

# 传感网中基于压缩感知的丢包 匹配数据收集算法

孙泽宇, 李传锋, 阎 奔

(洛阳理工学院计算机与信息工程学院, 河南洛阳 471023)

**摘 要:** 为了提高传感网中数据重构精度以及降低不可靠链路丢包对压缩感知(Compressive Sensing, CS)数据收集的影响,本文提出了一种基于压缩感知丢包匹配数据收集算法(Packet Loss Matching Data Gathering Algorithm Based on Compressive Sensing, CS-MDGA). 本文算法通过压缩感知技术构建了全网数据间的“关联效应”,并设计了基于丢包匹配的稀疏观测矩阵(Sparse Observation Matrix Based on Packet Loss Matching, SPLM),证明了该观测矩阵概率趋近于“1”时,满足的等距约束条件(Restricted Isometry Property, RIP),完成了节点间多路径路由数据的可靠交付. 仿真实验结果表明,本文算法在链路丢包率为60%情况下,相对重构误差仍小于5%,验证了本文算法不仅具有较高的重构精度,而且还可以有效缓解不可靠链路丢包对CS数据收集的影响.

**关键词:** 传感网; 压缩感知; 数据收集; 关联效应; 稀疏观测矩阵

**中图分类号:** TP393 **文献标识码:** A **文章编号:** 0372-2112 (2020)04-0723-11

**电子学报 URL:** <http://www.ejournal.org.cn> **DOI:** 10.3969/j.issn.0372-2112.2020.04.014

## CS-MDGA: A Packet Loss Matching Data Gathering Algorithm in Sensor Networks Based on Compressive Sensing

SUN Ze-yu, LI Chuan-feng, YAN Ben

(School of Computer Science and Engineering, Luoyang Institute of Science and Technology, Luoyang, Henan 471023, China)

**Abstract:** In order to improve the data reconstruction accuracy and alleviate the influence of packet loss over unreliable links on the Compressive Sensing (CS) data gathering in sensor networks, we propose a Packet Loss Matching Data Gathering Algorithm Based on Compressive Sensing (CS-MDGA) in this paper. This proposed algorithm establishes the correlation effect of the network data with the CS technique. We further design the Sparse Observation Matrix based on Packet Loss Matching (SPLM) in this paper. In addition, we prove that the designed observation matrix satisfies the Restricted Isometry Property (RIP) with a probability arbitrarily close to 1, which can guarantee the reliable delivery of the multi-path routing data among different nodes. The simulation results show that the relative reconstruction error of this proposed algorithm is still lower than 5% even when the packet loss rate of the link is as high as 60%. Therefore, it is verified that this proposed algorithm not only exhibits high reconstruction accuracy, but also effectively alleviates the influence of packet losses over unreliable links on the CS-based data collection.

**Key words:** sensor networks; compressive sensing; data gathering; correlation effect; sparse observation matrix

### 1 引言

无线传感网络中节点通常被密集布设,网内感知数据存在大量冗余,浪费了节点的能量资源. 压缩感知(Compressive Sensing, CS)技术是一种新兴的压缩理论,

能够极大降低采样频率,且压缩和采样同时进行,因此受到传感网研究者的广泛关注<sup>[1]</sup>. 为均衡和降低节点能耗,延长网络工作寿命,研究者们提出了基于压缩感知的数据收集算法. 目前,该类算法研究大多集中在如何有效降低网络能耗,延长网络寿命上<sup>[2]</sup>. 针对实际应

收稿日期:2018-08-13;修回日期:2019-02-22;责任编辑:孙瑶

基金项目:国家自然科学基金(No. U1604149);河南省教育厅高等学校青年骨干教师培养计划(No. 2016GGJS-158);河南省教育厅重点项目资助计划(No. 19A520006);河南省科技厅科技攻关计划(No. 182102210428);洛阳理工学院高层次人才资助计划(No. 2017BZ07)

用场景下的无线链路丢包问题,目前相关的压缩感知数据收集算法研究还相对较少,而由于无线链路的动态性和不对称性、信道干扰冲突、不恰当的天线方向和高度等原因,不可靠链路通常是传感网数据收集在实际应用中面临的关键问题<sup>[3]</sup>。

针对不可靠链路下的压缩感知数据收集问题,本文提出了一种基于压缩感知丢包匹配数据收集算法(Packet Loss Matching Data Gathering Algorithm Based on Compressive Sensing, CS-MDGA)。该算法将全网节点分为两类,一类为传统的中转收发节点(Traditional Forwarding, TF),另一类为采用 CS 观测的压缩感知节点;由于 TF 节点的丢包不具有关联性,而 CS 节点的丢包具有强关联性,因此在不可靠链路下,若 TF 节点发生丢包,则设计采用基于丢包匹配的稀疏观测矩阵观测其发送的感知数据,从而将链路丢包问题转变为压缩感知的稀疏匹配采样过程,而对于 CS 节点的丢包,则设计采用多路径备份传输机制来保障数据的可靠交付,避免丢包的“关联效应”,从而降低不可靠链路对 CS 数据收集过程的影响,保证感知数据的重构精度。

本文的贡献主要分为以下三点:(1)分析树状路由下链路丢包将严重恶化 CS 数据收集过程的数据重构精度,指出 CS 数据收集过程的链路丢包具有“关联效应”;(2)设计了(Sparse Measurement Matrix Based on Packet Loss Matching, SPLM)SPLM 稀疏观测矩阵,并证明该观测矩阵以趋近于“1”的概率满足 RIP 条件;(3)设计了基于混合压缩感知的 CS 节点间的多路径备份路由传输机制,保障压缩感知投影数据的可靠交付。

## 2 相关工作

压缩感知由于具有编码端简单,解码端复杂的特点,已经被广泛应用到无线传感网络数据收集领域中。目前,基于压缩感知的无线传感网数据收集算法研究主要集中在如何利用压缩感知技术降低无线传感网数据收集的网络能耗上,且大都假设网络链路为理想链路,不考虑链路丢包对 CS 数据收集过程的影响。如文献[4]讨论了采用稀疏观测矩阵,降低每次观测的通信代价;文献[5]利用感知数据的时空相关性,提高感知数据的压缩比,降低观测次数;文献[6]采用多级分层的簇状拓扑结构收集网内数据,减少各层节点的收发数据包数,从而减少整体网内的传输数据包数;文献[7]提出采用分块对角观测矩阵观测全网节点数据,在不降低重构精度的条件下,减少观测次数,延长网络寿命。近年来,随着 CS 数据收集理论的逐步成熟,研究者们也着手研究面向实际应用的 CS 数据收集算法。文献[8]考虑到真实数据集中,数据的稀疏度会随时空发生变化,因此提出采用自回归 AR 模型预测数据的变化,

自适应调整观测次数以达到重构性能最优。文献[9]认为无线链路上的环境噪声对网内传输的 CS“欠采样”数据具有较大影响,提出近似梯度下降算法重构噪声下的压缩数据。文献[10]对比了树状拓扑结构下应用压缩感知和不应用压缩感知的网内数据包发送量,提出了一种混合压缩感知的数据收集方法,将传统中转收发的数据收集方式和压缩感知数据收集方式相结合,进一步降低了网内的收发数据包量,降低网络能耗。文献[11]研究压缩感知技术在簇状路由结构下的数据收集应用,由于簇状拓扑为单跳的传输路由,网内的链路丢包不具有关联性,因此压缩感知数据收集算法对链路上的丢包不敏感。文献[12]提出了一种簇状拓扑下的不可靠链路压缩感知数据收集算法(Compressive Data Gathering, CDG),根据簇内的丢包节点调整观测矩阵的列向量,降低丢包对 CS 数据重构的影响。对于大规模的无线传感网络多采用树状多跳的网络路由拓扑结构,文献[13,14]研究基于多跳路由的压缩感知数据收集算法,该类算法不考虑无线链路的不可靠性,着重讨论了压缩感知的观测矩阵与树状路由结构的最优匹配问题;而多跳路由下传输 CS 数据包需加权叠加多个节点的感知数据,一旦发生丢包则造成多个原始感知数据丢失,因此多跳路由结构下的压缩感知数据收集对链路丢包十分敏感。文献[15]仿真验证了树状拓扑结构下链路丢包将大大降低 CS 数据的重构精度,提出了一种基于极稀疏随机调度的压缩感知数据算法(Sparsest Random Scheduling for Compressive Data Gathering in Lossy WSNs, SRS),该算法根据 Sink 端数据包接收情况构造极稀疏观测矩阵,并以此重构全网节点的原始感知数据,减少丢包对 CS 重构精度的影响,但该算法仅适用于全网感知数据具有较强空间相关性的场景中,应用相对局限。

传统无线传感网数据收集方法中保障链路可靠传输的方法主要有:多路径传输<sup>[16]</sup>,网络编码等<sup>[17]</sup>,而针对基于压缩感知的传感网可靠数据收集方法研究较少,并且 CS 数据收集方法对链路丢包的敏感性要远高于传统数据收集,因此研究不可靠链路下的压缩感知数据收集算法对压缩感知技术在实际传感网中的应用具有重要意义。

## 3 网络模型与问题描述

压缩感知技术是一种新兴的数据压缩理论,对于具有稀疏性的信号,压缩感知可以以低于奈奎斯特采样频率压缩采样,实现信号从高维到低维的投影变换,利用最优化重构算法实现压缩信号的高精度重构,由于其优异的压缩性能,已受到了各领域的广泛研究与应用<sup>[18,19]</sup>。

**定义 1** 链路状态矩阵 (Link State Matrix, LSM). LSM 记录链路状态信息的矩阵,其大小为  $M \times N$ ,其中  $M$  为观测次数, $N$  为全网节点个数,LSM 公式为:

$$L = (l(i, j))_{M \times N} = \begin{cases} 0, & x_j \notin X \\ 1, & x_j \in X \end{cases} \quad (1)$$

**定义 2** 稠密随机观测矩阵 (Dense Random Projections, DRP). DRP 为稠密观测矩阵,其每行包含  $O(N)$  个非零元素,通常选择式(2)构造稠密观测矩阵:

$$\Phi_d = (\delta(x, y))_{M \times N} \quad (2)$$

基于丢包匹配的稀疏观测矩阵利用实际链路中丢包的随机性构造具有随机特点的稀疏观测矩阵,其构造可通过 LSM 与 DRP 的矩阵点乘运算得到,如式(3)所示:

$$\Phi_s = L\Phi_d \quad (3)$$

若不可靠链路的丢包率为  $p$ ,则在 SPLM 矩阵中,各元素的取值如式(4):

$$\delta_s(i, j) = \begin{cases} +1, & \text{with prob. } (1-p)/2 \\ -1, & \text{with prob. } (1-p)/2 \\ 0, & \text{with prob. } p \end{cases} \quad (4)$$

设计观测矩阵需保证其与大多数正交基满足 RIP 条件,然而证明 RIP 条件是个 N-P 难问题,文献[11]指出若观测矩阵满秩,则经该矩阵投影后的数据将以趋于“1”的概率被精确重构.由于 SPLM 矩阵中各元素均满足式(4)所示的离散随机分布,则  $\Phi_s$  的每行看作为随机变量  $\zeta_n$  产生的随机序列,可用离散随机过程  $\{\zeta(n), n = 1, 2, \dots, N\}$  表示.

在传感网中,假设网络布设范围内随机部署  $N$  个节点,其采集的感知数据记为  $\mathbf{d} = (d_1, d_2, \dots, d_N)^T$ ,若  $\mathbf{d}$  在稀疏基  $\Psi_{N \times N}$  下稀疏,则观测矩阵为  $\Phi = (\delta)_{M \times N}$ ,观测向量  $\mathbf{Y}_{M \times 1}$  可表示为  $\mathbf{Y} = (\mathbf{y}_i)_{M \times 1} = \Phi - \Psi^T \cdot \mathbf{d}$ ,Sink 节点可通过式(5)和式(6)进行优化问题,在一定精度约束下重构原数据:

$$\mathbf{Y} = \Phi - \mathbf{S} = \Phi - \Psi^T \cdot \mathbf{d} = \Theta \cdot \mathbf{d} \quad (5)$$

$$\mathbf{d} = \operatorname{argmin} \|\mathbf{d}\|_p \quad (6)$$

其中,  $\Theta = \Phi - \Psi^T$  为传感矩阵,  $\|\mathbf{d}\|_p$  表示感知数据向量  $\mathbf{d}$  的  $l_p$  范数,其定义为:

$$\|\mathbf{d}\|_p = \begin{cases} \left( \sum_{i=1}^N |x_i|^p \right)^{1/p}, & 0 < p < +\infty \\ \max_{i=1, 2, \dots, N} |x_i| \end{cases} \quad (7)$$

在无线传感网络的数据收集过程中,CS 的每轮数据收集分为  $M$  次独立观测分别进行,其数学表达式可描述为:

$$\begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_M \end{bmatrix} = \begin{bmatrix} \varphi_{11} & \varphi_{12} & \cdots & \varphi_{1N} \\ \varphi_{21} & \varphi_{22} & \cdots & \varphi_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \varphi_{M1} & \varphi_{M2} & \cdots & \varphi_{MN} \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_N \end{bmatrix} \quad (8)$$

## 4 算法设计与实现

### 4.1 SPLM 观测矩阵设计

为解决参与压缩的感知数据与观测矩阵采样不匹配问题,本文设计了基于丢包匹配的稀疏观测矩阵 (Sparse Measurement Matrix Based on Packet-Loss Matching, SPLM). 该观测矩阵通过在每次观测过程中不观测丢包节点信息,从而将 CS 数据收集在树状路由下的数据丢失问题转换为基于稀疏匹配的观测矩阵投影问题,既实现了对网内数据的大量观测采样,又避免了 Sink 端对数据收集情况的误判,其具体实现过程如下:假设传感网由节点集合  $\{s_1, s_2, s_3, \dots, s_n\}$  和一个不可移动的 Sink 节点组成,感知节点被随机均匀部署在  $a \times a$  的监测区域内且不移动, Sink 位于监测区域的中心.感知节点周期地采集并上传感知数据至 Sink 节点,且可自适应动态调整自身的发射功率; Sink 节点具有较强的计算性能,周期地收集和重构全网感知数据,且可获知全网节点的位置信息.全网节点构建最小生成树 (Minimum Spanning Tree, MST) 路由执行数据收集,即生成连通的无向图  $G = (V, E(q))$ ,其中  $V = \{v_1, v_2, \dots, v_N\}$  代表感知节点集合,  $E(q) = \{e_1(q), e_2(q), \dots, e_N(q)\}$  表示 MST 的链路集合,  $e_i(q)$  表示该链路以概率  $q$  连通,若令  $p = 1 - q$ ,则  $p$  表示链路的丢包率.此外,传感网采用压缩感知的数据收集方式,其具有以下特点: (1) 感知数据向量的稀疏变换基  $\Psi_{N \times N}$  采用离散傅里叶变换 (Discrete Fourier Transform, DFT),稀疏变换和正交稀疏基分别如式(9)和(10)所示; (2) Sink 端接收到  $M$  个观测值后,采用正交匹配追踪 (Orthogonal Matching Pursuit, OMP) 法重构原始感知数据; (3) CS 重构精度采用如式(11)所示的相对误差  $\eta$  衡量,其值越小,其重构精度越高;定义若重构误差高于 5%,则认为本次重构不满足精度要求,即重构失败.

$$X(k) = \sum_{n=0}^{N-1} x(n) e^{-j\frac{2\pi}{N}kn} = \sum_{n=0}^{N-1} x(n) W_N^{kn} \quad (9)$$

$$\Psi_j(t) = 1/\sqrt{N} e^{j2\pi njt/N} \quad (10)$$

$$\eta = \|\hat{\mathbf{x}} - \mathbf{x}\|_2 / \|\mathbf{x}\|_2 = \sqrt{\sum_{n=0}^{N-1} (\hat{x}_n - x_n)^2} / \sqrt{\sum_{n=0}^{N-1} x_n^2} \quad (11)$$

不可靠链路下,树状拓扑网络的 CS 数据收集过程如图 1 所示.

公式  $\sum_{j=1}^5 \delta_{i,j} x_j$  表示为  $S_1$  至  $S_5$  全部数据的总和;若  $S_5$  与 Sink 间的链路发生丢包,  $S_5$  所对应的全部数据包丢失,即数据包  $\sum_{j=1}^5 \delta_{i,j} x_j$  全部数据包丢失.由于  $S_5$  与 Sink 节点之间链路的不可靠性,使得节点  $S_1$  至节点  $S_5$  所收



中 TF 节点仅以传统数据收集方式转发数据,丢包不具有关联性,而 CS 节点以 CS 数据收集方式收发数据,丢包具有关联性.因此,针对 TF 节点间的数据收集,采用 SPLM 观测矩阵观测采样仅可以克服丢包对 CS 数据重构的影响,因此,除采用 SPLM 矩阵观测采样外,需设计相应机制保障 CS 节点间的数据传输可靠性.

### 4.3 多路径传输机制

本文设计基于多路径备份路由的传输机制保障 CS 节点间的数据传输可靠性.在正常情况下,CS 节点采用 MST 路由收发数据包,若 CS 链路上发生数据包丢失,则该数据包的发送节点  $S_i$  将更换传输路径,采用备份路径向 Sink 发送数据包,节点  $S_i$  可看作源点, Sink 为目的节点.现有的从源点到目的节点的路由构建方式有很多,为降低备份路径传输的能耗,本文采用最小能耗生成树路由的构建方式,网络的能耗模型如式(19):

$$E(d) = 2\alpha_1 + \alpha_2 d^n \quad (19)$$

其中,  $\alpha_1$  为电路能耗系数,  $\alpha_2$  为功率放大系数,  $d$  传输距离,  $n$  为路径损耗因子 ( $2 \leq n \leq 5$ , 在自由空间中通常取  $n=2$ ), 因此最小能耗路由的构建问题可建模为如式(20)的最优化问题,其中  $d_0$  为源点到目的节点的距离,  $K$  为跳数,  $d_i$  为节点间的链路距离.

$$\min \left( \sum_{i=1}^K 2\alpha_1 + \alpha_2 d_i^n \right) \quad (20)$$

$$\text{s. t.} \quad \sum_{i=1}^K d_i = d_0$$

$$d_{\text{char}} = (2\alpha_1 / \alpha_2 (n-1))^{1/n} \quad (21)$$

利用拉格朗日乘数法求解上述最优化问题,当且仅当在目的节点和源点间每跳的距离都相等时,网络能耗最小,且给出特征距离  $d_{\text{char}}$  的取值如式(21)所示,则最优跳数  $K_{\text{opt}}$  取  $d$  与  $d_{\text{char}}$  的商上限.本文构建备份路径选择集中式的构建方法,具体分为以下四步:(1)由 Sink 端根据各节点的地理位置信息,计算由 Sink 节点至各 CS 节点的直线距离  $d$ ; (2) Sink 端比较各节点的直线距离  $d$  与特征距离  $d_{\text{char}}$  的大小,若  $d \leq d_{\text{char}}$ , 则构建单跳的备份路径,若  $d > d_{\text{char}}$ , 首先计算最优跳数  $K_{\text{opt}}$ , 根据最优跳数  $K_{\text{opt}}$ , 在 CS 节点至 Sink 的直线路径上,基于每跳距离均相等的原则,计算理论上的理想中继节点地理位置;(3) Sink 端根据理想中继地理坐标,选择距离该坐标最近的节点作为中继节点;(4) Sink 将构建好的 CS 节点最小能耗路由分发至 CS 节点和相应的中继节点.

CS-MDGA 算法首先构建全网节点 MST 路由树,树中节点  $i$  的全部孩子节点数目为  $\omega_i$ , 特别地,对于 Sink 节点  $\omega_{\text{sink}} = N$ , 对于位于链路末端的节点  $\omega_i = 0$ , 定义  $\varepsilon = M - 1$  为区分节点类型的阈值,各节点根据自身  $\omega_i$  的大小,若  $\omega_i > \varepsilon$ , 节点  $i$  以 CS 方式参与数据收集,称为 CS 节点,若  $\omega_i \leq \varepsilon$ , 节点以传统转发方式参与数据收集,

称为 TF 节点,则在数据收集过程中,节点  $i$  发送的数据包个数  $PN(i)$  为:

$$PN(i) = \begin{cases} \omega_i + 1, & \omega_i \leq \varepsilon \\ M, & \omega_i > \varepsilon \end{cases} \quad (22)$$

CS-MDGA 算法的数据收集过程中,TF 节点以传统转发方式沿 MST 路由收集数据,如图 2 中白色节点,其丢包并不具有“关联效应”;CS 节点以 CS 数据收集方式沿 MST 路由参与数据收集,如图 2 中黑色节点,其丢包具有“关联效应”.在 MST 路由中,对于与 TF 节点直接相连的 CS 节点,其每次完成对全部孩子节点的数据收集后,需将收集到的全部节点数据加权叠加,合并为一个适用于 CS 节点间传输的数据包格式,并在孩子数据包的节点 ID 部分标记相应的标志位,以标识各节点的数据接收情况.

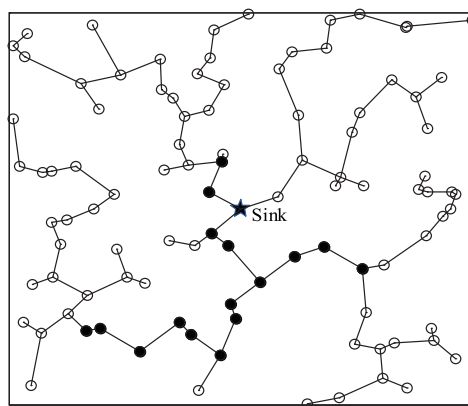


图2 CS-MDGA算法中不同类型节点分布

不可靠链路下,若 TF 节点间的链路发生丢包,则不做处理;若 CS 节点间的链路发生丢包,则采用最小能耗备份路径直接将数据包发送至 Sink 端.每轮数据收集结束后, Sink 节点结合每次观测过程的丢包情况构建 SPLM 观测矩阵,利用 SPLM 观测矩阵和 Sink 端接收到的  $M$  个观测值重构原始感知数据.

### 4.4 数据收集与分析

CS-MDGA 算法的数据收集过程分为 3 个阶段分别进行:第 1 阶段为传感网初始化阶段,该阶段完成的是全网节点的组网、网内各链路接收状态的先验信息积累和 CS 数据采样的节点观测向量配置;第 2 阶段为 CS 数据收集阶段,该阶段实现在有损链路下对网内各节点采集数据的有效 CS 采样和收集;第 3 阶段为采样数据的 CS 重构阶段,该阶段针对采样数据进行重构,以获得网内各节点的原始采集数据.算法流程如图 3 所示.

### 4.5 算法描述

**Step1** 传感网络初始化.首先,全网节点由 Sink 端发起,构建最小生成树 (Minimum Spanning Tree, MST) 路由,完成节点组网.

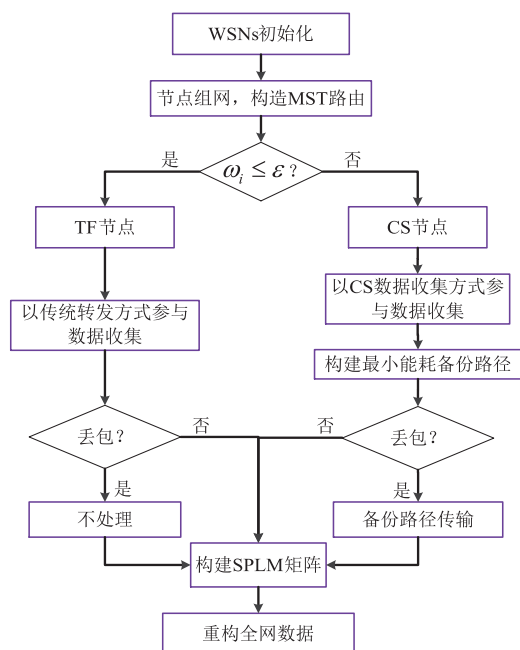


图3 CS-MDGA算法流程图

**Step2** Sink 向全网广播心跳数据包, 网内各节点在接收到 Sink 端心跳包后, 在时间段  $T_1$  内, 全网节点以传统的发送、转发的数据收集方式沿路由向 Sink 端传递节点自身心跳包。

**Step3** 各节点实时统计记录自身接收链路上数据包的接收情况, 并将结果存储在其内存中, 作为滑动窗丢包预判机制的先验信息。

**Step4** 时间  $T_1$  结束时, 各节点的接收序列  $\{X_i\}$ , Sink 先向全网广播随机种子  $\xi$ , 节点  $i$  收到随机种子  $\xi$  后与自身的节点 ID 号结合  $(\xi, ID_i)$ , 生成属于自身的一列观测矩阵  $(\delta_{i1}, \delta_{i2}, \dots, \delta_{iM})^T$  存储到自身内存中。

**Step5** 各节点按照路由将采集数据  $d_i$  与对应的观测系数  $\delta_{ij}$  相乘, 依次叠加并传送至 Sink 端。根据式(22)选择重传机制恢复丢包, 设定最大重传次数为  $\max\_num$ 。否则, 将预测值作为丢包节点接收到的数据包向下一跳节点传输, 完成本次 CS 观测采样。

**Step6** Sink 节点在收到一轮数据收集的  $M$  个观测值后, 构建观测向量  $\mathbf{Y} = (y_1, y_2, \dots, y_M)^T$ , Sink 根据随机种子  $\xi$  和全网节点 ID 号重建观测矩阵  $(\delta_{ij})_{M \times N}$ , 结合稀疏基  $\Psi_{N \times N}$ , 利用 CS 重构算法重构出稀疏矩阵  $\mathbf{S}$ , 通过计算  $\mathbf{d} = \Psi \cdot \mathbf{S}$  即可重构出原始信号向量  $\mathbf{d}$ 。

根据上述过程可知: 对于整个网络, CS-MDGA 算法执行一轮数据收集, 全网至少需收发  $N \times M$  个数据包, 其中,  $N$  为全网节点个数,  $M$  为观测次数。由于 CS 数据收集过程中, 收发任意一个数据包在算法执行上为一次乘法和加法线性运算, 并不涉及复杂的循环等操作, 且根据 CS 理论, 观测次数  $M \geq O(K \log N)$ ,  $K$  为数据稀

疏度, 可视为一个常量, 因此对于全网而言, CS-MDGA 算法的复杂度为  $O(M \log N)$ 。相较于传统的 CS 稠密投影数据收集算法, 该算法并不增加额外的计算开销。

## 5 性能评估

为进一步评估本文 CS-MDGA 算法的性能, 本文在以下五种不同环境中对本文算法进行了评估: (1) 不同丢包率下数据重构误差对比; (2) 三种算法在网络寿命和不同稀疏下丢包率对比; (3) 观测次数与 CS-MDGA 算法的性能对比; (4) 不同网络规模下 CS-MDGA 算法性能对比; (5) 不同路径下 CS-MDGA 算法性能对比等方面进行了仿真实验。该仿真实验以 MatlabR2014a 作为仿真平台, 基于本文的网络模型, 假设网络部署在  $200 \times 200$  的监测区域内, 全网内共有 500 个感知节点, 选择服从二维高斯分布的数据作为原始数据源。

### 5.1 实验一: 重构误差与网络丢包率对比

为评估 SPLM 矩阵的性能, 选择经典的 CS 数据收集 CDG 算法作为仿真算法, 对于链路丢包的处理分别采用三种不同的方法作为对比, 其中 CDG-DRP 表示稠密观测矩阵观测全网数据; CDG-SPLM 表示基于丢包标识的稀疏观测矩阵观测全网数据; CDG-SPLM-NC 表示不考虑丢包的前提下, 采用基于丢包标识的稀疏观测矩阵观测全网数据。

从图 4 中可以看出, CDG-SPLM 算法和 CDG-SPLM-NC 算法的重构精度明显高于采用稠密观测矩阵的 CDG-DRP 算法, 这是由于 CDG-DRP 算法存在 Sink 端对数据包接收情况的误判问题, 而采用 SPLM 矩阵观测的 CDG-SPLM 和 CDG-SPLM-NC 算法避免了该问题。可见, Sink 端对数据包接收情况的误判会严重影响 CS 数据收集的重构精度。这是由于当丢包率较高时, 丢包的“关联效应”将导致几乎全网数据的丢失, 进而造成 CS 观测采样数据的严重不足, 大大降低了重构精度。因此, 在存在丢包“关联效应”的数据收集过程中, 仅解决 Sink 端对数据包接收情况的误判问题并不能满足应用需求。

图 5 给出了 CS-MDGA 算法、SRS 算法和 CDG-SPLM 算法在不同丢包率下的重构精度对比。由图 5 可知, 三种算法的重构误差率均随丢包率的增加而增加, 但 CS-MDGA 算法的重构精度处于最优水平, 且随丢包率逐渐变大, CS-MDGA 算法的优势也更加明显。造成三种算法重构精度差距的原因为: SRS 算法采用极稀疏观测矩阵投影重构原始数据, 一方面, 观测采样的样本数据过少, 另一方面, 丢包将降低观测次数, 造成观测次数不足而降低重构精度; CDG-SPLM 算法虽能通过克服 Sink 端对数据包接收情况的误判问题, 但 CS 数据收集的丢包关联问题会严重恶化重构精度; 而本文 CS-MDGA 算法不仅实现了对全网数据尽可能多的观测采

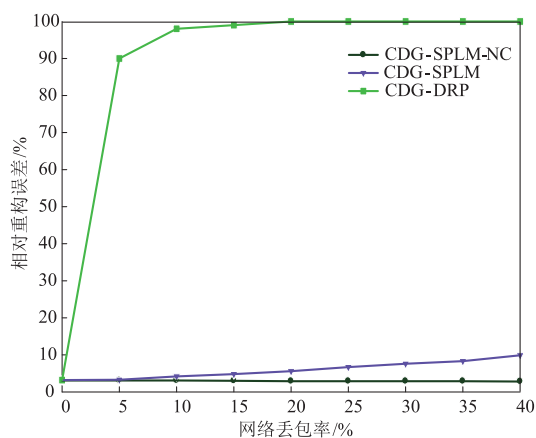


图4 丢包标识的稀疏观测矩阵性能对比

样,而且克服 Sink 端对数据包接收情况的误判问题和 CS 数据收集过程的丢包关联问题,多层次保障不可靠链路下的 CS 数据收集过程,实现了不可靠链路下的 CS 高精度重构.另外,根据仿真结果可知,本文所提出的 CS-MDGA 算法在 60% 的链路丢包率下,仍能实现 CS 数据收集的有效重构,适用范围较广.

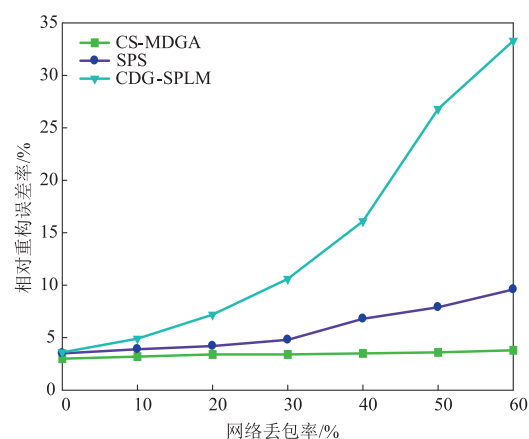


图5 不同丢包率下数据重构误差对比

评估算法在不同相关性数据集下的性能,分别选择数据稀疏度为 11 和 23 的数据集作为仿真数据.图 6 给出了 CS-MDGA 算法与 SRS 算法在不同数据稀疏度下的重构精度对比,其中链路丢包率设置为 20%.如图 6 所示,对于同一种算法,数据稀疏度越小,算法的重构精度越高,这是由于数据稀疏度越小,数据相关性越强,根据压缩感知理论,数据相关性越强,CS 重构性能就越优;而对于不同的算法,当数据稀疏度为 11 时,本文算法的重构精度略高于 SRS 算法,但当数据稀疏度为 23 时,CS-MDGA 算法的重构精度具有了较大优势,由于数据相关性弱时,SRS 算法仅基于少量的压缩样本并不能有效地评估全网节点数据,进而造成重构精度差的现象,而 CS-MDGA 算法通过尽可能多的观测采样全网节点数据,实现了对具有一般相关性数据集的 CS 有效重

构.因此,CS-MDGA 算法可有效降低不可靠链路下 CS 数据收集算法对感知数据集相关度的强依赖性,实现了针对具有一般相关性数据集的传感网在不可靠链路下的 CS 数据收集.

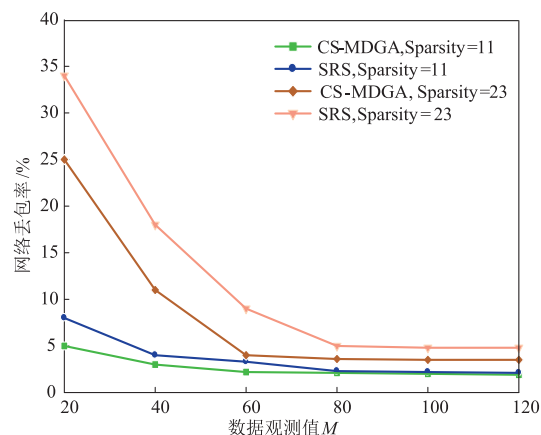


图6 不同稀疏度数据对算法性能影响对比

## 5.2 实验二:三种算法在网络寿命和不同稀疏下丢包率对比

图 7 给出了三种不同算法的网络寿命对比,其中,CDG-Retransmission 算法是指采用重传机制恢复丢包的 CDG 算法在不可靠链路中的应用;定义网络内第一个节点死亡的数据轮作为网络的最大生命周期.由图 6 可知,本文提出的 CS-MDGA 算法相较于 SRS 算法和 CDG-Retransmission 算法网络寿命延长 400% 和 140%,这是由于 SRS 算法中节点的能耗不均衡,存在“能力空洞”问题,造成网络提前终止,而 CDG-Retransmission 算法采用重传恢复丢包,增加网内的发包量,降低了网络寿命.

根据压缩感知理论,观测矩阵与数据稀疏基的相关度将影响算法的重构性能,为验证 CS-MDGA 算法在不同稀疏基下的性能,本文选择采用离散余弦变换 (Discrete Cosine Transform, DCT) 和离散傅立叶变换 (Discrete Fourier Transform, DFT) 两种稀疏基作为对比,如图 8 所示.由图 8 可知,不同的丢包率下,算法在 DCT 基下的重构性能均优于在 DFT 基下的性能,由此可知,SPLM 矩阵与 DCT 基间具有更低的相关性,选择 DCT 基作为感知数据稀疏变化的稀疏矩阵将获得更高精度的重构数据.

## 5.3 实验三:不同属性下 CS-MDGA 算法的性能对比

图 9 给出了不同观测次数下丢包率对算法性能的影响,CS-MDGA 算法观测次数分别取 60、80、120 和 140.由图 9 可知,当观测次数为 60 时,不同丢包率下 CS-MDGA 算法的重构相对误差相差较大,丢包率越高对应的重构误差越大;而随着观测次数的增加,CS-MDGA 算法的重构误差差距逐渐减小,当观测次数为 140 时,10%、30% 与 50% 的丢包率下的重构误差几乎

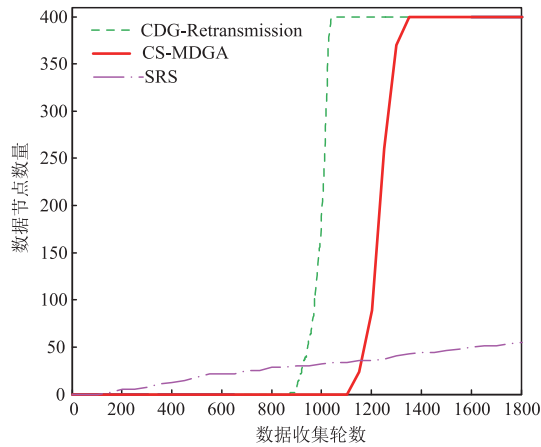


图7 三种不同算法的网络寿命对比

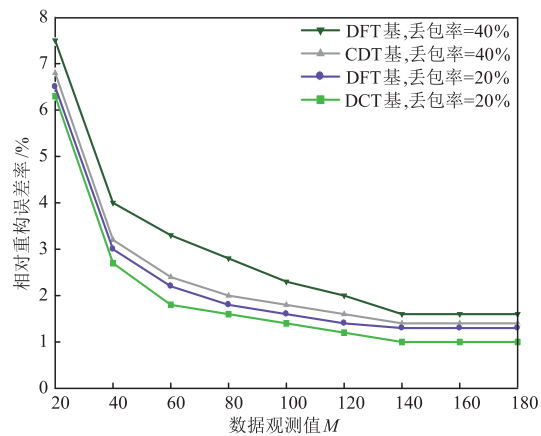


图8 不同稀疏基下CS-MDGA算法的性能对比

相等.造成该现象的原因是由于本文所提的CS-MDGA算法采用基于丢包匹配的观测矩阵观测采样全网数据,丢包率增加将导致观测矩阵稀疏度的增加,为保持重构数据精度不变,观测矩阵的稀疏度增加,需对应增加观测次数,因此,对于CS-MDGA算法,增加观测次数可在一定程度上降低丢包率对重构性能的影响.

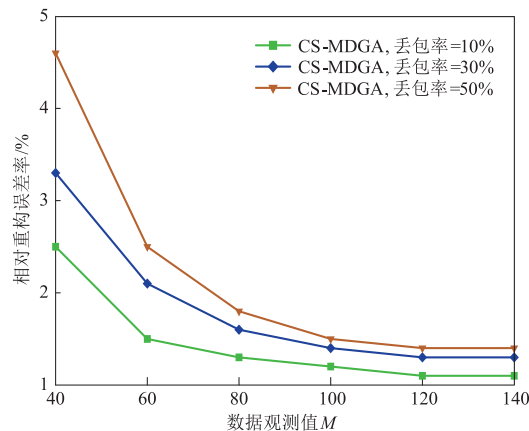


图9 不同丢包率下,观测次数对算法性能的影响

当事件源衰减系数  $n=0.01$ , 邻居范围  $r=2$  时,CS-MDGA 算法性能如图 10 所示. 在误码率较小  $P_b=10^{-5}$  时,3 种算法性能接近;当误码率较高为  $P_b=10^{-3}$  时,CS-MDGA 算法具有较好的性能.

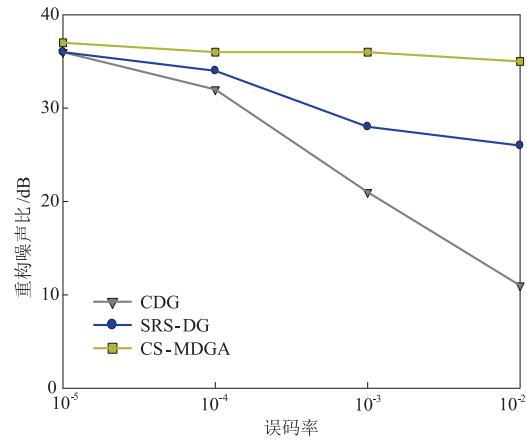


图10 CS-MDGA算法性能分析

由图 10 可知,此时 CDG 算法的重构数据信噪比为 27.33dB,高误码率对 CDG 算法具有较大的影响;SRS-DG 算法的数据信噪比为 29.72dB,由于 SRS-DG 算法是针对分组丢失设计的算法,数据分组一旦有错即丢弃,利用稀疏观测矩阵观测无错节点,导致每次数据观测信息量减少,该算法通过下一轮数据收集增加观测数来弥补丢失的数据分组,提高重构信噪比,所以本轮数据重构信噪比并不高;CS-MDGA 算法得到重构数据信噪比为 35.91dB,在一定条件下利用数据的空间相关性预测错传数据,避免丢弃使信息量减少.因此,在高误码率的无线环境中,CS-MDGA 算法没有增加额外的通信能耗,能够克服错传数据分组对数据重构的影响,具有有效性.

#### 5.4 实验四:不同网络规模下 CS-MDGA 算法性能对比

图 11 给出了在不同网络规模下 CS-MDGA 算法的性能分析,其中链路丢包率设置为 20%. 由图 11 中的 CS 节点比例变化曲线可知,随网内节点数增加,网内 CS 节点比例逐渐减少;其主要原因是在大规模传感网络中,树状拓扑路由的树枝将更茂盛,路由存在更多的细小分支,而根据混合压缩感知对节点分类的规则,通常只有处于主树枝上的节点才为 CS 节点,因此 CS 节点占据的比例也相对降低,则 CS 节点间构建的备份路径复杂性也相对降低. 由图 11 中的算法相对误差变化曲线可知,随着网内节点数的逐渐增加,CS-MDGA 算法的重构精度也逐渐增加;这是由于随着网内 CS 节点比例的降低,网内出现的丢包将更大概率不具有关联性,因此 CS-MDGA 算法的重构受备份路径可靠性保障机制性能的影响也更低,从而具有了更高的重构精度. 通

过上述仿真分析可知,CS-MDGA 算法在大规模传感网中将具有更好的性能.

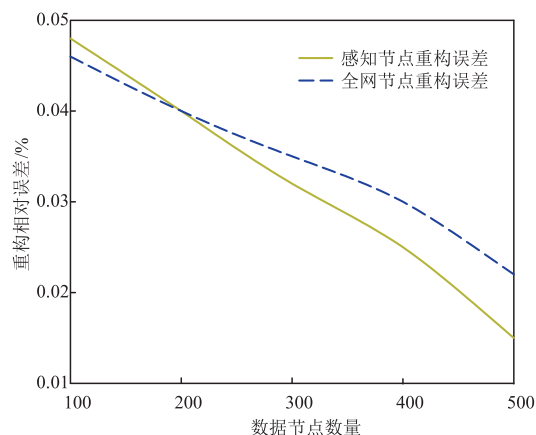


图11 不同网络规模对CS-MDGA算法性能影响

### 5.5 实验五:不同路径下 CS-MDGA 算法性能对比

在不同路径选取中,以图 1 为例,采用四组路径作为研究对象,分别是:第一组路径 L1:  $S_9 \rightarrow S_7 \rightarrow S_6 \rightarrow \text{Sink}$ ;第二组路径 L2:  $S_{10} \rightarrow S_7 \rightarrow S_6 \rightarrow \text{Sink}$ ;第三组路径 L3:  $S_8 \rightarrow S_6 \rightarrow \text{Sink}$ ;第四组路径 L4:  $S_1 \rightarrow S_3 \rightarrow S_5 \rightarrow \text{Sink}$ . 网络拓扑结构为 MST,感知节点时延为  $[0, 1\text{ms}]$ ;每条链路上的总计时延为  $[0, 10\text{ms}]$ ,上述两种时延均为随机产生.

图 12 给出了四种不同路径下的数据传输成功率对比示意图,从图 12 中可以看出,随着时间的推移,四种不同路径下的数据传输成功率均有所提升,其主要原因在于路径 L1 数据传输节点较少,对数据重构速度较快;数据在路径 L2 和 L3 传输节点与 L1 相比较多,重构速度以及路径时延较 L1 较长,因此在同一时间段内,其数据传输成功率低于 L1;对于 L4 来说,由于链路上存在不可靠性,设其权值较低,在相同的节点数量下,其数据传输的成功率小于上述三种情况.

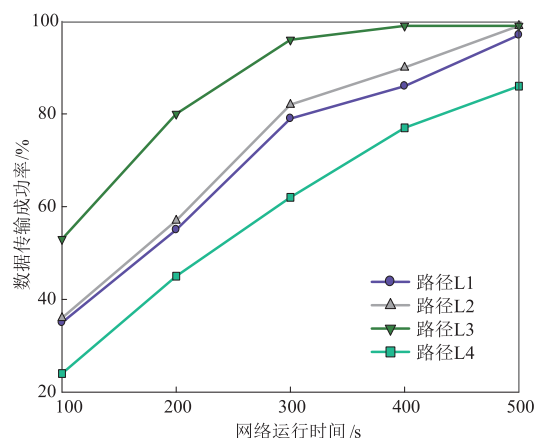


图12 四种不同路径下的数据传输成功率对比

图 13 给出了 CS-MDGA 算法与其它两种算法在网络吞吐量性能方面的对比实验.从图 13 中可以看出,随着时间的推移,三种算法的网络吞吐量均有所提升,但本文算法提升速度要快于其它两种算法,其主要原因在于,在本文采用观测矩阵通过在每次观测过程中不观测丢包节点信息,从而将 CS 数据收集在树状路由下的数据丢失问题转换为基于稀疏匹配的观测矩阵投影问题,既实现了对网内数据的大量观测采样,又避免了 Sink 端对数据收集情况的误判.因此,本文 CS-MDGA 在网络数据吞吐量方面高于其它两种算法.

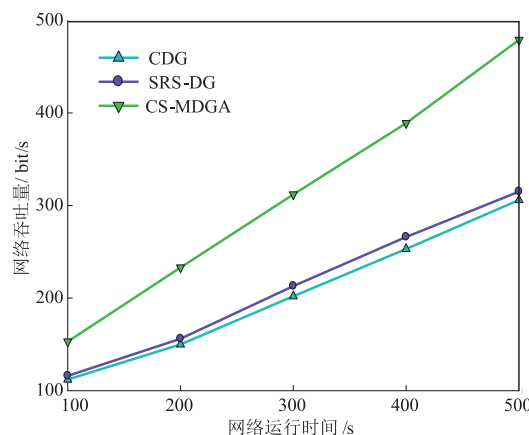


图13 不同算法下的网络吞吐量

## 6 结论

本文针对不可靠链路下 CS 数据收集问题,提出了 CS-MDGA 算法.首先,通过分析丢包对 CS 数据收集的影响,设计了 SPLM 观测矩阵.其次,本文算法将全网节点分为 TF 节点和 CS 节点,其中 TF 节点间的丢包不具有关联性,仅采用 SPLM 观测矩阵观测投影;CS 节点除利用 SPLM 矩阵观测投影外,采用基于最小能耗的备份路径保障数据传输可靠性,避免关联丢包出现.再次,证明了该观测矩阵概率趋近于“1”时,满足的等距约束条件,实现了节点间多路径路由数据的可靠交付.最后,仿真结果表明,CS-MDGA 算法在链路丢包率为 60% 的情况下,仍能保证压缩数据的有效重构,相对于其他算法,本文算法在重构精度和数据集稀疏度以及网络吞吐量等性能指标均有所提升,优化了网络资源利用率,抑制了不可靠链路丢包对 CS 数据收集的影响.下一步工作主要集中在当网络流量过大或过小时,可移动节点丢包率对算法重构精度的影响.

### 参考文献

- [1] 和志杰,杨春玲,汤瑞东.视频压缩中基于结构相似的帧间组稀疏表示重构算法研究[J].电子学报,2018,46(3):544-543.

- HE Zhi-jie, YANG Chun-ling, TANG Rui-dong. Research on structural similarity based inter-frame group sparse representation for compress video sensing[J]. *Acta Electronica Sinica*, 2018, 46(3): 544 – 553. (in Chinese)
- [2] BHOWMIK S, GIRI C. Convoy tree based fuzzy target tracking in wireless sensor networks[J]. *International Journal of Wireless Information Networks*, 2017, 24(4): 476 – 484.
- [3] 刘洲洲, 李士宁. 采用压缩感知和 GM(1,1) 的无线传感器网络异常检测方法[J]. *西安交通大学学报*, 2017, 51(2): 40 – 46.
- LIU Zhou-zhou, LI Shi-ning. An anomaly detection method for wireless sensor networks based on compressed sensing and GM(1,1)[J]. *Journal of Xi'an Jiaotong University*, 2017, 51(2): 40 – 46. (in Chinese)
- [4] YU Xiao-han, BAEK S J. Joint routing and scheduling for data collection with compressive sensing to achieve order-optimal latency [J]. *International Journal of Distributed Sensor Networks*, 2017, 13(10): 1 – 13.
- [5] ZHOU Fen, CHEN Zhen-zhong, GUO Song, LI Jie. Maximizing lifetime of data-gathering tree with different aggregation modes in WSNs[J]. *IEEE Sensors Journal*, 2016, 16(22): 8167 – 8177.
- [6] SINGH V K, SHARMA G, KUMAR M. Compressed sensing based acoustic event detection in protected area networks with wireless multimedia sensor [J]. *Multimedia Tools and Applications*, 2017, 76(18): 18531 – 18555.
- [7] CHEN Wei, WASSELL L J. Cost-aware activity scheduling for compressive sleeping wireless sensor networks [J]. *IEEE Transactions on Signal Processing*, 2016, 64(9): 2314 – 2323.
- [8] MISRA P, JU Wen, YANG Ming-rui, et al. Sparsity based efficient cross-correlation techniques in sensor networks [J]. *IEEE Transactions on Mobile Computing*, 2017, 16(7): 2037 – 2050.
- [9] WU Lian-tao, SUN P, XIAO Ming, et al. Sparse signal aloha: A compressive sensing-based method for uncoordinated multiple access [J]. *IEEE Communications Letters*, 2017, 21(6): 1301 – 1304.
- [10] 张策, 张霞, 李鸥. 不可靠链路下基于压缩感知的 WSN 数据收集算法[J]. *通信学报*, 2016, 37(9): 131 – 141.
- ZHANG Ce, ZHANG Xia, LI Ou. Compressive sensing based data gathering algorithm over unreliable links in WSN[J]. *Journal on Communications*, 2016, 37(9): 131 – 141. (in Chinese)
- [11] TALARI A, RAHNAVARD N. CStorage: decentralized compressive data storage in wireless sensor networks[J]. *Ad Hoc Networks*, 2016, 37: 475 – 485.
- [12] LUO Chong, WU Feng, SUN Jun, CHEN Chang-wen. Compressive data gathering for large-scale wireless sensor networks[A]. *The 15th Annual ACM International Conference on Mobile Computing and Networking [C]*. Beijing China: ACM, 2009. 145 – 156.
- [13] WU Xuan-gou, YANG Pan-long, JUNG T, et al. Compressive sensing meets unreliable link: sparsest random scheduling for compressive data gathering in lossy WSNs[A]. *The 15th ACM International Symposium on Mobile Ad Hoc Networking and Computing [C]*. Philadelphia USA: ACM, 2014. 13 – 22.
- [14] 韩哲, 张霞, 李鸥, 等. 面向有损链路的传感网压缩感知数据收集算法 [J]. *软件学报*, 2017, 28(12): 3257 – 3273.
- HAN Zhe, ZHANG Xia, LI Ou, et al. Data gathering algorithm based on compressive sensing under lossy WSN [J]. *Journal of Software*, 2017, 28(12): 3257 – 3273. (in Chinese)
- [15] XIAO Xue, SONG Xiao, LEI Quan. Efficient measurement method for spatiotemporal compressive data gathering in wireless sensor networks[J]. *KSI Transactions on Internet and Information Systems*, 2018, 12(4): 1618 – 1637.
- [16] 裴立业, 江桦, 李明. 一种基于能量的压缩感知稀疏度估计算法[J]. *电子学报*, 2017, 45(2): 285 – 290.
- PEI Li-ye, JIANG Hua, LI Ming. A sparsity order estimation algorithm based on measure signal's energy[J]. *Acta Electronica Sinica*, 2017, 45(2): 285 – 290. (in Chinese)
- [17] AZARNIA G, TINATI M A, REZAILI T Y. Cooperative and distributed algorithm for compressed sensing recovery in WSNs [J]. *IET Signal Processing*, 2018, 12(3): 346 – 357.
- [18] 徐佳, 冯鑫, 杨富贵, 等. 最大化最小能耗概率的移动 Sink 无线传感器网络数据收集方法 [J]. *电子学报*, 2015, 43(12): 2470 – 2475.
- XU Jia, FENG Xin, YANG Fu-gui, et al. A data collection method by maximizing minimum probability of energy consumption for mobile sink based WSNs[J]. *Acta Electronica Sinica*, 2015, 43(12): 2470 – 2475. (in Chinese)
- [19] 杨浩, 王喜玮. 基于区域化压缩感知的无线传感器网络数据收集方法 [J]. *计算机学报*, 2017, 40(8): 1933 – 1946.
- YANG Hao, WANG Xi-wei. Data gathering based on regionalized compressive sensing in WSN[J]. *Chinese Journal of Computers*, 2017, 40(8): 1933 – 1946. (in Chinese)

## 作者简介



**孙泽宇** 男,1977 年 7 月出生,吉林长春人,博士,副教授,硕士生导师,CCF 会员.2017 年 6 月毕业于西安交通大学,获工学博士学位.研究方向为传感网、物联网和并行计算.



**李传锋(通信作者)** 男,1976 年 8 月出生,河南周口人,博士,教授.2011 年毕业于华中科技大学,获工学博士学位.研究方向为智能控制、复杂网络和传感网.  
E-mail:Lcflit@163.com

**阎奔** 男,1976 年 4 月出生,河南洛阳人,博士,教授.研究方向为大数据、传感网和复杂网络.