Compressed Topology Tomography in Sensor Networks

Chunnan Yao
School of Electronic Information and
Electrical Engineering
Shanghai Jiao Tong University
Shanghai, China
Email: yaochunnan@sjtu.edu.cn

***Abstract*—Wireless sensor network(WSN) topology tomography is the kernel of routing maintenance, topology control, anomaly detection and load balance. As static routing tree estimation has been fully studied in recent years, WSN topology tomography has been proven non-trivial due to its instability. In this work, we study general WSN routing topology tomography from information gathered at the sink node only, where routing structure is dynamic. We first formulate the problem using a routing matrix. From the formulation, we divide the solution into two steps. First, we propose a refinement of Multi-hop Network Tomography(MNT) using compressive sensing to reconstruct the routing matrix. Second, we prove that for a routing matrix of a WSN which is very large and dynamic, link parameters can be reconstructed using l_1-minimization method with a high probability.**

***Index Terms*—wireless sensor networks, network tomography, compressive sensing, delay estimation, expander graphs**

I. Introduction

Wireless sensor networks(WSNs) are growing rapidly in both size and complexity because of wide deployments. Consequently, it becomes increasingly important to monitor the WSN structure and dynamics, namely, network tomography. This capability plays a significant role in routing improvement, topology control, anomaly detection and load balance. Although there have been lots of works on link estimation in traditional networks[4][3], studies of WSN topology tomography aimed to recover the dynamic routing structure and link parameters are limited. Our work is thus motivated by the needs to overcome the fragility of traditional network tomography methods in wireless scenarios.

In this paper, we divide the problem into two steps. First, we should reconstruct the per-packet routing path serves as the meta-information for later reconstruction of link parameters. There are lots of works dedicated to solve this problem.The method that identifies a certain path using heuristic path scanning triggers disastrous computation overhead[2]. Some methods reconstruct paths exploiting packet correlations[4]. However,experiments and analysis have confirmed that, in real WSNs, we observe non-negligible topology variation and instability. They lead to inaccuracy and even failure of some state-of-art methods. To solve these issues, we propose a method to use compressive sensing to optimize a state-of-art method: MNT(multi-hop network tomography), namely, MNT-CS method. We use simulations to verify that our algorithm is effective and practical. Statistics will show that loss rate, delay and reconstruction rate are all improved.

Second, once that topology is reconstructed, we shall use compressive sensing the second time to reconstruct the inner-link parameters such as delay. We consider the network to be dynamic and large, thus the routing matrix is very sparse. Using conclusions from[3] and knowledge of permutation and combination, we verified that for a random large WSN, the routing matrix generated by step1 is qualified to be the sensing matrix for step2 to reconstruct the inner-link parameters with a high probability. The total process can be briefly displayed in Figure 1.

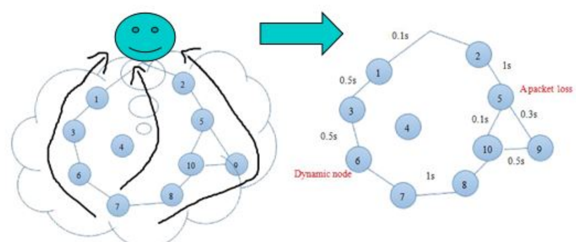

Fig. 1. WSN tomography illustration

The rest of this paper is organized as follows. Section II gives introduction on previous related works and their relationship with this work. Section III briefly gives the network routing model adopted in this work. Section IV represents the problem in a compressive sensing way. Section V proposes our optimization of state-of-art method MNT. Section VI verifies the effectiveness of l_1-minimization method in random large dynamic WSNs.It also gives an algorithm chart of our method. Section VII reports our simulations. Finally, Section VIII gives conclusions and outlines our future work.

II. Previous Related work

According to our knowledge, previous works related to WSN tomography can be divided into 3 classes.

The first class, which is classical, concerns about exploiting existing information in the network or add functional headers to each packet to collect inter-hop information. Some state-of-art algorithms fall into this class, such as MNT(multi-hop network tomography)[4], Pathfinder[7], Pathzip[8], ACO[9], etc. These methods are comparatively easy to deploy, but have their own constraints in scale or stability of the network.

The second class introduces compressive sensing to WSN tomography in an innovative way. Compressive sensing, which we will give a brief introduction to later, has been a new research field in recent ten years . Although intended to improve the efficiency of digital image processing, compressive sensing is also introduced in mobile communications for its flexibility and economical spacial performance. Lots of works concern about network tomography using compressive sensing[10][11][12]. Compressive sensing should be deployed in different ways according to different scenarios. But one thing is in common, we should find the sparsity in a specific scenario and use it to construct a compressive sensing model which is compatible with reconstruction algorithms.

The third class looks into the features of compressive sensing on graphs, and use graph theory and other mathematical tools to advance it. Typical ones of them concern about expander graphs[3].According to their conclusion, compressive sensing on expander graphs can exploit a more sparse sensing matrix and more efficient reconstruction methods. These are all preferred in resource-limited and time-limited scenes.

However, according to our knowledge, there are few works combining the above three kinds of work together. In fact, WSN tomography requires contribution of all the tree kinds of work, and through good combination, hopefully we can exploit the potential of previous related works. This work seeks to use state-of-art reconstruction method, compressive sensing and expander graph to achieve the goal of WSN tomography with high probability.

III. Routing Model

Assume that a given network N(V,E) has a total of n links(i.e., n=$|E|$), and γ(waiting to be reconstructed) is the set of paths between the source nodes of the network to the sink node and m = $|\gamma|$. Let $\gamma_{m\times n}$ denote the routing matrix, where there exists an isomorphism between the set γ and the corresponding routing matrix $\mathbf{R}$. For example, for the network in Figure 2, we have the routing matrix:

$$\mathbf{R} = \begin{array}{r} \\ P_1 : n_2->n_6->sink \\ P_2 : n_1->n_5->sink \\ P_3 : n_1->n_2->sink \\ P_4 : n_5->n_6->sink \end{array} \begin{array}{c} \begin{array}{ccccc} l_2 & l_2 & l_3 & l_4 & l_5 \end{array} \\ \left|\begin{array}{ccccc} 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{array}\right| \end{array}$$

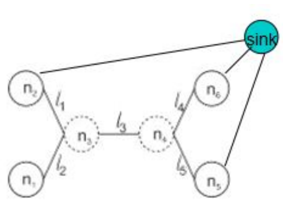

Fig. 2. packet routing model

For parameters such as delay, an addictive linear model adequately represents the relationship between a measured path and an individual link delay, i.e.,

$$y = \mathbf{R}x,$$

where x is the $n \times 1$(unknown) vector of the individual link mean delay. y is the measured r-vector of end-to-end path delays. We should notice that although this paper focuses on link delay, our method can be applied to any other link attributes(such as log of packet loss rate) which allows such a linear relationship with end-to-end measurements. For most cases, since we assume the WSN is very large, number of interval links of each packet is much more less than the total number of links in this network, so we assume that n≫ m. This enables us to view the problem of reconstructing x in a compressive sensing perspective.

IV. Formulation

We consider the entire dynamic routing topology G(V,E) to be reconstructed with routing matrix of $E = p_1 \cup p_2 \cup \cdots \cup p_M(M = n-1)$, where p_i is the path measurement of the i^{th} packet. The routing matrix is defined just as mentioned above:

$$\varphi_{ij} = \begin{cases} 1 & \text{the } i^{th} \text{ path traverses over the } j^{th} \text{ link} \\ 0 & \text{otherwise} \end{cases}$$

To adopt to the WSN scenario,where the routing topology structure is evolving along time, the total number of links n is the accumulation of all its dynamic possibilities in the total measurement period. This leads to a more sparse routing matrix. If we assume $\phi = \mathbf{R}$, the problem can be expressed in a compressive sensing way:

$$X = \text{argmin}\|X\|_0 \text{subject to} Y = \phi X$$

where l_0-norm $\|X\|_0$ is the number of nonzero elements in the vector X, $\phi_{m\times n}$ is a sparse sensing matrix and n≫m.X is a sparse vector of length n. Y is the sink node information of length m. X has 2 reasons to be sparse: 1. Inner-link parameters such as delay can only be large when congestion or other anomaly happens, which is not frequent. 2. As has mentioned above, the number of interval links of each packet is much more less than the total number of links in this network.

Compressive sensing is widely studied these years, many algorithms have been brought out to solve this NP problem. The most famous one is called l_1 -minimization. But the sensing matrix ϕ should satisfy RIP constraint to guarantee the reconstruction accuracy. When applied to networks, the routing matrix must satisfy other constraints, as explained in [3].

We know that in lots of previous related works, sensing matrix ϕ is known before the reconstruction process. Unfortunately, the ϕ in our problem formulation is completely unknown which requires path reconstruction. Only if all the paths are reconstructed and routing matrix(sensing matrix) ϕ is known, can we further reconstruct the inner-link parameters. So we can separate our algorithm into 2 steps:step1, path reconstruction;step2, reconstruction of link parameters.

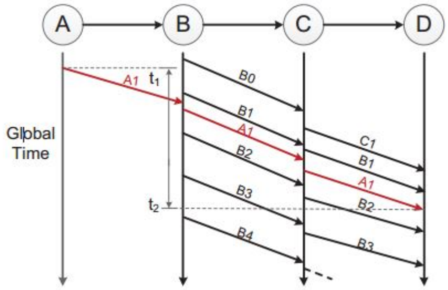

Fig. 3. A motivating example illustrating how MNT reconstructs the routing path of a packet(A,B,1). The packet is also denoted as A1 in the figure for short.

V. Path Reconstruction

The problem of path reconstruction can be viewed as a single topic in WSN tomography. A straightforward solution to reveal the packet path is to record the complete path during packet forwarding. But the introduced overhead linearly grows with the path length, which is unacceptable in large scale WSNs.

There have been many efforts made to address the path reconstruction problem in WSNs. One of them is state-of-art MNT(multi-hop network tomography)[4]. MNT exploits data already attached to each packet:(origin ID, parent ID, sequence number). Using time estimation and this tuple, the path of each packet can be reconstructed from the sink node hop by hop. Figure 3 shows a typical working process of MNT.

MNT infers the next hop of the parent hop by leveraging the parent hop of the other two packets generated at A's parent hop B. It assumes that all the three packets can be received by sink node D and the topology is so stable that B1 and B2 have the same parent node. By repeating the process, the path of A can be reconstructed hop by hop. Although MNT has low computation complexity and we impose no change on the headers of inner nodes, just as explained in [7], MNT fails to reconstruct the path when packet losses happens and the WSN is not stable. According to data from trace-driven simulations in Citysee, a real WSN deployed in Wuxi city, China, MNT's accuracy is below 60% in a certain WSN scenario, which is not satisfactory. To be clear, we denote the node where MNT fails as the "failure node", represented by p_f. We assume that the routing protocol enables the packet header to record the node ID where MNT fails, which is possible if the routing protocol is properly designed for this purpose.

While we confess this great weakness of MNT, we shall see that although MNT fails to reconstruct the whole path, it does reconstruct the partial path before the failure node p_f. It would be better if we can leverage this information and reconstruct the remaining path using other methods. Here, we think compressive sensing is a good choice again. Here we want to use compressive sensing based method to optimize the state-of-art MNT.

Compressive sensing based method can run independently to reconstruct the packet paths. Its principles are described as follows. We manually allocate a certain algorithm in all the nodes and the sink node. This algorithm can distribute the entries of a sensing matrix ϕ to each interval nodes, thus help to construct a compressive sensing scenario. We assume it is the i^{th} time that a packet generated from source A travels the same path to the sink node, then the j^{th} hop of the path is given ϕ_{ij} as the sensing parameter, which should be added to Acc, which is used to record the interval nodes of each packet in a space-economical way. The whole process can be shown in Figure 4[1].

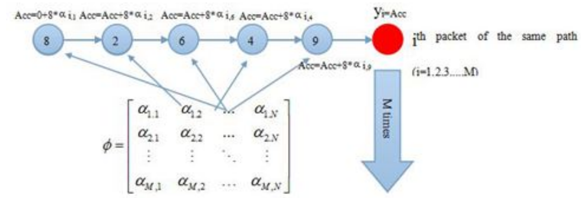

Fig. 4. Example of compressive sensing based path reconstruction method

Of course, different paths need to be classified. We can add a header bFlt using an L-bit array associated with H independent hash functions to space-efficiently record node IDs. Hash function is widely used in design of algorithms of computer networks. The headers of packets of the same classified path are included in the same "path group", denoted by $k(k = 1, 2, 3, 4, \ldots)$.The number of packets in path group k is denoted by G_k.

However, solely using compressive sensing based way is time-consuming, since we should wait M times' travel of the same path. According to latest study, M should be greater than $M_l = ck\log(\frac{N}{k})$ when c=1.5. k represents the sparsity of x.If we can reduce N, namely the width of the sensing matrix, we can reduce the reconstruction delay and loss rate as well as improving reconstruction ratio. Our motivation is to exploit the already-reconstructed partial path by MNT to reduce N in the compressive sensing method. We call the new method MNT-CS.

So our optimization method are described as follows. At the sink node, we implement MNT first to get the failure point p_f, then we use p_f to get a new sensing matrix(ϕ') and compressed information(y').

$$y = \phi x = [\phi_1 \phi_2] \times [\begin{array}{c} x_{mnt} \\ x_{cs} \end{array}] = y_1 + \phi_2 x_{cs}$$

$$(x_{mnt} = [p_1\ p_2\ p_3\ \cdots\ p_f], x_{cs} = [p_{f+1}p_{f+2}p_{f+3}\cdots\ p_n])$$

$$y' = y - y_1 = \phi_2 x_{cs}$$

$$(x_{mnt} : M \times f\ x_{cs} : M \times (n - f))$$

Since x_{mnt} is subtracted from x, the sparsity k of x_{cs} is will not increase, thus further proves the rationality of our motivation.

Then we just need to solve $y = \phi_2 \times x_{cs}$, in which the number of sensing matrix's width is reduced by f. Connect the MNT reconstructed x_{mnt} and CS reconstructed x_{cs} ,we can get the reconstructed path of packet A. After a certain time when enough paths of different packets have been reconstructed, the routing matrix can thus be generated.

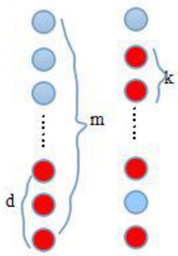

Fig. 5. The two columns have the same number of red nodes. Regard the red nodes in each column as 1, blue nodes as 0. Do the logical OR between the two columns. The answer column must has more than 1.5 times of red nodes than each original columns

VI. Reconstruction of Link Parameters

Through MNT-CS, we get the routing matrix $\mathbf{R}$. As we have already known in Section I that in our problem formation, $\mathbf{R}$ is very sparse, then we need to verify that for a very sparse random sensing matrix, l_1-minimization is qualified to reconstruct the inner-link parameters. But first, we need to use the conclusion in[3].That paper explores the potential of expander graphs to reconstruct to link parameters with a high probability.

First let's see the definition of expander graph, which is studied recent years and widely used self-correcting codes and computer networks.

Definition 1: A bipartite graph G(X,Y,H) with a left degree d(i.e., deg(v)=d $\forall$ v $\in$ X) is a (ϕ, d, ϵ)-expander if for any $\Phi \subset X$ with $|\Phi| \leq \phi$, the following condition holds:

$$|N(\Phi)| \geq (1-\epsilon)d|\Phi|,$$

where N(Φ) is a set of neighbors of Φ. ϕ and ϵ are the "expansion factor" and the "error parameter", respectively.

The author proposes that if the routing matrix of WSN can be viewed as an expander graph or the combination of several expander graphs with different left degrees, l_1-minimization method can be applied with high probability. Inspired by this conclusion, we can calculate the probability for a simplest (2,d,1/4) expander graph(two columns of routing matrix for simplicity) to be included in a random sparse matrix as the answer of a basic permutation and combination problem shown in Figure 5. The solution equation is,

$$P = \frac{C_M^d - \sum_{k=0}^{\text{floor}(0.5d)} C_d^{d-k} C_{M-d}^k}{C_M^d}$$

If P is close to 1, we can say confidently that link parameters can be reconstructed with high accuracy, using the routing matrix generated by MNT-CS.

The integral algorithm of our WSN tomography is listed as follows.

Algorithm 1 Procedures for MNT-CS based WSN tomography

M_p is the number of reconstructed paths
$M_p = 0$
for Each packet p_i received by sink node **do**
  Check its path header, bFlt
  if p_i belongs to path group k **then**
    G_k=G_k+1
  end if
  if p_i does not belong to path group **then**
    $G_{max(k)+1} = 1$
  end if
  if $G_k \geq m_l$ **then**
    Get MNT-CS switch point p_f
    Reconstruct path before node p_f using MNT
    Reconstruct the remaining path using compressive sensing, we get the whole path P_k
    $M_p = M_p + 1$
  end if
  if $M_p > n$, n is the total number of links mentioned in Section IV. We choose n to guarantee the sparsity of the sensing matrix **then**
    Use link delay accumulation data in the sink n-ode,run step2: reconstruction of inner-link delays
    Update the resultant database
  end if
end for

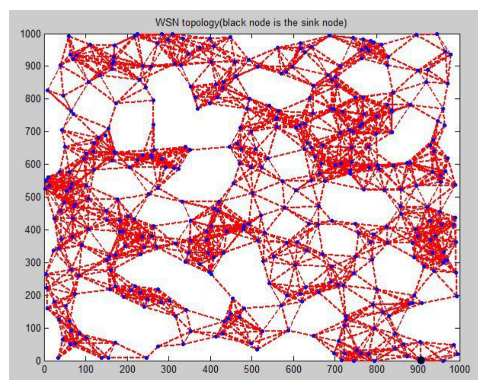

Fig. 6. WSN simulation topology(black node is the sink node). Note: This is the topology seen from the sink node in a specific time period, rather than a stasis topology of WSN

VII. Simulations

A. Simulation of MNT-CS

According to the algorithm and analysis of Section V, Section VI, we set up simulations independently for both two steps: MNT-CS reconstruction and reconstruction of inner-link parameters.

We implement our simulation on Matlab. We set 400 nodes(1 sink node) randomly distributed in a 1000*1000 area, as shown in Figure 6. Nodes in the radius of 65 can form a link. The number 65 is chosen to ensure each nodes are connected.

We simulate 40000 time units. In each time unit, 5 nodes transmit packets. Nodes are classified as 'active' and 'inactive', whose sending probability is 0.0125 and 0.001.The routing path is formed by the shortest path algorithm. We should notice that in real world, the paths of the packets generated from the same node will be different. To simplify our simulation, we assume that in a specific time of period, the paths of every

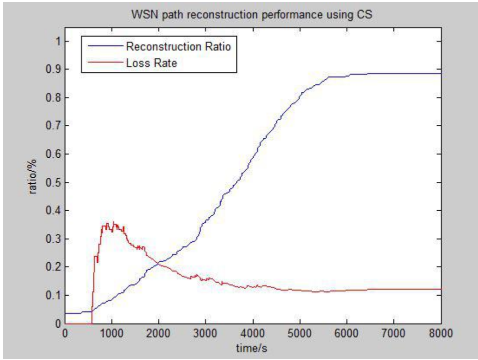

Fig. 7. Simulation result using only CS

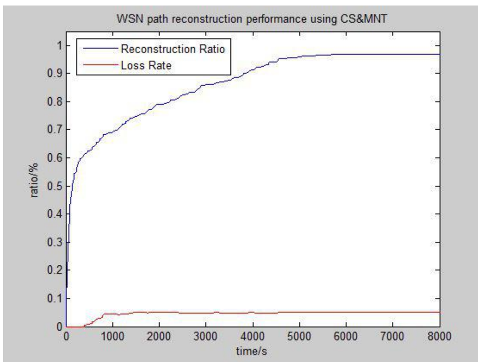

Fig. 8. Simulation result using CS & MNT

packet follows the principle of shortest path algorithm. Since our algorithm is not concerned about the routing algorithm, this assumption is reasonable. The result of our simulation will not be effected by the simplification.

The success rate of MNT is set as 70%, which means 30% packets' path need CS to reconstruct.

We plot the curve of reconstruction ratio and loss rate during the first 8000 time units. First we only implement compressive sensing method and get Figure 7.

Then we implement our refinement using the combination of MNT and CS and get Figure 8.

From Figure 7 and Figure 8 above, we can observe the improvements of our refinement:

1) Lower loss rate. When reconstruction procedure is finished, CS based reconstruction has a loss rate of 12.15%, while CS&MNT is 6.20%
2) Less delay. The time when CS based method reaches its reconstruction plateau is 6459, while CS&MNT is 5433.
3) Higher reconstruction rate. MNT can help CS to reconstruct the paths of inactive nodes. CS based method reconstructs 88.72% paths while CS&MNT reconstructs 96.74%.

B. Simulation of reconstruction of inner-link parameters

plot $P = \frac{C_M^d - \sum_{k=0}^{\text{floor}(0.5d)} C_d^{d-k} C_{M-d}^k}{C_M^d}$ in Matlab, where the x axis is the number of non-zero entries of each column, y axis is the probability. The total number of nodes is 400.The result is shown in Figure 9. As we can see, when x is small, the probability is very close to 1, which indicates the effectiveness of a random routing matrix in a large WSN to be the sensing matrix of link parameter reconstruction.

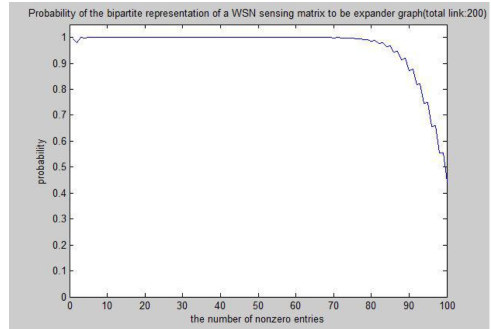

Fig. 9. The relationship between the effectiveness of routing matrix to be sensing matrix and the sparsity of the network

VIII. CONCLUSIONS

In this work, we put the problem of WSN topology tomography into a compressive sensing representation and separate its application in WSN into 2 steps, where routing matrix can not be pre-defined . In step1, we proposed a refinement on WSN path reconstruction problem based on MNT and CS. In step2, we tried to verify that for a very sparse WSN routing matrix, l_1-minimization method is reliable to reconstruct routing parameters.

We still have several future works to go if we want to get a more convincing result. First, we can find a way to predict the MNT&CS switch point in Step1 using network information.Second, we should implement a more reliable simulation and even deploy MNT & CS method in real wireless sensor networks. Moreover, the proper number of routes included in the routing matrix need to be clarified in a numerical way.

REFERENCES

[1] Xiaoyan Z, Houjun W, Zhijian D. Wireless sensor networks based on compressed sensing[C]//Computer Science and Information Technology (ICCSIT), 2010 3rd IEEE International Conference on. IEEE, 2010, 9: 90-92.
[2] ctt) Liang Y, Liu R. Compressed topology tomography in sensor networks[C]//Wireless Communications and Networking Conference (WCNC), 2013 IEEE. IEEE, 2013: 1321-1326.
[3] Firooz M, Roy S. Link delay estimation via expander graphs[J]. 2011.
[4] Keller M, Beutel J, Thiele L. Multi-hop network tomography: path reconstruction and per-hop arrival time estimation from partial information[C]//ACM SIGMETRICS Performance Evaluation Review. ACM, 2012, 40(1): 421-422.
[5] Xu W, Mallada E, Tang A. Compressive sensing over graphs[C]//INFOCOM, 2011 Proceedings IEEE. IEEE, 2011: 2087-2095.
[6] Berinde R, Indyk P. Sparse recovery using sparse random matrices[J]. preprint, 2008.
[7] Gao Y, Dong W, Chen C, et al. Zhejiang Provincial Key Laboratory of Service Robot, College of Computer Science, Zhejiang University, China[C]//Network Protocols (ICNP), 2013 21st IEEE International Conference on. IEEE, 2013: 1-10.
[8] Lu X, Dong D, Liao X, et al. PathZip: Packet path tracing in wireless sensor networks[C]//Mobile Adhoc and Sensor Systems (MASS), 2012 IEEE 9th International Conference on. IEEE, 2012: 380-388.
[9] Anamika,Neeru Singla,Neha,Sharma, A Dynamic Network Reconstruction Approach using ACO, IOSR Journal of Electronics and Communication Engineering (IOSRJECE)ISSN: 2278-2834 ISBN: 2278-8735Volume 3, Issue 5 (Sep. C Oct. 2012), PP 27-35.

[10] Siyari P, Rabiee H R, Salehi M, et al. Network Reconstruction under Compressive Sensing[C]//Social Informatics (SocialInformatics), 2012 International Conference on. IEEE, 2012: 19-25.
[11] Xu W, Mallada E, Tang A. Compressive sensing over graphs[C]//INFOCOM, 2011 Proceedings IEEE. IEEE, 2011: 2087-2095.
[12] Mao Y, Kschischang F R, Li B, et al. A factor graph approach to link loss monitoring in wireless sensor networks[J]. Selected Areas in Communications, IEEE Journal on, 2005, 23(4): 820-829.
[13] Gnawali O, Fonseca R, Jamieson K, et al. Collection tree protocol[C]//Proceedings of the 7th ACM Conference on Embedded Networked Sensor Systems. ACM, 2009: 1-14.
[14] Simek M, Moravek P, SA S. Wireless Sensor Networking in Matlab: Step-by-Step[C]//proceedings of International Conference on Teleinformatics. 2011: 185-190