

Article

Efficient Data Gathering Methods in Wireless Sensor Networks Using GBTR Matrix Completion

Donghao Wang, Jiangwen Wan *, Zhipeng Nie, Qiang Zhang and Zhijie Fei

School of Instrumentation Science and Opto-Electronics Engineering, Beihang University, Beijing 100191, China; dhwang@buaa.edu.cn (D.W.); buaapeng_1417@buaa.edu.cn (Z.N.); youleyuanzq@buaa.edu.cn (Q.Z.); feizhijie@163.com (Z.F.)

* Correspondence: jwwan@buaa.edu.cn; Tel: +86-10-8233-9889

Academic Editor: Xue-Bo Jin

Received: 21 July 2016; Accepted: 13 September 2016; Published: 21 September 2016

Abstract: To obtain efficient data gathering methods for wireless sensor networks (WSNs), a novel graph based transform regularized (GBTR) matrix completion algorithm is proposed. The graph based transform sparsity of the sensed data is explored, which is also considered as a penalty term in the matrix completion problem. The proposed GBTR-ADMM algorithm utilizes the alternating direction method of multipliers (ADMM) in an iterative procedure to solve the constrained optimization problem. Since the performance of the ADMM method is sensitive to the number of constraints, the GBTR-A2DM2 algorithm obtained to accelerate the convergence of GBTR-ADMM. GBTR-A2DM2 benefits from merging two constraint conditions into one as well as using a restart rule. The theoretical analysis shows the proposed algorithms obtain satisfactory time complexity. Extensive simulation results verify that our proposed algorithms outperform the state of the art algorithms for data collection problems in WSNs in respect to recovery accuracy, convergence rate, and energy consumption.

Keywords: wireless sensor networks; data gathering; compressive sensing; matrix completion; graph based transform; ADMM; A2DM2

1. Introduction

Wireless sensor networks (WSNs) are composed of large-scale, self-organized sensor nodes, which are capable of sensing, data storage, and communication. WSNs have lots of applications, such as remote environment sensing, industrial automation, smart city, and military monitoring. In practical applications, lots of ordinary nodes are deployed in an unattended mode. These ordinary nodes perform data collection tasks individually and transmit the raw data to the sink node in multi-hop access. Since it is difficult to recharge or replace the limited power supply of ordinary nodes, developing energy efficient data gathering methods is becoming crucial.

A large number of data collection methods have been proposed to reduce the energy consumption with different levels of data reconstruction precision in the literature [1–3]. These obtained data in WSNs possess spatial and temporal correlations, which are intrinsic characteristics of a physical environment. A previous article [1] proposed a clustered aggregation (CAG) algorithm for data collection, which utilizes the spatial and temporal correlations of the sensed data. Pham et al. [2] presented a divide and conquer approximating (DCA) algorithm to reduce power consumption. Since the sensed data require to be transmitted to the sink node in multihop communication, Rosana et al. [3] proposed a novel algorithm to construct spanning trees for efficient data gathering in wireless sensor networks. Unfortunately, data gathering methods in traditional mode have limitations. Firstly, the clustering methods (or the spanning tree construction methods) reflect high computational

cost, as well as the dynamic maintaining of clusters (or trees). Secondly, the energy consumption is not balanced since the nodes close to the sink consume more energy.

Compressive Sensing (CS) [4–7] theory has brought about a new approach for efficient data gathering in WSNs. Since the sensed data have temporal and spatial correlations, they can be sparsely represented in an appropriate transform basis. CS theory states that a small number of linear measurements can accurately reconstruct the sparse signals when the sensing matrix satisfies the restricted isometry property (RIP). Thus, the number of data transmissions for one measurement is largely reduced. Since the high energy-intensive reconstruction algorithm is implemented at the sink node, the computational cost between these ordinary nodes is quite low. Benefiting from the merits of CS, the energy consumption is balanced and reduced for data gathering problem in WSNs. Thus, many papers [8–11] have been published about efficient data gathering methods based on CS theory in recent years. Luo et al. [8] first proposed to apply compressive sensing for data gathering in WSNs. The idea of the proposed compressive data gathering (CDG) is that the intermediate nodes transmit the weighted sums of father nodes and their own data. In [9], a distributed compressive sampling method was presented. The method is quite efficient since in-network compression is employed, and each node individually determines the transmit scheme to minimize the total number of transmissions. Liu et al. [10] presented the compressed data aggregation (CDA) method to reconstruct the original signals with high precision. Meanwhile, the energy consumption is reduced in comparison with the CDG method. In [11], the authors proposed the compressive data collection (CDC) method to collect data in wireless sensor networks. The scheme reduces the necessary number of measurements, thus the network lifetime is prolonged.

However, the real application of CS in WSNs is difficult. Firstly, CS assumes the data are sparse or could be sparsely represented in a transform basis. Nevertheless, the appropriate sparse matrix basis is not always available. Secondly, the spatial correlation and the temporal correlation cannot be employed together since the sensed data are expressed in the vector form.

As a more efficient data gathering method, matrix completion (MC) [12] considers recovering the incomplete data matrix by observing a small part of the matrix elements. Actually, MC is an extension of the CS theory. In CS, the signals are represented in the vector form, while MC formulates the signals in the matrix form. The sensor data are commonly denoted as matrix, such as the image signals and the video samples. Thus, these two-dimensional signals can be computed more efficiently in the matrix form, although they could be transformed into the form of a vector. In comparison with CS theory, MC do not require to seek a priori sparse basis, and the necessary sampling ratio could be even lower. Since the sensed data collected in WSNs have spatial and temporal correlation, they show low-rank properties. In [13], the singular value thresholding (SVT) algorithm was proposed by approximating the low-rank matrix with a nuclear norm minimization method. To measure large-scale traffic datasets, Roughan et al. [14] proposed the spatial and temporal matrix completion algorithm, which was called the sparsity regularized matrix factorization (SRMF). In [14], intensive analysis of the massive traffic data resulted in the optimal choice of the spatial and temporal constraint matrices. SRMF can be extended to solve various matrix completion based problems, such as data interpolation, and missing matrix elements inference. To further take advantage of the low-rank feature and the short-term stability property of the sensed data, Cheng et al. [15] proposed the STCDG method. The recovery accuracy is improved, and the power consumption is reduced by applying STCDG for data gathering in WSNs.

Actually, the sensor nodes are deployed in a finite area. Therefore, the features of the sensed data are coupled with network topology information. In our analysis, the sensed data are found to be sparse under the graph based transform (GBT) basis. The GBT basis is composed of the eigenvector of the Laplacian matrix when the whole network is represented as a graph. To the best of our knowledge, this is the first time the GBT sparsity has been applied to a matrix completion problem. In consideration of both the GBT sparsity and the low-rank feature of the sensed data, the GBTR-ADMM and the GBTR-A2DM2 algorithm are proposed. The time complexity of our proposed algorithms are also

analyzed, which shows that they have a low complexity. Simulation results show our proposed algorithms outperform the state of art algorithms for data collection problems in respect to recovery accuracy, convergence rate, and energy consumption.

The main contributions of the paper are as follows:

- (1) The features of sensor datasets are analyzed in consideration of their topology information, which reveals that the data matrix is sparse under the graph based transform.
- (2) The graph based transform regularized (GBTR) Matrix Completion problem is formulated. To reconstruct the missing values efficiently, the GBTR by Alternating Direction Method of Multipliers (GBTR-ADMM) algorithm is proposed. Simulation results reveal that GBTR-ADMM outperforms the state of art algorithms in view of the recovery accuracy and the energy consumption.
- (3) To accelerate the convergence of GBTR-ADMM, GBTR-A2DM2 algorithm is proposed, which benefits from a restart rule and the fusion of multiple constraints.
- (4) The time complexity of our proposed algorithms is analyzed, which shows that the complexity is low.

The structure of the paper is concluded as follows: In Section 2, the problem formulation about matrix completion is given. Section 3 presents the features of the real datasets and the synthesized dataset. The proposed GBTR-ADMM and GBTR-A2DM2 algorithms are expatiated individually in Sections 4 and 5. Section 6 shows the time complexity of the proposed algorithms. In Section 7, the performances of the proposed algorithms are studied. The conclusions and the future works are summarized in Section 8.

2. Problem Formulation

In this section, we introduce the related issues in respect to matrix completion theory. The main notations of the paper are summarized in Table 1.

Table 1. Summary of notations.

M	Number of time slots
N	Number of sensor nodes
τ	The observed ratio
r	The matrix rank
λ	The GBT sparsity regularization parameter
β	The Lagrange penalty parameter
X	The original data matrix
\hat{X}	The reconstructed data matrix
M	The observed data matrix
D	The degree matrix
A	The adjacency matrix
L	The Laplacian matrix
Ψ	The GBT matrix
W	The introduced auxiliary variable
Z	The Lagrange multiplier

Suppose there are N sensor nodes in the WSNs. Using $\{x^i\}_{i=1}^N$ denotes the sensor data, where $x^i \in \mathbb{R}^M$ represents a data vector collected by node i in time slot t_1, t_2, \dots, t_m . The sample interval is assumed to be equal. Thus, the data matrix $X \in \mathbb{R}^{M \times N}$ can be used to represent the sensor data gathered by N sensor nodes in M time slots.

In order to reduce energy consumption in resource-constrained WSNs, only a small amount of sensor data is transmitted to the sink node. Let $\Omega \subset \{1, 2, \dots, M\} \times \{1, 2, \dots, N\}$ denotes the indices of the corresponding observed data of X . Similarly, let Ω^c denotes the indices of omitted value. Let π_Ω

be the linear projection operator that keeps the entries in Ω invariant and adjusts the entries in Ω^c to zero, that is:

$$(\pi_{\Omega}(\mathbf{X}))_{ij} = \begin{cases} X_{ij} & , \forall (i, j) \in \Omega \\ 0 & , \forall (i, j) \in \Omega^c \end{cases} \quad (1)$$

Suppose matrix $M \in \mathbb{R}^{M \times N}$ is the observed data, which is the incomplete version of matrix X with entries those outside Ω zeros. That is $\pi_{\Omega}(X_{ij}) = \pi_{\Omega}(M_{ij}), \forall (i, j) \in \Omega$.

Our goal is to reduce the amount of data transmission to the sink node, and to design relevant matrix completion algorithm to reconstruct the original data matrix X as closely as possible. The observed ratio is defined as:

$$\tau = \frac{\sum_{(i,j) \in \Omega} |\Omega(i,j)|}{MN} \quad (2)$$

Next, the features of the datasets are studied in detail, which would be utilized in our designed algorithms.

3. Exploring the Features of Datasets

In this subsection, three datasets are utilized for analysis. The first two datasets are gathered by the GreenOrbs [16] system, which is deployed in the forest environment with up to 330 nodes. Its topology is exhibited in Figure 1. We mainly consider the temperature and the humidity data collected between 3 and 5 August 2011. The third dataset is the smooth data generated with a second order Autoregressive (AR) model. In detail, the AR filter $H(z) = \frac{1}{1+a_1z^{-1}+a_2z^{-2}}$ is used, where a_1 and a_2 is predefined as -0.1 and -0.8 individually. Five hundred nodes assigned with the generated data are randomly deployed in a $1000 \text{ m} \times 1000 \text{ m}$ area, which is shown in Figure 2. The detailed information about these datasets is given in Table 2.

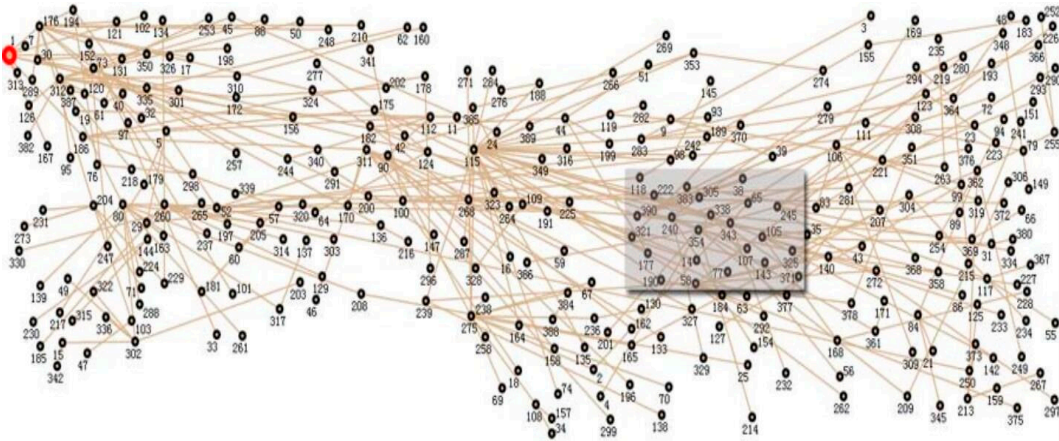


Figure 1. The real deployment topology of GreenOrbs.

Table 2. The experimental datasets.

Data Name	Data Types	Selected Data Matrix	Time Interval
GreenOrbs	Temperature	326×500	5 min
GreenOrbs	Humidity	326×500	5 min
Synthesized	AR model	500×500	-

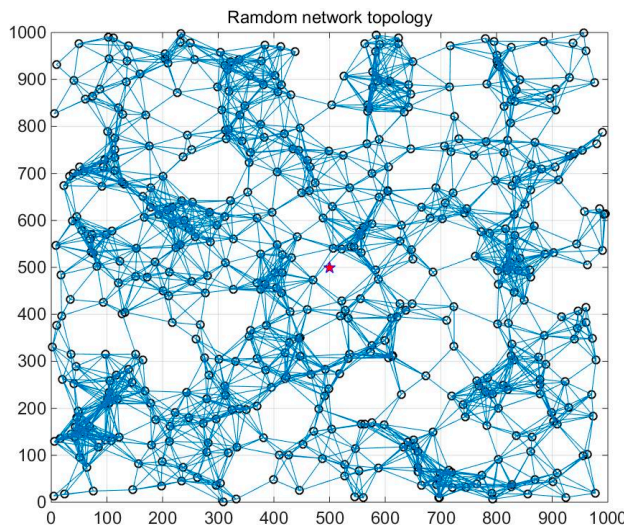


Figure 2. The random topology of synthesized data with 500 nodes in a 1000 m × 1000 m area.

3.1. Low-Rank Property

Since sensor readings have spatial and temporal correlations, the rank r of data matrix X would be small, such as $r \ll \min(M, N)$. This low-rank property of the sensor data has been studied in previous papers [14,15,17]. Let $\text{rank}(X)$ denote the rank of matrix X . Candes et al. [18] proposed to solve the matrix completion problem by minimizing $\text{rank}(X)$ under suitable constraints. However, the minimization problem of rank function cannot be figured out with a global solution in polynomial time because of its non-convexity. Fortunately, the nuclear norm $\|X\|_*$, which can be solved using various convex programming methods, is the tightest convex relaxation of the rank function [12]. Also, the relationship between rank function and nuclear norm in matrix completion is similar to the relation between l_0 norm and the convex l_1 norm in compressive sensing.

3.2. GBT Sparsity

Since these datasets are coupled with their topologies, we construct the graph-based transform (GBT) to sparsely represent them. The sensor network is represented by a graph $G = (V, E)$, which consists of a vertex set V (sensor nodes) and an edge set E (sensor links). The link $e(i, j)$ is supposed to exist if the Euclidean distance between any two nodes (node i and node j) is smaller than the communication range. The topological graph is mathematically denoted with the adjacent matrix A :

$$A(i, j) = \begin{cases} 1, & \text{if } e(i, j) \text{ is an element of } E \\ 0, & \text{otherwise.} \end{cases} \quad (3)$$

The degree matrix D is a diagonal matrix, where the diagonal elements $D(i, i)$ denote the number of links connected to node i , and $D(i, j) = 0, \forall i \neq j$. Thus, the Laplacian matrix can be induced as:

$$L = D - A \quad (4)$$

Since matrix L is symmetric, the eigenvalue decomposition can be obtained. Then, the eigenvector of the Laplacian matrix L constitutes the columns of the graph-based transform matrix, which is denoted as Ψ . Clearly, the GBT matrix Ψ is orthogonal. Detailed information about GBT basis can be found in [19].

The performance of the GBT matrix as a sparse basis is demonstrated in Figure 3. As can be seen, nearly 10% of the sorted GBT coefficients assemble more the 99% of the energy. Therefore, the matrix

$\Psi^{-1}X$ is extremely sparse. In other words, the l_1 norm of matrix $\Psi^{-1}X$, represented as $\|\Psi^{-1}X\|_1$, is very small.

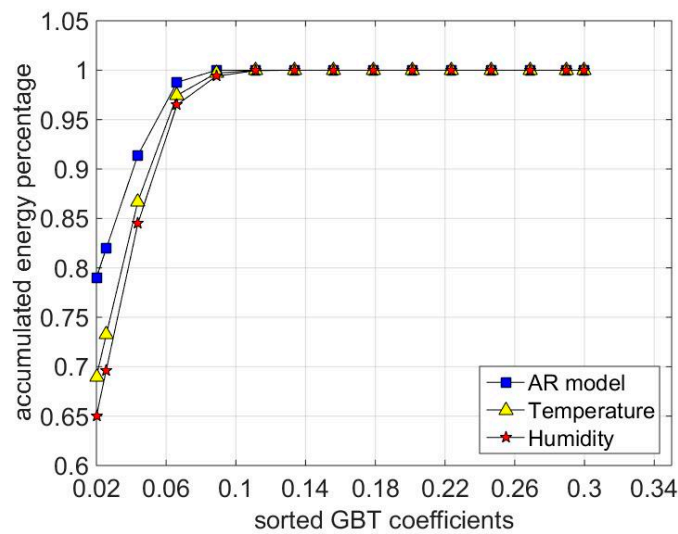


Figure 3. The sorted GBT coefficients of the datasets.

4. The Proposed Optimization Algorithm

Previous matrix completion algorithm is not realistic, as the overfitting problem cannot be avoided when the sampling rate is low. Thus, the recovery error could be large due to the overfitting problem. To obtain exact reconstruction of the missing values, the GBT sparsity and the low-rank property of the data matrix are utilized. Finally, the following optimization problem is formulated:

$$\min_X \left\{ \|X\|_* + \lambda \|\Psi^{-1}X\|_1 \right\}, \text{ subject to } \pi_\Omega(X) = \pi_\Omega(M) \tag{5}$$

where λ is the GBT sparsity regularization parameter.

ADMM [20] algorithm can blend the decomposability of dual ascent with an extra variable. Benefitted from the method of multiplier, the algorithm has superior convergence. With an introduced auxiliary variable $W \in \mathbb{R}^{M \times N}$, the above problem is rewritten as follows:

$$\min_X \left\{ \|X\|_* + \lambda \|\Psi^{-1}W\|_1 \right\}, \text{ subject to } X = W, \pi_\Omega(W) = \pi_\Omega(M) \tag{6}$$

In the following, some prerequisite properties are presented to serve as the foundation to solve the above problem.

Definition 1. Define the soft-thresholding operator is:

$$S_\varepsilon(\sigma) = \begin{cases} \sigma - \varepsilon, & \text{if } \sigma > \varepsilon, \\ \sigma + \varepsilon, & \text{if } \sigma < -\varepsilon, \\ 0, & \text{otherwise,} \end{cases} \tag{7}$$

where $\varepsilon > 0$, and the operator can be applied to vectors or matrices in an element-wise manner.

In consideration of Definition 1, the following helpful theorem is given, as stated in [13].

Theorem 1. Suppose the Singular Value Decomposition (SVD) of a matrix $\mathbf{Y} \in \mathbf{R}^{M \times N}$ is defined as $\mathbf{Y} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}$, $\mathbf{\Sigma} = \text{diag}(\{\sigma_i\}_{1 \leq i \leq \min(M,N)})$, where $\mathbf{U} \in \mathbf{R}^{M \times r}$ and $\mathbf{V} \in \mathbf{R}^{N \times r}$ are orthonormal matrix, and $r = \text{rank}(\mathbf{Y})$. Then, for any matrix $\mathbf{X} \in \mathbf{R}^{M \times N}$ and $\forall \lambda \geq 0$, the following equations are available:

$$S_\lambda(\mathbf{Y}) = \underset{\hat{\mathbf{X}} \in \mathbf{R}^{M \times N}}{\text{argmin}} \left\{ \frac{1}{2} \|\mathbf{X} - \mathbf{Y}\|_F^2 + \lambda \|\mathbf{X}\|_1 \right\} \quad (8)$$

and

$$\mathbf{U}S_\lambda(\mathbf{\Sigma})\mathbf{V}^T = \underset{\hat{\mathbf{X}} \in \mathbf{R}^{M \times N}}{\text{argmin}} \left\{ \frac{1}{2} \|\mathbf{X} - \mathbf{Y}\|_F^2 + \lambda \|\mathbf{X}\|_* \right\} \quad (9)$$

where $\mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$ is the SVD of matrix \mathbf{Y} .

Lemma 1. Suppose $\mathbf{\Psi} \in \mathbf{R}^{N \times N}$ is the unit orthogonal real matrix, and $\mathbf{\Psi}^T$ denotes the transpose matrix of $\mathbf{\Psi}$. Then, the Frobenius norm of any matrix \mathbf{A} is invariant under a unitary transformation, that is:

$$\|\mathbf{A}\mathbf{\Psi}\|_F^2 = \|\mathbf{\Psi}\mathbf{A}\|_F^2 = \|\mathbf{A}\|_F^2 \quad (10)$$

Proof. Since matrix $\mathbf{\Psi}$ is orthogonal, we have $\mathbf{\Psi}^T\mathbf{\Psi} = \mathbf{\Psi}\mathbf{\Psi}^T = \mathbf{I}_N$. The inverse of matrix $\mathbf{\Psi}$ is equal to $\mathbf{\Psi}^T$. Then, following the definition of trace, we have:

$$\|\mathbf{A}\mathbf{\Psi}\|_F^2 = \text{Tr}(\mathbf{\Psi}^T\mathbf{A}^T\mathbf{A}\mathbf{\Psi}) = \text{Tr}(\mathbf{A}^T\mathbf{A}\mathbf{\Psi}\mathbf{\Psi}^T) = \text{Tr}(\mathbf{A}^T\mathbf{A}) = \|\mathbf{A}\|_F^2 \quad (11)$$

and

$$\|\mathbf{\Psi}\mathbf{A}\|_F^2 = \text{Tr}(\mathbf{A}^T\mathbf{\Psi}^T\mathbf{\Psi}\mathbf{A}) = \text{Tr}(\mathbf{A}^T\mathbf{A}) = \|\mathbf{A}\|_F^2 \quad (12)$$

□

Then, the GBTR Regularization by Alternating Direction Method of Multipliers (GBTR-ADMM) is proposed to solve Problem (6). To get rid of the linear constraint in Problem (6), the augmented Lagrangian function is formulated as:

$$L(\mathbf{X}, \mathbf{Z}, \mathbf{W}, \beta) = \|\mathbf{X}\|_* + \lambda \|\mathbf{\Psi}^{-1}\mathbf{W}\|_1 + \langle \mathbf{Z}, \mathbf{X} - \mathbf{W} \rangle + \frac{\beta}{2} \|\mathbf{X} - \mathbf{W}\|_F^2 \quad (13)$$

where \mathbf{Z} is the Lagrange multiplier and $\beta > 0$ is the penalty parameter. For a comparative analysis, a fixed parameter β and an adaptive update strategy for optimal β are all studied in the experimental analysis section. GBTR-ADMM updates the variables in a three-step iterative approach under the constraint of fixed β . The augmented Lagrangian function $L(\mathbf{X}, \mathbf{Z}, \mathbf{W}, \beta)$ is minimized in respect of the variables in a Gauss-Seidel manner. In each step, a single variable is updated by fixing the rest of the variables. By updating the variables alternately, each subproblem is solved with a closed form solution. More specifically, the iterations of GBTR-ADMM are formulated as follows:

Firstly, the variable \mathbf{X}_{k+1} is computed with fixed value of \mathbf{Z}_k and \mathbf{W}_k . Then $L(\mathbf{X}, \mathbf{Z}, \mathbf{W}, \beta)$ is minimized as follows:

$$\begin{aligned} \mathbf{X}_{k+1} &= \underset{\mathbf{X}}{\text{argmin}} L(\mathbf{X}, \mathbf{Z}_k, \mathbf{W}_k, \beta) \\ &= \underset{\mathbf{X}}{\text{argmin}} \|\mathbf{X}\|_* + \lambda \|\mathbf{\Psi}^{-1}\mathbf{W}_k\|_1 + \langle \mathbf{Z}_k, \mathbf{X} - \mathbf{W}_k \rangle + \frac{\beta}{2} \|\mathbf{X} - \mathbf{W}_k\|_F^2 \end{aligned} \quad (14)$$

Removing the constant term in Function (14), this function can be rewritten as:

$$\begin{aligned} \mathbf{X}_{k+1} &= \underset{\mathbf{X}}{\text{argmin}} \|\mathbf{X}\|_* + \langle \mathbf{Z}_k, \mathbf{X} - \mathbf{W}_k \rangle + \frac{\beta}{2} \|\mathbf{X} - \mathbf{W}_k\|_F^2 \\ &= \underset{\mathbf{X}}{\text{argmin}} \|\mathbf{X}\|_* + \frac{\beta}{2} \|\mathbf{X} - (\mathbf{W}_k - \frac{1}{\beta}\mathbf{Z}_k)\|_F^2 \end{aligned} \quad (15)$$

Obviously, the above optimization problem has the same form as defined in Theorem 1. Thus, a closed form solution is obtained as follows:

$$\mathbf{X}_{k+1} = \mathbf{U} \mathbf{S}_{\frac{1}{\beta}}(\boldsymbol{\Sigma}_P) \mathbf{V} \tag{16}$$

where \mathbf{U} , \mathbf{V} , and $\boldsymbol{\Sigma}_P$ are obtained from the SVD of matrix P , and P equals to $\mathbf{W}_k - \frac{1}{\beta} \mathbf{Z}_k$.

Secondly, the variable \mathbf{W}_{k+1} is updated in the choice of default values \mathbf{X}_{k+1} and \mathbf{Z}_k . The minimization of $L(\mathbf{X}, \mathbf{Z}, \mathbf{W}, \beta)$ goes as follows:

$$\begin{aligned} \mathbf{W}_{k+1} &= \underset{\mathbf{W}}{\operatorname{argmin}} L(\mathbf{X}_{k+1}, \mathbf{Z}_k, \mathbf{W}, \beta) \\ &= \underset{\mathbf{W}}{\operatorname{argmin}} \|\mathbf{X}_{k+1}\|_* + \lambda \|\boldsymbol{\Psi}^{-1} \mathbf{W}\|_1 + \langle \mathbf{Z}_k, \mathbf{X}_{k+1} - \mathbf{W} \rangle + \frac{\beta}{2} \|\mathbf{X}_{k+1} - \mathbf{W}\|_F^2 \end{aligned} \tag{17}$$

Ignoring the constant term in this step, we can obtain the following optimization problem:

$$\begin{aligned} \mathbf{W}_{k+1} &= \underset{\mathbf{W}}{\operatorname{argmin}} L(\mathbf{X}_{k+1}, \mathbf{Z}_k, \mathbf{W}, \beta) \\ &= \underset{\mathbf{W}}{\operatorname{argmin}} \lambda \|\boldsymbol{\Psi}^{-1} \mathbf{W}\|_1 + \langle \mathbf{Z}_k, \mathbf{X}_{k+1} - \mathbf{W}_k \rangle + \frac{\beta}{2} \|\mathbf{X}_{k+1} - \mathbf{W}\|_F^2 \\ &= \underset{\mathbf{W}}{\operatorname{argmin}} \lambda \|\boldsymbol{\Psi}^{-1} \mathbf{W}\|_1 + \frac{\beta}{2} \|\mathbf{W} - (\mathbf{X}_{k+1} + \frac{1}{\beta} \mathbf{Z}_k)\|_F^2 \end{aligned} \tag{18}$$

Taking into consideration of the orthogonal invariance of the Forbenius norm, which is defined in Lemma 1, we obtain the following theorem.

Theorem 2. *The closed form solution of Problem (18) is defined as follows:*

$$\begin{aligned} \mathbf{W}_{k+1} &= \underset{\mathbf{W}}{\operatorname{argmin}} L(\mathbf{X}_{k+1}, \mathbf{Z}_k, \mathbf{W}, \beta) \\ &= \boldsymbol{\Psi} S_{\lambda/\beta} \left(\boldsymbol{\Psi}^{-1} \left(\mathbf{X}_{k+1} + \frac{1}{\beta} \mathbf{Z}_k \right) \right) \end{aligned} \tag{19}$$

Proof. Since matrix $\boldsymbol{\Psi}^{-1}$ is orthogonal, the following equation is obvious in combination with Lemma 1. \square

$$\frac{\beta}{2} \|\mathbf{W} - (\mathbf{X}_{k+1} + \frac{1}{\beta} \mathbf{Z}_k)\|_F^2 = \frac{\beta}{2} \|\boldsymbol{\Psi}^{-1} \mathbf{W} - \boldsymbol{\Psi}^{-1} (\mathbf{X}_{k+1} + \frac{1}{\beta} \mathbf{Z}_k)\|_F^2 \tag{20}$$

and defining $\mathbf{Q} = \boldsymbol{\Psi}^{-1} \mathbf{W}$, we have:

$$\begin{aligned} \mathbf{Q}_{k+1} &= \underset{\mathbf{Q}}{\operatorname{argmin}} L(\mathbf{X}_{k+1}, \mathbf{Z}_k, \mathbf{Q}, \beta) \\ &= \underset{\mathbf{Q}}{\operatorname{argmin}} \lambda \|\mathbf{Q}\|_1 + \frac{\beta}{2} \|\mathbf{Q} - \boldsymbol{\Psi}^{-1} (\mathbf{X}_{k+1} + \frac{1}{\beta} \mathbf{Z}_k)\|_F^2 \end{aligned} \tag{21}$$

By Theorem 1, the closed form solution of the above Problem (21) is obtained as follows:

$$\begin{aligned} \mathbf{Q}_{k+1} &= \underset{\mathbf{Q}}{\operatorname{argmin}} L(\mathbf{X}_{k+1}, \mathbf{Z}_k, \mathbf{Q}, \beta) \\ &= S_{\lambda/\beta} \left(\boldsymbol{\Psi}^{-1} \left(\mathbf{X}_{k+1} + \frac{1}{\beta} \mathbf{Z}_k \right) \right) \end{aligned} \tag{22}$$

Substituting $\mathbf{Q} = \boldsymbol{\Psi}^{-1} \mathbf{W}$ to Problem (22), then the following closed form solution is available:

$$\begin{aligned} \mathbf{W}_{k+1} &= \underset{\mathbf{W}}{\operatorname{argmin}} L(\mathbf{X}_{k+1}, \mathbf{Z}_k, \mathbf{W}, \beta) \\ &= \boldsymbol{\Psi} S_{\frac{\lambda}{\beta}} \left(\boldsymbol{\Psi}^{-1} \left(\mathbf{X}_{k+1} + \frac{1}{\beta} \mathbf{Z}_k \right) \right) \end{aligned} \tag{23}$$

In view of the second constraint term in Problem (6), the final form of W_{k+1} is defined as:

$$W_{k+1} = \pi_{\Omega}(M) + \pi_{\Omega^c}(W_{k+1}) \quad (24)$$

Thirdly, with the derived value of X_{k+1} and W_{k+1} in the above two steps, the calculation of Lagrange multiplier is updated as:

$$Z_{k+1} = Z_k + \beta(X_{k+1} - W_{k+1}) \quad (25)$$

The main procedure of GBTR-ADMM is shown in Algorithm 1. Note that the choice of penalty parameter β has high influence on the performance of ADMM algorithm. As it is difficult to choose an optimal value, the adaptive renewal mechanism is preferred in practical application. The performance difference of GBTR-ADMM with varying penalty value is studied in Section 7.1. What is more, the convergence of the ADMM based method is theoretically demonstrated in [20].

Algorithm 1: The proposed GBTR-ADMM algorithm.

Initialization: $X_1 = \pi_{\Omega}(D)$, $W_1 = X_1$, $Z_1 = X_1$, β, λ

While $\|X_{k+1} - X_k\| > \xi$ **do**

1: $X_{k+1} = US_{1/\beta}(\Sigma_P)V^T$

2: $W_{k+1} = \Psi S_{\lambda/\beta}(\Psi^{-1}(X_{k+1} - \frac{1}{\beta}Z_k))$

In consideration of the constraint in Problem (6)

$W_{k+1} = \pi_{\Omega}(M) + \pi_{\Omega^c}(W_{k+1})$

3: $Z_{k+1} = Z_k + \beta(X_{k+1} - W_{k+1})$

5. The Proposed Method for Accelerated Convergence

The performance of ADMM is highly sensitive to the number of variables and the number of constraints. As is stated in [21,22], more memory is required, and the rate of convergence is reduced with multiple variables constraints. What is more, the convergence property is not proven theoretically when the number of variables is greater than or equal to 3. In optimization Problem (6), these two constraints are considered separately, as shown in Algorithm 1. This may slow down the convergence speed.

In this section, a new approach is proposed to solve Problem (6) with fast constringency speed. Firstly, the two constraints in Problem (6) are merged together in a linear operator. Thus, the convergence rate is accelerated with only one constraint. Then, we introduce the GBT Regularization by accelerated alternating direction method of multipliers (GBTR-A2DM2) As we know, the convergence rate of A2DM2 [23,24] algorithm is $O(\frac{1}{k^2})$ while the convergence rate of ADMM (as Algorithm 1) is $O(\frac{1}{k})$.

5.1. The Fusion of Two Constraints

In consideration of the two constraints in Problem (6), $X = W$ and $\pi_{\Omega}(W) = \pi_{\Omega}(M)$, two linear operators, which are represented as \mathcal{A} and \mathcal{B} : $\mathbb{R}^{M \times N} \rightarrow \mathbb{R}^{2M \times 2N}$, are defined as follows:

$$\mathcal{A}(X) = \begin{pmatrix} X & 0 \\ 0 & 0 \end{pmatrix}, \mathcal{B}(W) = \begin{pmatrix} -W & 0 \\ 0 & \pi_{\Omega}(W) \end{pmatrix}, C = \begin{pmatrix} 0 & 0 \\ 0 & \pi_{\Omega}(M) \end{pmatrix} \quad (26)$$

where $C \in \mathbb{R}^{2M \times 2N}$ is a constant matrix.

Thus, Problem (6) is reformulated as follows:

$$\min_{X, W} \left\{ \|X\|_* + \lambda \|\Psi^{-1}W\|_1 \right\}, \text{ subject to } \mathcal{A}(X) + \mathcal{B}(W) = C \quad (27)$$

Also, the Lagrange function for the above optimization problem is:

$$L(\mathbf{X}, \mathbf{Z}, \mathbf{W}, \beta) = \|\mathbf{X}\|_* + \lambda \|\Psi^{-1}\mathbf{W}\|_1 + \langle \mathbf{Z}, \mathcal{A}(\mathbf{X}) + \mathcal{B}(\mathbf{W}) - C \rangle + \frac{\beta}{2} \|\mathcal{A}(\mathbf{X}) + \mathcal{B}(\mathbf{W}) - C\|_F^2 \quad (28)$$

Similar to Algorithm 1, GBTR-A2DM2 decomposes the minimization of $L(\mathbf{X}, \mathbf{Z}, \mathbf{W}, \beta)$ into several subproblems. In each subproblem, GBTR-A2DM2 updates a variable keeping in mind that the other variables are fixed. Specifically, the optimization scheme of GBTR-A2DM2 for Problem (28) is resolved in the following steps:

$$\begin{aligned} \mathbf{X}_{k+1} &= \underset{\mathbf{X}}{\operatorname{argmin}} L(\mathbf{X}, \mathbf{Z}_k, \mathbf{W}_k, \beta) \\ &= \underset{\mathbf{X}}{\operatorname{argmin}} \|\mathbf{X}\|_* + \langle \mathbf{Z}_k, \mathcal{A}(\mathbf{X}) + \mathcal{B}(\mathbf{W}_k) - C \rangle + \frac{\beta}{2} \|\mathcal{A}(\mathbf{X}) + \mathcal{B}(\mathbf{W}_k) - C\|_F^2 \\ &= \frac{\beta}{2} \|\mathcal{A}(\mathbf{X}) + \mathcal{B}(\mathbf{W}) - C + \frac{1}{\beta} \mathbf{Z}_k\|_F^2 \end{aligned} \quad (29)$$

$$\begin{aligned} \mathbf{W}_{k+1} &= \underset{\mathbf{W}}{\operatorname{argmin}} L(\mathbf{X}_{k+1}, \mathbf{Z}_k, \mathbf{W}, \beta) \\ &= \underset{\mathbf{W}}{\operatorname{argmin}} \lambda \|\mathbf{G}^{-1}\mathbf{W}\|_1 + \langle \mathbf{Z}_k, \mathcal{A}(\mathbf{X}_{k+1}) + \mathcal{B}(\mathbf{W}) - C \rangle + \frac{\beta}{2} \|\mathcal{A}(\mathbf{X}_{k+1}) + \mathcal{B}(\mathbf{W}) - C\|_F^2 \\ &= \underset{\mathbf{W}}{\operatorname{argmin}} \lambda \|\mathbf{G}^{-1}\mathbf{W}\|_1 + \frac{\beta}{2} \|\mathcal{A}(\mathbf{X}_{k+1}) + \mathcal{B}(\mathbf{W}) - C + \frac{1}{\beta} \mathbf{Z}_k\|_F^2 \end{aligned} \quad (30)$$

$$\mathbf{Z}_{k+1} = \mathbf{Z}_k + \beta (\mathcal{A}(\mathbf{X}_{k+1}) + \mathcal{B}(\mathbf{W}_{k+1}) - C) \quad (31)$$

The pseudocode of GBTR-A2DM2 algorithm is shown in Algorithm 2. Next, we will discuss the accelerated technique of GBTR-A2DM2 with a restarting rule.

Algorithm 2: GBTR-A2DM2 algorithm using restarting rule.

Initialization

$\mathbf{W}_0 = \hat{\mathbf{W}}_0 \in \mathbb{R}^{M \times N}, \mathbf{Z}_0 = \hat{\mathbf{Z}}_0 \in \mathbb{R}^{M \times N}, \tau > 0, a_0 = 1, \eta = 0.999$

While $\|\mathbf{X}_{k+1} - \mathbf{X}_k\| > \zeta$ **do**

1: Update \mathbf{X}_k by Equation (29)

2: Update \mathbf{W}_k by Equation (30)

3: Update \mathbf{Z}_k by Equation (31)

4: $m_k \equiv \frac{1}{\tau} \|\mathbf{Z}_k - \hat{\mathbf{Z}}_k\|_F^2 + \tau \|\mathcal{B}(\mathbf{W}_k - \hat{\mathbf{W}}_k)\|_F^2$

If $m_k < \eta m_{k-1}$ **Then**

5: $a_{k+1} = \frac{1 + \sqrt{1 + 4a_k^2}}{2}$

6: $\hat{\mathbf{W}}_{k+1} = \mathbf{W}_k + \frac{a_k - 1}{a_{k+1}} (\mathbf{W}_k - \mathbf{W}_{k-1})$

7: $\hat{\mathbf{Z}}_{k+1} = \mathbf{Z}_k + \frac{a_k - 1}{a_{k+1}} (\mathbf{Z}_k - \mathbf{Z}_{k-1})$

Else

8: $a_{k+1} = 1, \hat{\mathbf{W}}_{k+1} = \mathbf{W}_k, \hat{\mathbf{Z}}_{k+1} = \mathbf{Z}_k$

9: $m_k = \eta^{-1} m_{k-1}$

End if

End While

5.2. The Accelerated Technique

Since the objective function in Problem (27) is not very convex, the accelerated ADMM method with a restart rule is employed. To determine when to restart the value assignment, the primal error and the dual error are combined:

$$m_k \equiv \frac{1}{\tau} \|\mathbf{Z}_k - \hat{\mathbf{Z}}_k\|_F^2 + \tau \|\mathcal{B}(\mathbf{W}_k - \hat{\mathbf{W}}_k)\|_F^2 \quad (32)$$

where $\hat{\mathbf{Z}}_k$ and $\hat{\mathbf{W}}_k$ represent the second updated step in iteration steps 6–7 of Algorithm 2. For each iteration, m_k is compared with m_{k-1} and if $m_k < \eta m_{k-1}$, where η is defined equal to 0.999, the algorithm is accelerated with steps 5–7. Otherwise, the method is restarted in process of steps 8–9. In comparison with GBTR-ADMM, Algorithm 2 has a higher convergence rate. Also, the convergence property of Algorithm 2 is guaranteed by A2DM2 with a restarting rule [23].

6. Time Complexity Analysis

In this part, the computational complexity of the proposed algorithms is discussed. The calculation of an inverse matrix cost much, which has the time complexity of $O(n^3)$ (n is the dimension of an invertible matrix). Since matrix $\mathbf{\Psi}$ is orthogonal, the expensive computation of matrix inversion in our implementation can be substituted by its transposition. Thus, the dominated computational cost of GBTR-ADMM and GBTR-A2DM2 is the execution of matrix SVD in each iteration. As pointed out in [25], the time complexity of SVD operation is $O(MN^2)$. In our implementation, the famous PROPACK [26] is utilized to perform partial SVD for the proposed algorithms. Since the low-rank property of the objective matrix, it is inefficient to compute the full SVD. To obtain the dominated energy of the objective matrix, only those singular values exceeding than a certain threshold are necessary. The limitation of PROPACK is that it cannot automatically determine the necessary calculations, except for a predefined number. Thus we are supposed to estimate the number of singular values and assign the number to PROPACK in each iteration.

Suppose svp^k is the number of positive singular values of \mathbf{X}_k , and sv^k is the number of singular value to be measured at k -th iteration. Then, the following updated strategy [27] is used,

$$sv^{k+1} = \begin{cases} svp^k + 1, & \text{if } svp^k < sv^k \\ svp^k + 5, & \text{if } svp^k = sv^k \end{cases} \quad (33)$$

where the initial estimated value of sv^0 is 10. Benefiting from the software package PROPACK, the time complexity for a $M \times N$ matrix with rank of r is $O(rMN)$. Hence, the total time complexity of our proposed algorithms is $O(rMN)$. Nevertheless, the state of art algorithms for matrix completion problem [15,17] demand a complexity of $O(r^2MN)$ for each iteration.

7. Performance Evaluation

In this section, we evaluate the performances of GBTR-ADMM and GBTR-A2DM2. The experimental datasets and their topological structures are described in Section 3. Since the proposed algorithms are heavily influenced by several input parameters, it is necessary to choose the optimal parameters to maximize the algorithm performance. With the optimal parameters for GBTR-ADMM and GBTR-A2DM2, the recovery accuracy and the convergence properties are compared with the state of art algorithms. At last, the energy consumption of the proposed algorithms are compared with the state of art data gathering methods for WSNs. Simulation results show that GBTR-ADMM and GBTR-A2DM2 can highly reduce energy consumption in WSNs. Thus, the network lifetime is prolonged.

To measure the performance of the proposed algorithm, the reconstructed data matrix $\hat{\mathbf{X}}$ is achieved. Thus, the recovery performance is estimated by the Normalized Mean Absolute Error (NMAE):

$$NMAE = \frac{\sum_{(i,j) \in \Omega^c} |(\hat{\mathbf{X}}(i,j) - \mathbf{X}(i,j))|}{\sum_{(i,j) \in \Omega^c} |\Omega^c(i,j)|} \quad (34)$$

7.1. Parameter Setting

In this subsection, the choice of optimal parameters for GBTR-ADMM is discussed. Previous studies have shown that global convergence for ADMM algorithm holds for any fixed $\beta > 0$. However, different parameter values result in various convergence speeds. Thus, the input values, as listed in

Table 3, to Algorithm 1 are selected by experience to obtain the best performance. The performance of GBTR-ADMM is also studied with different parameter values of β . The variation of the objective function values of Problem (13) with the increase of iteration numbers is shown in Figure 4. As we can see, $\beta = \beta_0$ achieves the best performance. Meanwhile, the descending speed becomes slower when the choice of parameter β is too large or too small. This is because the penalty parameter β trades off between minimizing the primal residual and the residual of the dual problem. A large penalty value may drop the primal residual, but at the expense of an increase of the dual residual, and vice versa.

Table 3. Input values to Algorithm 1.

Parameter Name	λ	β
Set Value	0.01	$\beta_0 = 3.0/\min(m, n)$

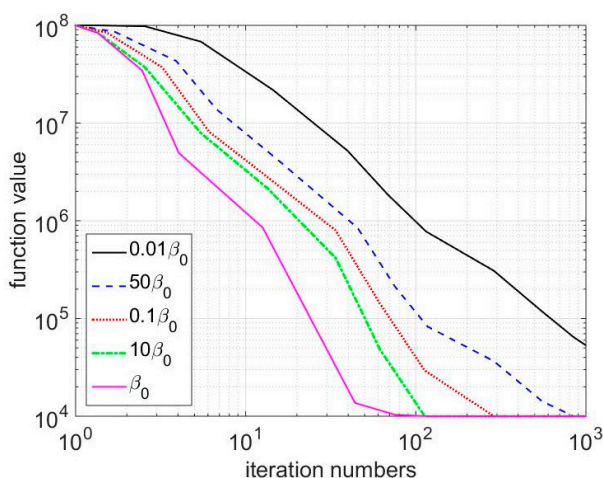


Figure 4. The performance of GBTR-ADMM in respect to different β .

Figure 4 demonstrates the results in the synthesized datasets, and the optimal chosen value of β changes randomly in other datasets. Instead, an adaptive penalty update method is preferred, which is based on previous study [20]. The update strategy is formulated as follows:

$$\beta_{k+1} = \min(\beta_{\max}, \rho\beta_k), \tag{35}$$

where β_{\max} denotes the maximum value of the β_k . The value of variable ρ is updated as:

$$\rho = \begin{cases} \rho_0, & \text{if } \frac{\beta_k \max\{\|X_{k+1} - X_k\|_F^2, \|W_{k+1} - W_k\|_F^2\}}{\|C\|_F} < \kappa \\ 1, & \text{otherwise} \end{cases} \tag{36}$$

where $\rho_0 > 0$ is a constant and κ is the predefined threshold value. Obviously, when the residual value between $\|X_{k+1} - X_k\|_F^2$ and $\|W_{k+1} - W_k\|_F^2$ is less than the threshold, the value of β_{k+1} increases to $\rho_0\beta_k$. Thus, the convergence speed is improved in this way.

The effect of the sparsity regularization parameter λ is also analyzed. Figure 5 shows the variation of recovery error with different parameter value. As can be seen, the recovery error is quite large with small value of λ . With the increase of λ , recovery error declines rapidly, and remains stable as $\lambda > 0.01$. So, the optimal value for the sparsity regularization parameter λ is set as 0.01 in our experiments. Since similar trends are obtained for GBTR-A2DM2, we just omit it here.

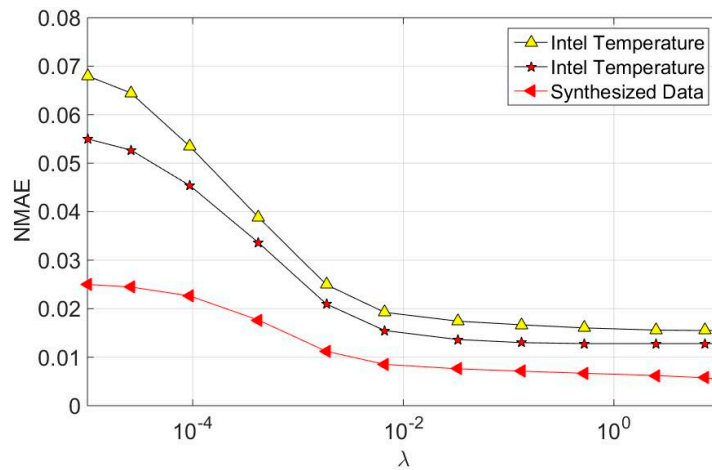


Figure 5. The effect of the sparsity regularization parameter λ .

7.2. Recovery Accuracy

In this subsection, we compare the recovery accuracy of GBTR-ADMM and GBTR-A2DM2 with the state of art algorithms for matrix completion. The first chosen method is the Spatio-Temporal Compressive Data Collection (STCDG) [15]. The second method is the Compressive Data Collection (CDC) [11]. GBTR-ADMM is considered to solve the optimization Problem (6), while GBTR-A2DM2 is used for the optimization Problem (27) with only one constraint.

Simulations are executed on both the real datasets and the synthesized dataset, which are exploited in detail in Section 3. For each parameter setting of the simulation, the results are averaged for 50 independent trials.

Figures 6–8 show that our GBTR based methods can reconstruct the missing values with high accuracy. In general, the recovery errors of all reconstruction algorithms decrease rapidly with the increase of the sampling ratio. When the sampling ratio is high enough, all reconstruction methods achieve smaller recovery error. Since our proposed two GBTR based methods are used to solve the same matrix completion problems, their performances are nearly the same. Figure 6 shows the recovery errors on GreenOrbs temperature data. As can be seen, our proposed methods achieve about 25% recovery error while the error of other two algorithms is more than 80%, when the sampling ratio is 1%.

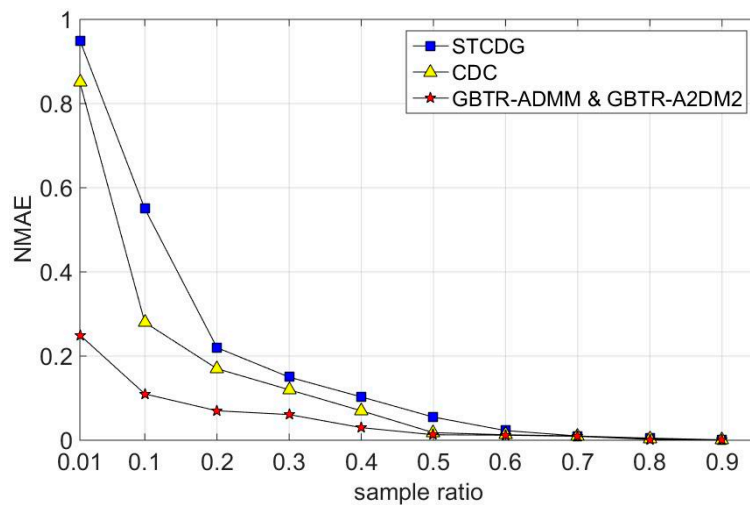


Figure 6. Recovery errors on temperature dataset.

Similar results can be obtained from Figure 7, which is simulated on the humidity dataset. The recovery error of GBTR based methods is still much less than STCDG and CDC when the sampling ratio is small. When the sampling is 1%, GBTR-ADMM and GBTR-A2DM2 can reconstruct the original missing values with recovery errors of less than 20%. Meanwhile, the recovery errors of STCDG and CDC are nearly 100%.

In Figure 8, the experiment results show that GBTR based methods outperform STCDG and CDC by a larger margin. Compared with Figures 6 and 7, Figure 8 shows that GBTR-ADMM and GBTR-A2DM2 achieve the best performance on the synthesized dataset. The recovery error on synthesized dataset is smaller than the real datasets at the same sampling ratio. The reason is that the synthesized data have much better sparsity than the other two real datasets under the GBT basis.

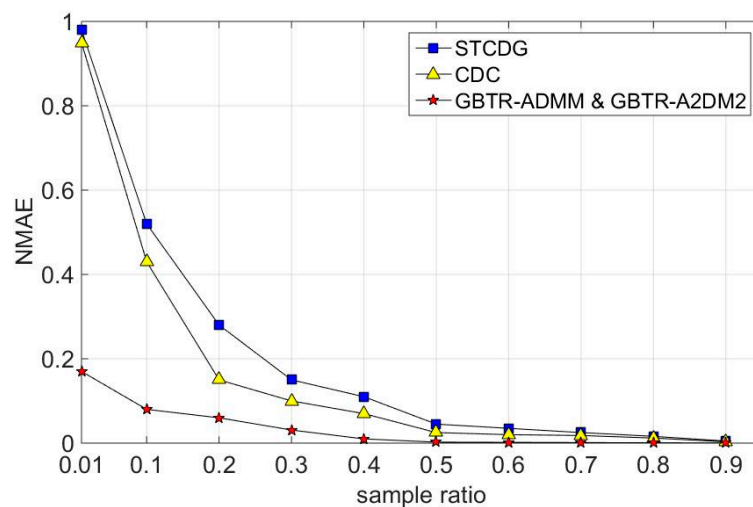


Figure 7. Recovery errors on humidity dataset.

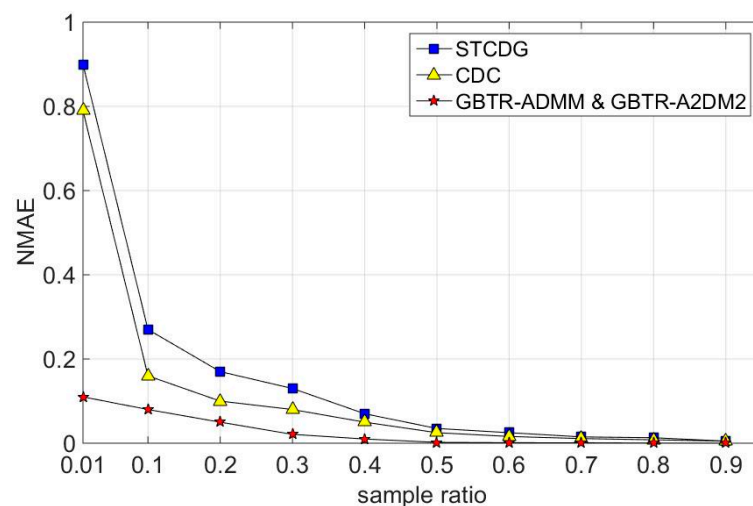


Figure 8. Recovery errors in the synthesized dataset.

7.3. Convergence Behavior Analysis

In this subsection, the convergence performances are studied in the synthesized dataset. The compared methods are SVD and STCDG. For each method, we set the same stop conditions, where the tolerance error ζ is 10^{-4} . Figure 9 shows the necessary number of iterations to obtain accurate reconstruction at different sampling ratios. As can be seen from Figure 9, the convergence speed of our proposed two methods surpasses the SVD and STCDG. SVD has the slowest convergence

rate of the four methods. Also, STCDG converges faster than SVD. Although the recovery accuracy of GBTR-A2DM2 and GBTR-ADMM behaves similar, as shown in Figures 6–8, GBTR-A2DM2 converges much faster than GBTR-ADMM. Note that GBTR-ADMM needs nearly 160 iterations to converge at the sampling ratio 0.9, while only about 40 iterations leads to the convergence of GBTR-A2DM2. Even when the sampling ratio is 0.5, the necessary number of iterations for GBTR-A2DM2 is about 125, which is less than half of GBTR-ADMM.

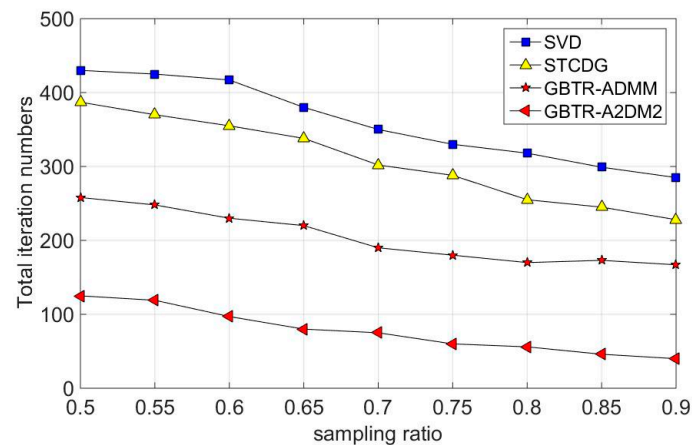


Figure 9. Necessary number of iterations for different algorithms.

Next, when the sampling ratio is fixed as 0.6, the relative recovery errors of all the compared methods are analyzed. As can be seen from Figure 10, Compared to SVD and STCDG, both GBTR-ADMM and GBTR-A2DM2 gain less error in several iteration processes. Clearly, GBTR-A2DM2 converges much faster, which converges in about 100 iterations. Also, with no more than 250 iterations, both GBTR-ADMM and GBTR-A2DM2 terminate as the relative recovery errors drop below the tolerance error. Meanwhile, the relative errors of SVD and STCDG are about one order of magnitude larger than GBTR-ADMM and GBTR-A2DM2. In general, compared with the state of the art methods for matrix completion, our proposed methods can achieve smaller recovery error at the same number of iterations.

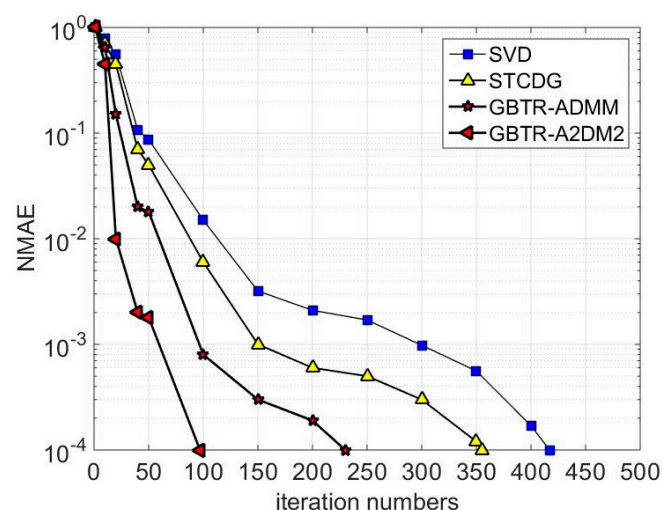


Figure 10. Variation of recovery errors in respect to iteration numbers for different algorithms.

7.4. Energy Consumption and Network Lifetime

In this subsection, the energy consumption of the proposed algorithms for data gathering problems is analyzed. Five nodes are randomly deployed in a 1000 m × 1000 m area. The topology is shown in Figure 2, where the sink node is deployed in the center. The data transmission and recovery process are fulfilled in three steps. First, the sink node broadcasts the sampling ratio through the whole network. In our setting, the sampling ratio determines the probability of node to gather data. In the second step, the selected nodes transmit the gathering data to the sink node. Finally, these missing data are reconstructed by implementing GBTR-ADMM and GBTR-A2DM2 at the sink node. The compared methods are CDC and STCDG. In the traditional data gathering method, all sensor nodes are required to transmit their sampling data to the sink node. Thus, the traditional method is selected as a baseline method for comparison.

The energy consumption model in paper [28] is employed in our simulation. The detailed simulation parameters are presented in Table 4. The initial energy for every sensor node is 2 J. Each packet contains 64 bits. The network is supposed to be symmetrical. The energy consumption for one bit transmission, defined as E_{Tx} , is 100 nJ. Meanwhile, the reception of a packet consumes $E_{Rx} = 120$ nJ. E_{Amp} is the unit energy consumption of the power amplifying circuit.

Table 4. Experimental parameters.

Parameter Name	Value
Nodes number	500
Transmission range	100 m
Initial energy	2 J
Data Size	64 bits
E_{Tx}	100 nJ/bit
E_{Rx}	120 nJ/bit
E_{Amp}	0.1 nJ/(bit·m ²)

The synthesized dataset is exploited in our experiment. The network lifetimes of CDC, STCDG, and our proposed methods are evaluated. Detailed information is revealed in Figure 11. Note that the total energy consumption of the baseline method is relatively constant. That is because the baseline method transmits all the sensor data to the sink node no matter how the sampling ratio varies. To be different, the total power consumptions of CDC, STCDG, and GBTR based methods increase with enlargement of the sampling ratio. The reason is that more sensor data are needed to be transmitted when the sampling ratio increases. Thus, the network lifetime decreases. GBTR-ADMM and GBTR-A2DM2 outperform CDC at the same setting value of sampling ratio. However, the energy consumption of our proposed methods is equal to the baseline method when the sampling ratio is exactly 1. Note that the lifetime of CDC is smaller than the baseline when the sampling ratios are higher than 75%. This phenomenon could be explained as per below. In CDC, all sensor nodes transmit the data M times, which is the necessary number to reconstruct original signals. At the same time, the ordinary nodes need only one data transmission and necessary relay transmissions for each sampling in the baseline method. Thus, when the sampling ratio is higher than a specific threshold, the total numbers of transmission for CDC is larger than that in the baseline method. In addition, since STCDG and our proposed methods are all based on the matrix completion theory, the curve variations over their lifetime coincide with each other.

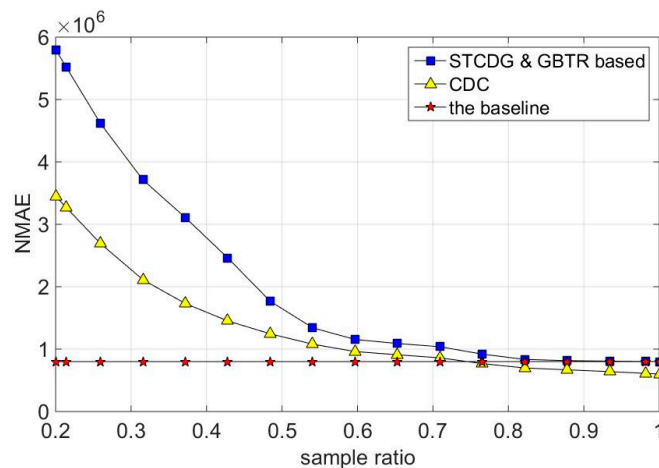


Figure 11. Network lifetime comparison.

8. Conclusions and Future Works

In this paper, the data gathering problem based on Matrix Completion theory is studied. Except for the low-rank property, the sensed data are observed to be sparse under the graph based transform. By taking full advantage of these features, two novel reconstruction algorithms (named GBTR-ADMM and GBTR-A2DM2) are proposed. The time complexity is also analyzed, which shows their complexity is low. Several experiments on both real datasets and synthesized datasets are carried out. The experiment results show that our proposed methods outperform the state of the art algorithm for data gathering problems in WSNs. Furthermore, it is observed that GBTR-A2DM2 converges much faster than GBTR-ADMM. For future works, will focus on applying our proposed algorithms to other datasets of real networks, which may exhibit complex topological information other than random networks, such as the scale-free or the small-world networks.

Acknowledgments: This work is supported in part by the National Natural Science Foundation of China under Grant No. 61371135 and by Beihang University Innovation & Practice Fund for Graduate under Grant YCSJ-02-2016-04. The authors are thankful to the anonymous reviewers for their earnest reviews and helpful suggestions.

Author Contributions: Donghao Wang made deduced the original optimization methods, and implemented detailed algorithm design. Jiangwen Wan revised the manuscript. Zhipeng Nie performed the experiments. Qiang Zhang and Zhijie Fei analyzed the data.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Yoon, S.; Shahabi, C. The Clustered AGgregation (CAG) technique leveraging spatial and temporal correlations in wireless sensor networks. *ACM Trans. Sens. Netw.* **2007**, *3*, 3. [[CrossRef](#)]
2. Pham, N.D.; Le, T.D.; Park, K.; Choo, H. SCCS: Spatiotemporal clustering and compressing schemes for efficient data collection applications in WSNs. *Int. J. Commun. Syst.* **2010**, *23*, 1311–1333. [[CrossRef](#)]
3. Lachowski, R.; Pellenz, M.E.; Penna, M.C.; Jamhour, E.; Souza, R.D. An efficient distributed algorithm for constructing spanning trees in wireless sensor networks. *Sensors* **2015**, *15*, 1518–1536. [[CrossRef](#)] [[PubMed](#)]
4. Candes, E.J.; Romberg, J.; Tao, T. Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. *IEEE Trans. Inf. Theory* **2006**, *52*, 489–509. [[CrossRef](#)]
5. Donoho, D.L. Compressed sensing. *IEEE Trans. Inf. Theory* **2006**, *52*, 1289–1306. [[CrossRef](#)]
6. Gleichman, S.; Eldar, Y.C. Blind compressed sensing. *IEEE Trans. Inf. Theory* **2011**, *57*, 6958–6975. [[CrossRef](#)]
7. Li, S.X.; Gao, F.; Ge, G.N.; Zhang, S.Y. Deterministic construction of compressed sensing matrices via algebraic curves. *IEEE Trans. Inf. Theory* **2012**, *58*, 5035–5041. [[CrossRef](#)]

8. Luo, C.; Wu, F.; Sun, J.; Chen, C.W. Compressive data gathering for large-scale wireless sensor networks. In Proceedings of the 15th ACM International Conference on Mobile Computing and Networking, Beijing, China, 20–25 September 2009; pp. 145–156.
9. Caione, C.; Brunelli, D.; Benini, L. Distributed compressive sampling for lifetime optimization in dense wireless sensor networks. *IEEE Trans. Ind. Inf.* **2012**, *8*, 30–40. [[CrossRef](#)]
10. Xiang, L.; Luo, J.; Rosenberg, C. Compressed data aggregation: Energy-efficient and high-fidelity data collection. *IEEE ACM Trans. Netw.* **2013**, *21*, 1722–1735. [[CrossRef](#)]
11. Liu, X.Y.; Zhu, Y.; Kong, L.; Liu, C.; Gu, Y.; Vasilakos, A.V.; Wu, M.Y. CDC: Compressive data collection for wireless sensor networks. *IEEE Trans. Parallel Distrib. Syst.* **2015**, *26*, 2188–2197. [[CrossRef](#)]
12. Candes, E.J.; Recht, B. Exact matrix completion via convex optimization. *Found. Comput. Math.* **2009**, *9*, 717–772. [[CrossRef](#)]
13. Cai, J.F.; Candes, E.J.; Shen, Z.W. A singular value thresholding algorithm for matrix completion. *SIAM J. Optim.* **2010**, *20*, 1956–1982. [[CrossRef](#)]
14. Roughan, M.; Zhang, Y.; Willinger, W.; Qiu, L.L. Spatio-temporal compressive sensing and internet traffic matrices. *IEEE ACM Trans. Netw.* **2012**, *20*, 662–676. [[CrossRef](#)]
15. Cheng, J.; Ye, Q.; Jiang, H.; Wang, D.; Wang, C. STCDG: An efficient data gathering algorithm based on matrix completion for wireless sensor networks. *IEEE Trans. Wirel. Commun.* **2013**, *12*, 850–861. [[CrossRef](#)]
16. Liu, Y.; He, Y.; Li, M.; Wang, J.; Liu, K.; Li, X. Does wireless sensor network scale? A measurement study on GreenOrbs. *IEEE Trans. Parallel Distrib. Syst.* **2013**, *24*, 1983–1993. [[CrossRef](#)]
17. Kong, L.; Xia, M.; Liu, X.Y.; Chen, G.; Gu, Y.; Wu, M.Y.; Liu, X. Data loss and reconstruction in wireless sensor networks. *IEEE Trans. Parallel Distrib. Syst.* **2014**, *25*, 2818–2828. [[CrossRef](#)]
18. Candes, E.; Recht, B. Exact matrix completion via convex optimization. *Commun. ACM* **2012**, *55*, 111–119. [[CrossRef](#)]
19. Shuman, D.I.; Narang, S.K.; Frossard, P.; Ortega, A.; Vandergheynst, P. The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains. *IEEE Signal Process. Mag.* **2013**, *30*, 83–98. [[CrossRef](#)]
20. Boyd, S.; Parikh, N.; Chu, E.; Peleato, B.; Eckstein, J. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Found. Trends Mach. Learn.* **2011**, *3*, 1–122. [[CrossRef](#)]
21. He, B.; Tao, M.; Yuan, X. Alternating direction method with Gaussian back substitution for separable convex programming. *SIAM J. Optim.* **2012**, *22*, 313–340. [[CrossRef](#)]
22. Hu, Y.; Zhang, D.; Ye, J.; Li, X.; He, X. Fast and accurate matrix completion via truncated nuclear norm regularization. *IEEE Trans. Pattern Anal. Mach. Intell.* **2013**, *35*, 2117–2130. [[CrossRef](#)] [[PubMed](#)]
23. Goldstein, T.; O'Donoghue, B.; Setzer, S.; Baraniuk, R. Fast alternating direction optimization methods. *SIAM J. Imaging Sci.* **2014**, *7*, 1588–1623. [[CrossRef](#)]
24. Kadkhodaie, M.; Christakopoulou, K.; Sanjabi, M.; Banerjee, A. Accelerated alternating direction method of multipliers. In Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Sydney, Australia, 10–13 August 2015; pp. 497–506.
25. Golub, G.H.; Van Loan, C.F. *Matrix Computations*; JHU Press: Baltimore, MD, USA, 2012.
26. Larsen, R.M. PROPACK-Software for Large and Sparse SVD Calculations. Available online: <http://sun.stanford.edu/~rmunk/PROPACK> (accessed on 19 September 2016).
27. Toh, K.C.; Yun, S. An accelerated proximal gradient algorithm for nuclear norm regularized linear least squares problems. *Pac. J. Optim.* **2010**, *6*, 615–640.
28. Heinzelman, W.R.; Chandrakasan, A.; Balakrishnan, H. Energy-efficient communication protocol for wireless microsensor networks. In Proceedings of the 33rd Annual Hawaii International Conference on System Sciences, Maui, HI, USA, 4–7 January 2000; p. 223.

