modeling and monitoring framework," *IEEE Transactions on Wireless Communications*, vol. 11, no. 10, pp. 3447–3461, October, 2012. Article (CrossRef Link).

[24] M. Leinonen, M. Codreanu, and M. Juntti, "Sequential compressed sensing with progressive signal reconstruction in wireless sensor networks," *IEEE Transactions on Wireless Communications,* vol. 14, no. 3, pp.1622–1635, March, 2015. Article (CrossRef Link).

[25] He J, Sun G, Li Z, et al.,"Compressive data gathering with low-rank constraints for Wireless Sensor networks," *Signal Processing*, vol. 131, no. C, pp. 73-76, February, 2017. Article (CrossRef Link).

[26] Gong B, Cheng P, Chen Z, et al., "Spatiotemporal compressive network coding for energy-efficient distributed data storage in wireless sensor networks," *IEEE Communications Letters*, vol. 19, no.5, pp. 803-806, May,2015. Article (CrossRef Link).

[27] Wang C, Cheng P, Chen Z, et al., "Practical spatiotemporal compressive network coding for energy-efficient distributed data storage in wireless sensor networks," in *Proc. of IEEE 81st. Vehicular Technology Conference*, pp. 1-6, May 11-14, 2015. Article (CrossRef Link).

[28] M. Leinonen, M. Codreanu, and M. Juntti, "Distributed correlated data gathering in wireless sensor networks via compressed sensing," in *Proc. of Asilomar Conference on Signals, Systems and Computers*, pp. 418–422, November 3–6, 2013. Article (CrossRef Link).

[29] Y. Rivenson and A. Stern, "Compressed imaging with a separable sensing operator," *IEEE Signal Processing Letters,* vol. 16, no. 6, pp. 449–452, June, 2009. Article (CrossRef Link).

[30] Duarte M F and Baraniuk R G., "Kronecker product matrices for compressive sensing," in *Proc. of IEEE International Conference on Acoustics Speech and Signal Processing,* pp. 3650-3653, March 14-19, 2010. Article (CrossRef Link).

[31] Duarte M F and Baraniuk R G., " Kronecker Compressive Sensing," *IEEE Transactions on Image Processing,* vol. 21, no. 2, pp. 494-504, February, 2012. Article (CrossRef Link).

[32] E. Candès and J. Romberg, "Sparsity and incoherence in compressive sampling," *Inverse Problems*, vol. 23, no. 3, pp. 969-985, April, 2007. Article (CrossRef Link).

[33] Yang X, Tao X, Dutkiewicz E, et al., "Energy-Efficient Distributed Data Storage for Wireless Sensor Networks Based on Compressed Sensing and Network Coding," *IEEE Transactions on Wireless Communications*, vol. 12, no. 10, pp. 5087–5099, October, 2013. Article (CrossRef Link).

[34] I. F. Akyildiz, M. C. Vuran, and O.B. Akan, "On exploiting spatial and temporal correlation in wireless sensor networks," in *Proc. of WiOpt 2004: Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks,* pp. 71–80, March 24–26, 2004. Article (CrossRef Link).

[35] Luo C, Sun J, Wu F., "Compressive Network Coding for Approximate Sensor Data Gathering," in *Proc. of IEEE Global Telecommunications Conference*, pp. 1-6, December 5-9, 2011. Article (CrossRef Link).

[36] M. A. T. Figueiredo, R. D. Nowak, and S. J. Wright, "Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems," *IEEE Journal of Selected Topics in Signal Process*ing, vol. 1, no. 4, pp. 586–597, December, 2007. Article (CrossRef Link).

Xiao Xue is currently pursuing a doctoral degree at the School of Communication Engineering in Xidian University, Xi'an, Shaanxi, P. R. China. She received her M.S. degree in Applied Mathematics from Northwest University, Xi'an, Shaanxi, P. R. China, in 2009. Her current research interests include Compressive Sensing and its applications in wireless sensor networks.

Song Xiao is a professor in Xidian University, Xi'an, Shaanxi, P. R. China. She received her M.S degree in Communication and Information System and PhD degree in Signal and Information Processing in Xidian University, Xi'an, China in 2001 and 2004, respectively. Her research interests are in the fields of Joint Source Channel Coding, Network Coding, Compressive Sensing, Robust image and Video Communication etc.

Lei Quan is currently a post doctorate at Xidian University, Xi'an, Shaanxi, P. R. China. He received his B.E. degree of Telecommunication Engineering and PhD degree of Information and Communication Engineering from Xidian University, Xi'an, Shaanxi, P. R. China.His current research interests include Intelligent Information Processing, Compressive Sensing and Inverse Problems.