

# 链路不可靠下稀疏投影无线传感器网络数据收集研究

吴宣够<sup>1)</sup> 储昭斌<sup>1)</sup> 郑 啸<sup>1)</sup> 王修君<sup>1)</sup> 杨盘隆<sup>2)</sup>

<sup>1)</sup>(安徽工业大学计算机科学与技术学院 安徽 马鞍山 243032)

<sup>2)</sup>(中国科学技术大学计算机科学与技术学院 合肥 230027)

**摘 要** 基于压缩感知的无线传感器网络数据收集被认为是一种有效的网内数据压缩技术,其具有网络通信代价小与均衡网络能量消耗的特点. 现有基于压缩感知数据收集方案主要针对可靠无线传感器网络环境,即链路不存在网络丢包现象. 然而现实网络运行结果表明网络丢包是一种常见现象,网络丢包会导致这些方案数据恢复性能易受干扰. 为了提高压缩感知在不可靠链路下数据的鲁棒性,该文提出一种联合路由的稀疏投影数据收集方案,其可以有效降低网络丢包对感知数据质量的影响. 该方案主要是通过两个方面技术来提高不可靠链路下数据的鲁棒性. 首先,根据路由设计稀疏投影矩阵,降低网络丢包对测量值数据的破坏;其次,进行稀疏投影矩阵限制下的低相干稀疏表示基设计,保证被压缩数据有效恢复与重构. 为了满足稀疏表示基与稀疏投影路由矩阵之间的低相干性,该文给出一种稀疏表示基学习算法. 除此之外,该文还给出分布式网络环境下稀疏随机投影数据收集实现算法. 真实网络数据实验结果表明,文中提出的稀疏表示基学习算法得到的稀疏表示基可以有效压缩网络数据,同时即便在丢包率为20%情况下网络感知数据仍能被有效恢复. 能量仿真结果表明,对比现有压缩感知数据收集方案,文中提出的方案可以有效减少网络能量消耗.

**关键词** 无线传感器网络;压缩感知;数据收集;链路不可靠;低相干性

**中图法分类号** TP393 **DOI号** 10.11897/SP.J.1016.2019.00388

## Sparse Random Projection Compressive Data Gathering in Lossy Wireless Sensor Networks

WU Xuan-Gou<sup>1)</sup> CHU Zhao-Bin<sup>1)</sup> ZHENG Xiao<sup>1)</sup> WANG Xiu-Jun<sup>1)</sup> YANG Pan-Long<sup>2)</sup>

<sup>1)</sup>(Department of Computer Science and Technology, Anhui University of Technology, Maanshan, Anhui 243032)

<sup>2)</sup>(Department of Computer Science and Technology, University of Technology and Science of China, Hefei 230027)

**Abstract** Compressed sensing is considered as a promising data sampling and compression technology compared with the traditional transformed based data compression technology. Existing research results show that compressive sensing based data gathering approaches can reduce network communication costs and balance network energy consumption in wireless sensor networks. Although many existing approaches have been proposed to reduce data transmission cost and balance network energy consumption in wireless sensor networks, these approaches mainly focus on reliable wireless sensor networks. That is, no network packet loss occurs on the network links between all sensor nodes. However, the experimental results of real sensor network system show that the network packet loss is a common phenomenon in wireless sensor networks under network packet loss retransmission mechanism. Unfortunately, our experimental results show that the

收稿日期:2017-09-24;在线出版日期:2018-04-08. 本课题得到国家自然科学基金(61672038,61402009,61502520,61502010,61402008)、安徽省高校优秀拔尖人才培养项目(gxgwfz2018017)、江苏省杰出青年自然科学基金(BK20150030)和安徽省科技重大专项(16030901060)资助. 吴宣够,男,1979年生,博士,副教授,中国计算机学会(CCF)会员,主要研究方向为无线网络、压缩感知. E-mail: wxgou@ahut.edu.cn. 储昭斌,男,1993年生,硕士研究生,主要研究方向为无线网络. 郑 啸(通信作者),男,1975年生,博士,教授,主要研究领域为计算机网络、服务计算. E-mail: xzheng@ahut.edu.cn. 王修君,男,1982年生,博士,讲师,主要研究方向为计算机算法、算法优化. 杨盘隆,男,1976年生,博士,教授,主要研究领域为无线网络、群智感知网络.

recovery data performance of these approaches would be disturbed severely by the network packet loss during wireless sensor network data gathering. Even if there is a small amount of network packet loss on wireless sensor networks, the recovery sensory data of the traditional compressive sensing based data gathering approaches will be still arbitrarily bad. To make compressed sensing meet the data gathering requirements of lossy wireless sensor network, we propose a sparse random data gathering approach with packet loss consideration. Our proposed approach can effectively cope with the impact of network packet loss on the recovery quality of sensory data in wireless sensor network data gathering. In our approach, we mainly exploit two aspects of techniques to achieve the robustness of sensory data in lossy wireless sensor networks. First, unlike the existing random sparse projection approaches, we design our sparse projection matrix according to the network routing in wireless sensor networks. Our designed sparse projection matrix can effectively reduce the damage of gathered measurements caused by packet loss during wireless sensor network data gathering. Second, in order to satisfy the conditional requirements of compressed sensing theory for sparse projection matrix and sparse representation basis, we design a low-coherence sparse representation basis with sparse projection matrix constraint. In other words, our designed sparse projection matrix and sparse representation basis guarantee that the sensory data can be reconstructed successfully even if some packets are lost during wireless sensor network data gathering. Meanwhile, we present a sparse representation learning algorithm by the historical sensory data, which satisfies the low-coherence between the sparse representation matrix and the sparse projection routing matrix. Furthermore, we also give the corresponding distributed data gathering algorithms of our approach in wireless sensor network. The experimental results of real network data show that the sparse representation basis obtained by our proposed representation base learning algorithm can compress the network sensory data effectively, the proposed joint sparse projection data gathering approach can recover and reconstruct the sensory data accurately, even when the network packet loss rate is up to 20% in wireless sensor networks. At the same time, the simulation results also show that the proposed scheme can effectively reduce network energy consumption compared with the existing compressed sensing data gathering approaches in wireless sensor network data gathering.

**Keywords** wireless sensor networks; compressive sensing; data gathering; link unreliable; low-coherence

## 1 引言

无线传感器网络(Wireless Sensor Networks, WSNs)已广泛用于各种数据收集应用,如环境监测、医疗护理、军事安全、目标跟踪等相关领域<sup>[1]</sup>. 由于传感器节点受能量、通信能力、带宽等资源的限制,其难以大规模部署和长时间运行. 如何减少网络能量消耗和延长网络运行时间是大规模 WSN 研究的关键问题之一<sup>[2]</sup>. 而现有研究表明传感器网络能量消耗主要集中在数据包的发送上<sup>[3]</sup>. 因此,在数据收集过程中减少网内数据包的传输数目,可以有效延长网络运行时间. 同时,无线传感器网络常被用于

监控同一区域的物理量,通常这些监控数据具有一定的时空相关性,这使得利用网内数据压缩技术来减少传输数据包的数目成为可能<sup>[4]</sup>.

压缩感知(Compressive Sensing, CS)作为一种新的数据采样与压缩理论<sup>[5]</sup>,被用于进行 WSN 数据收集时具有网络通信代价小和均衡网络能量消耗的特点<sup>[6]</sup>. 因此,基于压缩感知的数据收集技术可以有效延长网络运行时间. 目前,已有多种基于压缩感知的 WSN 数据收集方案被提出<sup>[7-12]</sup>. 如 Luo 等人在文献[7]中首次将压缩感知理论应用到大规模 WSN 数据收集,对比传统数据收集技术,其具有平衡网内能量消耗的优点. Luo 等人在文献[8]中指出稠密投影数据收集并不一定能带来网络通信代价

的减少,同时作者提出一种混合压缩感知数据收集方案.针对降低网络通信代价方面,已存在利用稀疏投影进行 WSN 数据收集的诸多工作.如 Wang 等人在文献[9]中提出一种分布式随机稀疏投影数据收集方案,减少参与测量值的传感器节点数目.吴晓培等人在文献[10]中提出一种在单个节点上的随机稀疏采样数据收集方法,从而减少数据收集代价.王新兵等人在文献[11]中提出一种用于分布式网络环境下的随机游走数据收集方法,构造稀疏投影矩阵.我们在文献[13]中提出最稀疏随机投影数据收集方案,实现投影矩阵的每一行只有一个非零元素,可以进一步减少单个压缩感知测量值的收集代价.现有这些基于压缩感知的数据收集方法主要从数据压缩性和投影矩阵的稀疏性两个方面考虑,实现数据压缩的同时降低网络通信代价.

然而,现有基于压缩感知的无线传感器网络数据收集研究主要集中在可靠网络环境下,即网络链路之间不存在网络丢包现象.但真实无线传感器网络是一种易错网络,网络丢包普遍存在.例如,GreenOrbs 系统数据显示即便存在丢包重传机制,网络丢包率有时仍可高达 20%<sup>[14]</sup>.此外,我们前期研究结果表明现有基于压缩感知数据收集方法的性能易受网络丢包影响.例如,在 2% 的网络丢包率下,现有数据收集方法就难以实现对网络数据的有效恢复<sup>[15]</sup>.尽管我们也提出相应的稀疏随机投影数据收集方法,且该方法可以降低网络丢包对数据恢复性能的影响,但我们并没有给出有效的稀疏表示基设计方法,其缺乏通用性.

为了在不可靠网络环境下,基于压缩感知的数据收集技术能实现对网络数据的有效重构,本文提出一种联合路由的稀疏投影数据收集方案,其可以有效应对网络丢包对感知数据质量的影响.由于现有方法是直接将投影矩阵与网络路由被动匹配,导致参与单个测量值节点过多,造成压缩感知测量值易受网络丢包影响.本文方案主要是通过两个方面技术来实现提高不可靠链路下数据的鲁棒性.首先,根据路由设计稀疏投影矩阵,从而降低网络丢包对测量值数据的破坏;其次,进行稀疏投影矩阵限制下的低相干稀疏表示基设计,保证被压缩数据有效恢复与重构.为了满足稀疏表示基与稀疏投影路由矩阵之间的低相干性,本文给出一种稀疏表示基学习算法.除此之外,本文还给出分布式网络环境下稀疏随机投影数据收集实现算法.总结起来,本文具体贡献如下:

(1) 本文提出一种联合路由的自适应稀疏投影

数据收集方案,该方案是在给定稀疏投影矩阵下进行稀疏表示基设计.因此,投影矩阵可以与路由主动匹配,从而减少网络丢包对测量值的破坏.

(2) 本文提出一种限制条件下低相干性稀疏表示基学习方法,使得学习到的表示基既有良好的数据压缩性能又能与稀疏路由投影矩阵满足低相干性,从而保证感知数据能被有效重构.

(3) 本文给出具体的分布式网络下联合路由稀疏投影数据收集算法,并进行网络传输代价理论分析.分析结果表明本文提出的方法网络通信代价为  $O(hk\log(N))$ .

(4) 通过 CitySee 系统真实数据实验,实验结果表明本文提出的数据收集方案在网络丢包率高达 20% 的情况下仍能对感知数据进行有效恢复.能量仿真结果显示,对比现有数据收集方案,本文提出的方案能有效减少网络能量消耗.

本文第 2 节总结相关工作;第 3 节给出联合路由与稀疏投影数据收集问题的描述;第 4 节给出具体的低相干稀疏表示基学习算法;第 5 和 6 节分别给出具体网络算法实现和网络通信代价分析;最后,第 7 和 8 节分别给出实验评价结果和本文总结.

## 2 相关工作

### 2.1 可压缩 WSN 数据收集

无线传感器网络常用于监控同一区域的物理量,其收集的数据通常具有一定程度的时空相关性.利用数据压缩技术来减少网络中传输的数据包数目,可以降低网络资源消耗<sup>[4]</sup>.由于压缩感知技术在网络数据收集过程中具有减少网络通信代价和平衡网内能量消耗的特点,基于压缩感知的无线传感器网络数据收集被广泛研究<sup>[7-18]</sup>.

现有基于压缩感知的数据收集技术主要是将投影矩阵分布到网络各节点之中,在数据收集过程中进行感知数据的线性组合.根据所使用投影矩阵类型可以将其分为稠密投影和稀疏投影.如 Luo 等人在文献[7]中首次将压缩感知理论应用到大规模 WSN 数据收集中,理论分析与实验结果均表明其具有均衡网内能量消耗特点.在文献[16]中,Jun 等人根据收集数据的时空关系提出一种最小能量消耗的数据收集方案,其可以有效减少网络能量消耗.在文献[17]中,Markus 等人提出序列压缩感知 WSN 数据收集方案,其能有效减少网络传输代价同时保证数据恢复质量.以上方案都是利用稠密投影矩阵

进行数据收集,且单个测量值的收集需要所有节点参与。

针对稀疏投影数据收集,其主要是通过减少参与单个测量值的传感器节点数目来降低网络通信代价。如在文献[8]中,Luo 等人提出一种混合形式的稀疏投影数据收集策略,以减少部分测量值的参与节点数目。在文献[9]中,Wang 等人从理论上证明了稀疏投影也可以实现压缩数据的有效恢复,并给出测量值数目与投影矩阵稀疏程度之间的关系。在文献[10]中,吴晓培等人利用稀疏投影实现单个传感器节点感知数据的高精度重构,并首次尝试根据稀疏投影矩阵进行稀疏表示基设计,以满足应用场景的需求。在文献[11]中,王新兵等人利用随机游走方法实现分布式环境下的 WSN 数据收集,并给出相应的数据恢复性能的理论证明。Liu 等人在文献[12]中提出一种基于随机投影的数据模型估计算法,在每个测量值中只需少量的几个传感器节点的值,就能实现网络通信代价的降低。在文献[13]中,我们利用每一行只有一个非零元素的投影矩阵进行感知数据的收集,将参与单个测量值收集的节点数目降到最低,从而有效减少数据收集过程中的通信代价。

现有工作无论是稠密投影还是稀疏投影,主要是针对可靠网络环境下的压缩感知数据收集。然而在不可靠网络环境下,现有方案的数据恢复性能易受网络丢包影响。同时,我们前期的研究表明,即使在 2% 的网络丢包率下,现有基于压缩感知的数据收集方案无法实现对感知数据的有效恢复<sup>[15]</sup>。针对不可靠网络环境,文献[18]提出一种提高稀疏投影的数据收集方案,从而改善网络丢包对测量值的破坏。我们在文献[14]中提出一种最稀疏随机投影数据收集方案来降低网络丢包对数据恢复性能的影响。尽管我们前期工作和文献[18]已开展这方面的工作,但并没有给出有效的感知数据稀疏表示基设计方法,其不具备通用性。

## 2.2 信号稀疏表示

信号稀疏表示的目的是在给定的超完备字典中用尽可能少的原子来表示信号,从而可以获得更为简洁的信号表示方式。目前,信号稀疏表示技术已经广泛用于信号处理、图像处理、压缩感知等多个领域。且已存在多种稀疏表示基学习算法用于对信号的稀疏表示,如 MOD<sup>[19]</sup>、K-SVD<sup>[20]</sup> 等经典信号稀疏表示算法<sup>[21]</sup>。但这些算法主要是从编码的角度来考虑如何提高编码压缩效率,缺少提高解码效率方

面的考虑。

近年来,已存在针对压缩感知编码和解码方面进行稀疏表示基学习相关算法。例如,Mailh 等人在文献[22]中提出 INK-SVD,该方法可以从一个给定稀疏表示基中获得一个低相干性的稀疏表示基。Barchiesi 等人在文献[23]中通过迭代投影和旋转方法进行稀疏表示基学习。该方法可以获得比 INK-SVD 更低相干性的稀疏表示基。在文献[24]中,Bai 等人提出一种迭代最优稀疏表示基学习算法。该方法既可以获得最优稀疏表示基还可以获得最优投影矩阵。Ding 等人在文献[25]中提出根据张量压缩感知进行联合最优感知矩阵和表示基学习。

尽管现有这些方法可以进行低相干性稀疏表示基学习,但他们主要集中在稠密投影矩阵情况下的最优稀疏表示基学习。也就是说,这些方法并没有考虑在分布式网络环境中,单个测量值的收集代价是不可以被忽略的。在分布式无线传感器网络环境中,投影矩阵的稀疏程度直接决定网络数据收集的通信代价和能量消耗。因此,这些方法并不适用于分布式无线传感器网络数据收集。所以,本文主要是针对给定稀疏投影矩阵下的最优稀疏表示基学习问题的研究。

## 3 问题描述

### 3.1 基于压缩感知的数据收集

假设有一个无线传感器网络包含  $N$  个节点,这些节点分别表示为  $sr_1, sr_2, \dots, sr_N$ ,它们在某一时刻的采样值分别表示为  $x_1, x_2, \dots, x_N$ 。在基于压缩感知的数据收集过程中,每个节点将自己的采样值分别乘上  $M$  个权值发给其下一跳节点,最后汇集点(sink)收到  $M$  个测量值(感知数据的线性组合值)。基于压缩感知的数据收集过程可以用下面数学形式来进行表示:

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_M \end{bmatrix} = \begin{bmatrix} \phi_{11} & \phi_{12} & \cdots & \phi_{1N} \\ \phi_{21} & \phi_{22} & \cdots & \phi_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \phi_{M1} & \phi_{M2} & \cdots & \phi_{MN} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix} \quad (1)$$

即  $\mathbf{y} = \Phi \mathbf{x}$ ,其中  $\mathbf{y} = [y_1, y_2, \dots, y_M]^T$  为 sink 收集到的测量值向量, $\mathbf{x} = [x_1, x_2, \dots, x_N]^T$  为  $N$  个节点的采样值, $\Phi$  为  $M \times N$  的投影矩阵(权值系数矩阵), $\Phi$  通常为高斯随机矩阵或等概率的  $\pm 1$  Bernoulli 矩阵。

为了简单说明,图 1 给出了链式多跳无线传感器网络拓扑下的基于压缩感知数据收集示例。如果

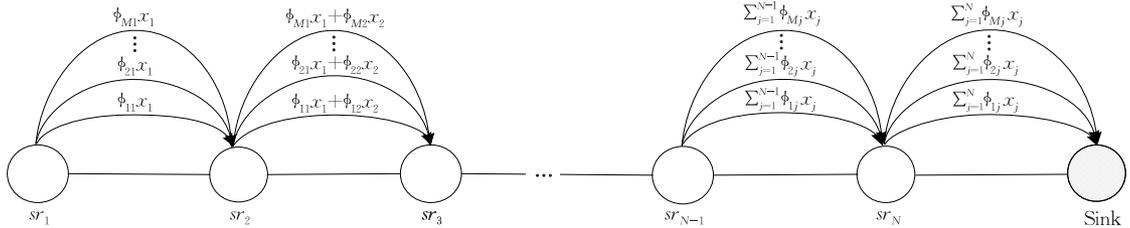


图 1 基于压缩感知的数据收集

$x$  为可压缩信号, 即  $x$  可表示为  $x = \Psi s$  且  $s$  为  $k$  稀疏信号 ( $k \ll N$ ), 则可以通过  $y$  来重构  $x$ .  $x$  可以通过下面优化问题求解:

$$\hat{s} = \operatorname{argmin} \|s\|_1 \quad \text{s. t.} \quad \|y - \Phi \Psi s\|_2 \leq \epsilon \quad (2)$$

其中  $\epsilon$  为恢复所允许的噪音界. 目前已存在许多算法对问题(2)求解, 如 OMP 算法<sup>[26]</sup>、BP 算法<sup>[27]</sup>、CoSaMP 算法<sup>[28]</sup>等.

### 3.2 稀疏表示基学习

在无线传感器网络监控中, 由于其监控的感知数据通常具有一定程度的时空相关性. 因此, 大部分无线传感器网络数据被认为是可压缩数据. 目前, 在基于压缩感知的数据收集研究中, 绝大部分假设感知数据在通用表示基下可以转化为稀疏信号, 如离散余弦变换表示基 (Discrete Cosine Transform, DCT)、离散傅立叶变换表示基 (Discrete Fourier Transform, DFT)、离散小波变换表示基 (Discrete Wavelet Transform, DWT) 等. 然而, 在很多情况下感知数据在通用表示基下并非是稀疏信号, 需要对历史感知数据进行学习得到特定的稀疏表示基. 假设有一组收集到的数据  $X = [x_1, x_2, \dots, x_n]$  作为稀疏表示基的学习数据, 则稀疏表示基学习可表示为

$$\min_{\Psi, C} \|X - \Psi C\|_F^2, \quad \text{s. t.} \quad \forall i, \|C_i\| \leq k \quad (3)$$

其中,  $\|\cdot\|_F$  为矩阵 Frobenius 范数,  $\Psi$  是要寻找的稀疏表示矩阵,  $k$  为稀疏度, 矩阵  $C$  中的每一列都是  $k$  稀疏向量. 通过求解优化问题(3)就可以得到稀疏表示基  $\Psi$ . 对问题(3)的求解有 K-SVD 等表示基学习方法<sup>[20]</sup>.

问题(3)是从给定数据中求解最稀疏表示基问题. 而在基于压缩感知的数据恢复过程需要稀疏表示基与投影矩阵满足等距离约束性质或低相干性, 否则会存在以下两个方面的问题: 一是被重构的感知数据质量无法得以保证; 二是需要过多的测量值来进行感知数据的重构. 而在无线传感器网络数据收集过程中, 数据收集代价是由单个测量值传输代价和所需测量值数目决定的.

### 3.3 联合路由的稀疏投影数据收集

由式(1)可知, 在基于压缩感知的数据收集过程中, 投影矩阵  $\Phi$  每一行为单个测量值的收集路由过程. 即  $\Phi$  中非零元素个数为参与测量值的节点数目. 如果减少单个测量值的收集代价, 则需要减少参与单个测量值的节点数目, 也就是需要将  $\Phi$  变得更加稀疏. 正如我们前期工作所提出的最稀疏随机投影压缩感知数据收集, 实现投影矩阵的每一行只有一个非零元素, 有效实现链路不可靠情况下的数据收集<sup>[15]</sup>. 但现有工作并没有考虑网络中的路由情况, 同时也没有给出满足稀疏投影矩阵的表示基学习方法.

接下来, 我们将给出联合路由的稀疏投影数据收集问题模型, 并给出符合路由稀疏投影矩阵下的稀疏表示基的求解问题. 假设路由矩阵为  $R$ , 如果节点  $i$  的数据包需要经过节点  $j$  来转发才能到达 sink 节点, 则  $R_{ij} = 1$ , 否则  $R_{ij} = 0$ .  $B$  为随机权值稀疏矩阵, 通常可以是元素为  $\pm 1$  的 Bernoulli 矩阵. 因此, 投影矩阵  $\Phi$  可以表示为  $\Phi = R \circ B$ , 其中  $\circ$  为矩阵元素乘积运算, 即  $\Phi_{ij} = R_{ij} \cdot B_{ij}$ . 由于  $R$  为稀疏矩阵, 所以投影矩阵  $\Phi$  也是稀疏矩阵. 因此, 数据收集过程可以表示如下:

$$y = P_\Omega(\Phi)x = P_\Omega(\Phi)\Psi s \quad (4)$$

其中,  $P_\Omega(\cdot)$  为从给定矩阵中随机均匀选取行向量,  $\Psi$  为  $x$  的稀疏表示基. 根据压缩感知理论可知,  $\Phi$  与  $\Psi$  之间的相干性应该尽可能的低, 才能保证被压缩数据能被有效重构<sup>[26]</sup>. 根据以上定义和限制性条件, 联合路由稀疏投影的表示基求解问题可以表示如下:

$$\begin{aligned} \text{(P)} \quad F(\Psi, S) = & \min_{\Psi, S} \|X - \Psi S\|_F^2 + \alpha \\ & \|\Phi(X - \Psi S)\|_F^2 + \beta \\ & \|\Psi^T \Phi^T \Phi \Psi - I\|_F^2 \\ \text{s. t.} \quad & \forall i, \|S_i\|_0 \leq k \end{aligned}$$

其中,  $\|X - \Psi S\|_F^2$  表示尽可能稀疏化  $X$ ,  $\|\Phi(X - \Psi S)\|_F^2$  表示采样数据的误差尽可能小, 即式(2)中的噪音界尽可能小,  $\|\Psi^T \Phi^T \Phi \Psi - I\|_F^2$  表示投影矩

阵和表示基尽可能的不相干,  $\mathbf{I}$  为单位矩阵,  $\alpha$  和  $\beta$  为权衡因子. 问题(P)是非凸优化问题, 且该问题是一个难解问题. 在接下来部分, 我们给出问题(P)的具体求解算法.

## 4 低相干稀疏表示基学习

尽管已存在多种针对压缩感知的最优稀疏表示基学习研究, 但现有研究主要集中在最优投影下的最优稀疏表示基学习, 如文献[22-24]. 而本文研究的是投影矩阵限制性下的稀疏表示基学习. 由于问题(P)是非凸优化难解问题, 针对问题(P), 本文采用交替迭代方法求解, 其包含两个主要步骤: 稀疏逼近和限制条件下的低相干稀疏表示基更新.

在稀疏逼近阶段, 采取固定稀疏表示基  $\Psi$  求解  $\mathbf{X}$  在  $\Psi$  稀疏变换下的  $\mathbf{S}$ , 其过程如下:

$$g(\mathbf{S}) = \min_{\mathbf{S}} \|\mathbf{X} - \Psi\mathbf{S}\|_F^2 \quad (5)$$

$$\text{s. t. } \forall i, \|\mathbf{S}_i\|_0 \leq k.$$

针对问题(5), 已存在多种求解算法, 我们选择 K-SVD 算法求解其对应的稀疏信号.

在稀疏表示基更新阶段, 固定稀疏系数  $\mathbf{S}$ , 求解低相干稀疏表示基  $\Psi$ , 其求解过程如下:

$$\mathcal{L}(\Psi) = \min_{\Psi} \|\mathbf{X} - \Psi\mathbf{S}\|_F^2 + \alpha \|\Phi(\mathbf{X} - \Psi\mathbf{S})\|_F^2 + \beta \|\Psi^T \Phi^T \Phi \Psi - \mathbf{I}\|_F^2. \quad (6)$$

针对问题(6), 我们提出投影矩阵限制下的低相干稀疏表示基学习算法. 因为问题(6)是凸优化问题, 其可以通过求极值来求解.

为了便于计算, 假设  $\mathcal{L}_1(\Psi) = \|\mathbf{X} - \Psi\mathbf{S}\|_F^2$ ,  $\mathcal{L}_2(\Psi) = \|\Phi(\mathbf{X} - \Psi\mathbf{S})\|_F^2$  和  $\mathcal{L}_3(\Psi) = \|\Psi^T \Phi^T \Phi \Psi - \mathbf{I}\|_F^2$ . 因此, 可以得到  $\partial \mathcal{L}_1(\Psi) / \partial \Psi = 2(\Psi\mathbf{S}\mathbf{S}^T - \mathbf{X}\mathbf{S}^T)$ ,  $\partial \mathcal{L}_2(\Psi) / \partial \Psi = -2\alpha \Phi^T \Phi (\mathbf{X} - \Psi\mathbf{S})\mathbf{S}$  和  $\partial \mathcal{L}_3(\Psi) / \partial \Psi = 4\beta \Phi^T \Phi \Psi ((\Phi\Psi)^T \Phi\Psi - \mathbf{I})$ .

令  $\partial \mathcal{L}(\Psi) / \partial \Psi = 0$ , 则有:

$$2(\Psi\mathbf{S}\mathbf{S}^T - \mathbf{X}\mathbf{S}^T) - 2\alpha \Phi^T \Phi (\mathbf{X} - \Psi\mathbf{S})\mathbf{S}^T + 4\beta \Phi^T \Phi \Psi ((\Phi\Psi)^T \Phi\Psi - \mathbf{I}) = 0 \quad (7)$$

由式(6)可得

$$\Psi = (\alpha \Phi^T \Phi (\mathbf{X} - \Psi\mathbf{S})\mathbf{S}^T + 2\beta \Phi^T \Phi \Psi ((\Phi\Psi)^T \Phi\Psi - \mathbf{I}) + \mathbf{X}\mathbf{S}^T) (\mathbf{S}\mathbf{S}^T)^{-1} \quad (8)$$

根据式(7)可得如下递推公式:

$$\Psi_{(k+1)} = (\alpha \Phi^T \Phi (\mathbf{X} - \Psi_{(k)}\mathbf{S})\mathbf{S}^T + 2\beta \Phi^T \Phi \Psi_{(k)} ((\Phi\Psi_{(k)})^T \Phi\Psi_{(k)} - \mathbf{I}) + \mathbf{X}\mathbf{S}^T) (\mathbf{S}\mathbf{S}^T)^{-1} \quad (9)$$

根据递推公式(9), 则对问题(P)的求解可以用如下

算法 1 来表示.

**算法 1.** 交替迭代求解低相干稀疏表示基.

输入:  $\mathbf{R}, \mathbf{B}, Iter$

输出:  $\Psi$

1.  $\Phi = \mathbf{R} \cdot \mathbf{B}$ ;
2.  $\Psi_{(0)} = \text{randn}(N)$ ;
3.  $\Psi_{(0)} = \text{fun\_normalized}(\Psi_{(0)})$ ;
4.  $i = 0$ ;
5. while ( $i \leq Iter$ ) {
6.  $\mathbf{S}_{(i)} = \text{fun\_KSVD}(\mathbf{X}, \Psi_{(i)}, k)$ ;
7.  $\Psi_{(i+1)} = \text{fun\_update\_Psi}(\Psi_{(i)})$ .
8.  $\Psi_{(i+1)} = \text{fun\_normalized}(\Psi_{(i+1)})$
9.  $i = i + 1$ ;
10. }
11.  $\Psi = \Psi_{(i+1)}$ ;

在算法 1 中, 输入为稀疏路由矩阵  $\mathbf{R}$  和  $\pm 1$  Bernoulli 矩阵  $\mathbf{B}$ . 第 1 行  $\Phi$  为联合路由的稀疏投影矩阵. 第 2~3 行为稀疏表示基的初始化值, 其通过  $\text{randn}(N)$  产生一个  $N \times N$  的随机矩阵. 第 5~10 行为交替迭代稀疏表示基求解过程. 其中, 第 6 行为通过  $\text{fun\_KSVD}(\mathbf{X}, \Psi_{(i)}, k)$  函数来实现 K-SVD 算法, 并求解信号  $\mathbf{X}$  对应的稀疏信号  $\mathbf{S}_{(i)}$ . 第 7 行通过函数  $\text{fun\_update\_Psi}(\Psi_{(i)})$  函数实现等式(9)的迭代过程. 第 8 行通过  $\text{fun\_normalized}(\Psi_{(i+1)})$  函数实现归一化更新后的稀疏表示基. 第 11 行最后一次迭代得到的表示基为输出的稀疏表示基.

## 5 路由联合稀疏投影数据收集

以上两节中, 本文仅给出联合路由的稀疏投影压缩感知数据收集模型和低相干稀疏表示基学习算法, 并没有给出在分布式传感器网络中具体实现. 接下来, 我们将给出具体的联合路由稀疏投影数据收集的分布式网络实现.

在给出具体算法实现之前, 我们对无线传感器网络和数据收集过程做如下假设: (1) 网络中所有传感器节点负责数据的采样、压缩、转发与上报, 汇集点/服务器负责压缩数据的收集与恢复; (2) 网络数据收集以轮为单位, 在每轮时间内由汇集点收集齐所有节点上报的压缩数据; (3) 汇集点/服务器拥有网络中所有节点的路由信息. 为了便于表述, 本文不区分汇集点与网络数据服务器. 通常汇集点只负责网络信息的汇集并转发给服务器, 服务器执行数据的其它操作. 针对假设(3), 在真实网络环境中汇集点可能并不拥有各节点路由信息, 但在节点位置固定的传感器网络中, 网络路由信息并不经常改变.

如果节点路由信息发生改变,可以将改变的路由信息上报给汇集点或服务器,从而使得汇集点拥有所有节点路由信息.针对网络路由动态变化情况,我们将在本工作之后的下一步工作中进行考虑.

由压缩感知理论可知,在给定待收集数据的表示基和投影矩阵的情况下,待收集数据的恢复性能

取决于收集到的测量值数目.而网络中待收集数据的稀疏度是未知的,且可能会随时间推移而发生变化.因此,汇集点需要收集的测量值数目和汇集点所使用的稀疏表示基也是动态的.根据以上分析,本文提出一种自适应随机稀疏投影网络数据收集方案,其结构框图如图 2 所示.

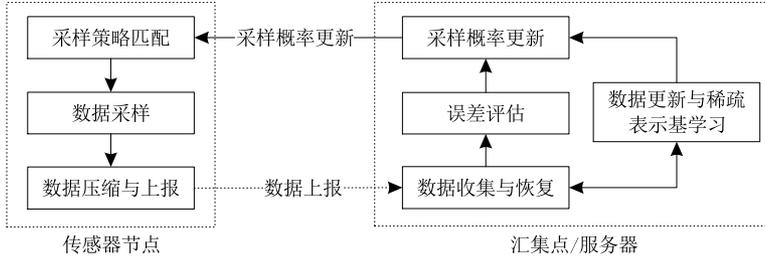


图 2 自适应随机稀疏投影数据收集框图

在本文提出的数据收集方案中,汇集点负责稀疏表示基学习、全网数据恢复、网络采样率的更新与分发.对于传感器节点,将其分为测量值发起节点和测量值参与节点两种类型.在每轮数据收集过程中,各节点以相同概率作为测量值发起节点.如果一个节点不是测量值发起节点,但在发起节点数据上报的路由链路上,则该节点为测量值参与节点.无论是测量值发起节点还是参与节点都需要进行网络数据的采样、压缩与上传操作.传感器节点数据收集流程如图 3 所示,其具体执行过程如算法 2 和算法 3 所示,汇集点的执行过程如算法 4 所示.

2. if ( $P_r \leq P_s$ ) {
3. sampling  $x_i$ ;
4.  $y_j = \phi_{ji} x_i$ ; //calculating the compressed value
5.  $s\_id = i$ ;
6. send ( $y_j, s\_id$ ) to the next hop node;
7. }
8. else
9. sleep( );

**算法 3.** 测量值参与节点数据收集算法(节点  $t$ ).

1. if (received ( $y_j, s\_id$ )) {
2. sampling  $x_t$ ;
3.  $y_j = y_t + \phi_{jt} x_t$ ; //calculating the compressed value
4. send ( $y_j, s\_id$ ) to the next hop node;
5. }

**算法 4.** 汇集点数据收集算法.

输入:  $\Phi, \Psi, \epsilon_{ub}, \epsilon_{lb}, \Delta P$

输出:  $\hat{x}$

1.  $P_\Omega \leftarrow 0$ ; //initializing measurement matrix
2.  $y \leftarrow 0$ ; //initializing measurement vector
3. do{
4. if (received data packet ( $y_j, s\_id$ )) {
5.  $P_\Omega \leftarrow s\_id$ ;
6.  $y \leftarrow y_j$ ; }
7. } until the round timeout;
8.  $\Theta = P_\Omega (\Phi) \Psi$ ;
9.  $(\hat{s}_\epsilon) = \text{fun\_CS\_rec}(\Theta, y)$ ; //CS recovery
10.  $\hat{x} = \Psi \hat{s}_\epsilon$ ;
11.  $P_s = |P_\Omega| / N$ ;
12. if ( $\epsilon > \epsilon_{ub}$ ) {
13.  $P_s = P_s + \Delta P$ ;
14. broadcast  $P_s$  to all sensors; }
15. else if ( $\epsilon < \epsilon_{lb}$ ) {
16.  $P_s = P_s - \Delta P$ ;

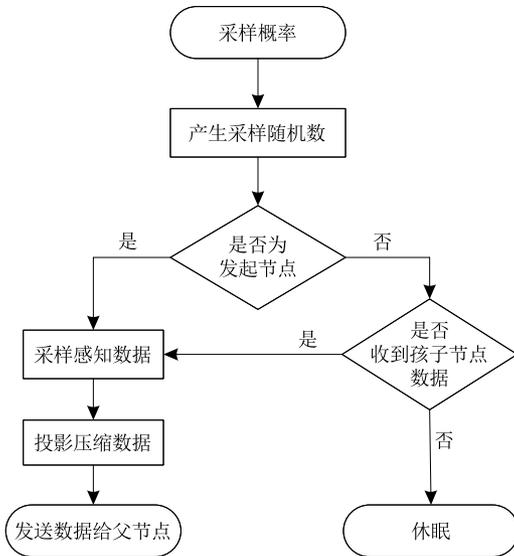


图 3 传感器节点数据收集流程图

**算法 2.** 测量值发起节点数据收集算法(节点  $i$ ).

输入: 采样概率  $P_s$

输出: 测量值  $y_j$

1.  $P_r \leftarrow \text{rand}()$ ; //generating a random value (0,1)

- 17. broadcast  $P_s$  to all sensors;
- 18. }

算法 2 是测量值发起节点数据收集算法. 网络中的每个节点首先通过  $rand()$  函数产生一个  $(0, 1)$  之间的随机数  $P_r$ . 当  $P_r \leq P_s$  时, 该节点作为测量值的发起节点, 其中  $P_s$  为发起节点的概率阈值, 由整个网络采样概率决定. 测量值发起节点将自己的采样值乘上对应投影矩阵中的权值, 并将得到的测量值与自己的节点号发给其下一跳节点.  $send(y_j, s\_id)$  函数来实现此功能, 其中  $y_j$  为第  $j$  个测量值,  $s\_id$  为当前节点的 ID 号. 针对非测量值发起节点, 如果其在测量值发起节点的路由链路上, 则为测量值参与节点. 算法 3 为测量值参与节点算法, 这些节点只要将自己的感知数据乘上对应的投影矩阵权值系数, 然后与收到的压缩数据值相加, 再转发给其下一跳节点.

在汇集点执行操作的算法 4 中, 假设汇集点的稀疏表示基是由历史感知数据离线学习得到. 经过一段时间采样数据积累后, 再进行稀疏表示基学习样本更新, 同时更新数据恢复过程中的稀疏表示基. 在算法 4 中,  $\Phi$  和  $\Psi$  分别为网络中的投影矩阵和学习到的稀疏表示基.  $\varepsilon_{ub}$  和  $\varepsilon_{lb}$  分别为收集数据的恢复误差上界和下界.  $\Delta P$  动态调整测量值发起节点的步长概率. 第 1~2 行表示在每轮开始时初始化投影矩阵和收集到的测量值向量. 第 3~7 行是收集保存上传来的测量值并记录对应的测量值发起点 ID 号. 8~10 行是进行压缩感知解码恢复过程, 即重构网络监控数据  $\hat{x}$ , 其中  $\Theta = P_\Omega(\Phi)\Psi$  计算投影感知矩阵,  $P_\Omega$  为随机采样操作. 函数  $fun\_CS\_rec$  为压缩感知重构算法, 本文采用 OMP 算法. 第 11 行计算当前轮发起测量值节点概率. 第 12~17 行是根据数据恢复精度来动态调整网络发起测量值的节点概率.

## 6 网络通信性能分析

在本小节, 我们将给出本文提出的数据收集方案网络通信代价分析. 为了体现本文提出算法的通信性能, 我们选择另外 3 种数据收集方案与本文提出的方案进行对比, 即不进行数据压缩收集、随机稠密投影数据收集<sup>[7]</sup>和随机稀疏投影数据收集<sup>[9]</sup>. 为了便于进行量化与分析, 假设无线传感器网络包含  $N$  个节点, 网络待收集数据的稀疏度为  $k$ . 数据收集过程中采用树形路由进行多跳数据传输策略, 且节点到汇集点的平均跳数为  $h$ .

因此, 针对不压缩数据收集网络传输代价为  $TC_{non} = O(hN)$ . 针对稠密投影数据收集, 单个测量值收集传输代价为  $O(N)$ . 由压缩感知理论可知, 稀疏度为  $k$  的数据至少需要收集  $O(k \log(N))$  个测量值才能对感知数据进行有效恢复, 则稠密投影数据收集的网络传输代价为  $TC_{drp} = O(N \cdot k \log(N)) = O(kN \log(N))$ . 针对随机稀疏投影数据收集, 由于稀疏投影矩阵中每行有  $\log(N)$  个随机非零元素, 则单个测量值收集代价为  $O(h \cdot \log(N))$ . 根据文献[8]可知, 数据收集网络传输代价为  $TC_{srp} = O(h \log(N) k^2 \log(N)) = O(hk^2 \log^2(N))$ . 对于本文提出的方案, 由于每一个测量值收集过程中只有一个发起节点, 其它参加节点都在其路由路径上, 所以单个测量值的收集代价仅为  $O(h)$ . 收集齐  $O(k \log(N))$  个测量值所需的网络传输代价为  $TC_{our} = O(hk \log(N))$ .

在图 4 中, 我们给出树形网络拓扑下的 3 种数据收集策略示例, 节点链路上的数字为网络中传输的数据包数目. 图 4(a) 展示的是不进行任何压缩情景下的数据收集情况, 每个节点除发送自己数据外还需转发其孩子节点发来的数据. 所以, 在离 sink 节点越近的节点其链路上传输的数据包数目越多, 全网需要传输 47 个数据包. 图 4(b) 展示的是稠密投影数据收集, 如果 14 个节点的网络需要收集 4 个测量值, 则每条链路上需要传输 4 个数据包, 全网需

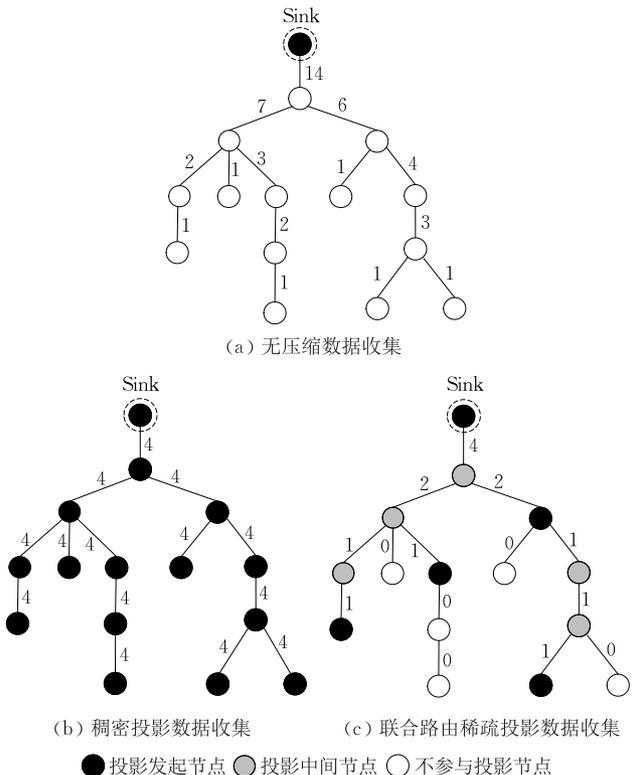


图 4 树形网络路由下的不同数据收集通信代价对比

要传输 56 个数据包,从图 4(a)与图 4(b)对比可以看出,虽然稠密投影数据收集可以均衡网络传输代价,但网络传输总的数据包数目会超过不压缩数据收集策略.图 4(c)展示的是本文提出的联合路由稀疏投影数据收集策略,其只要随机选择 4 个测量值发起节点,在各发起节点的路由路径上的节点参与其测量值收集,所以全网只需上传 14 个数据包,其可以有效减少网络中传输的数据包数目.

综上分析,对比其它 3 种数据收集策略,本文提出数据收集策略具有最少的网络数据收集传输代价,其传输代价为  $O(hk\log(N))$ ,而稠密投影数据收集传输代价为  $O(kN\log(N))$ ,稀疏投影数据收集传输代价为  $O(hk^2\log^2(N))$ .

## 7 实验评估

### 7.1 实验环境与数据

在本文的实验中,实验数据来源于真实的大规模无线传感器网络系统 CitySee 系统<sup>[29]</sup>. CitySee 系统是部署在无锡太湖新区,用于监控城市温度、湿度、二氧化碳排放等环境参数.系统中所有节点每隔 10 min 进行一次数据采样.我们选取同一区域的 256 个节点的温度数据作为本文的实验数据.

为了更好地评价本文提出的联合路由稀疏投影数据收集方案性能,我们将从以下 4 个方面来进行实验对比:稀疏表示基压缩性能,可靠网络环境下数据恢复性能,不可靠网络环境下数据恢复性能和网络能量消耗.针对数据压缩与恢复性能方面,本文统一采用均方误差作为性能评价指标,其表示为

$$\epsilon_{mse} = \frac{1}{N} \sum_{i=1}^N (x_i - \hat{x}_i) \quad (10)$$

其中,  $N$  为节点数目,  $x_i$  为第  $i$  个节点的数据,  $\hat{x}_i$  为第  $i$  个节点的估计值.

在数据恢复的过程中,我们选择 OMP 算法作为压缩感知的恢复算法<sup>[26]</sup>. 由于在基于压缩感知的数据收集过程中,数据收集代价与两个方面相关:一是单个测量值的收集代价;另一方面是重构感知数据所需的测量值数目.在单个测量值收集代价给定的情况下,数据的压缩性能直接决定所需的测量值数目.所以,本部分首先对提出的稀疏表示基的压缩性能进行评价,其次是对可靠和不可靠网络环境下不同数据恢复性能评价,最后是对不同数据收集方案对网络能量消耗的评价.

### 7.2 压缩性能评估

由压缩感知理论可知,待收集数据的稀疏程度直接决定其所需的测量值数目.在多跳无线传感器网络中,测量值数目直接决定网络中传输数据包的数目和网络能量消耗.

在本文实验中,我们将传统的 DCT、DFT 和 K-SVD 学习到的稀疏表示基与本文提出的稀疏表示基进行压缩性能对比.图 5 显示的是 4 种不同稀疏表示基的压缩性能对比,其中  $\alpha=0.1$  和  $\beta=10$ . 图 5 中 X-轴是选取变换后的最大系数数目(即稀疏度), Y-轴为压缩数据的截断误差.  $\Psi_{dct}$ ,  $\Psi_{dft}$ ,  $\Psi_{ksvd}$  和  $\Psi_{our}$  分别表示 DCT、DFT、K-SVD 和本文提出的方法得到的表示基.从图 5 显示的结果可以看出通用 DCT 和 DFT 表示基的压缩性能最差,与 K-SVD 和本文提出的稀疏表示基性能相差甚大.而本文方法学习到的稀疏表示基与 K-SVD 稀疏表示基在压缩性能上比较接近.所以,对比传统通用稀疏表示基,本文提出的稀疏表示基生成方法具有良好的感知数据压缩性能.

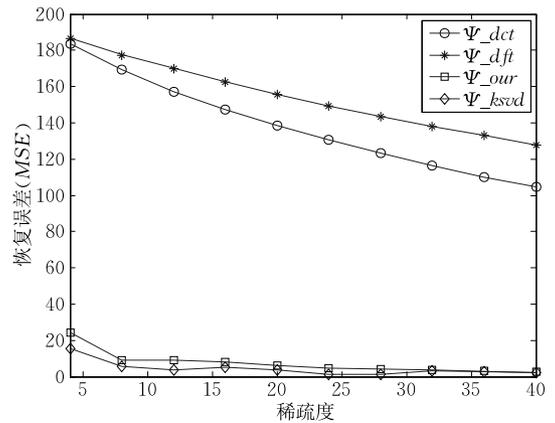


图 5 不同稀疏表示基压缩性能对比

图 6 显示的是不同  $\alpha$  和  $\beta$  的取值下,利用 OMP 算法进行感知数据的恢复误差实验对比结果.其中

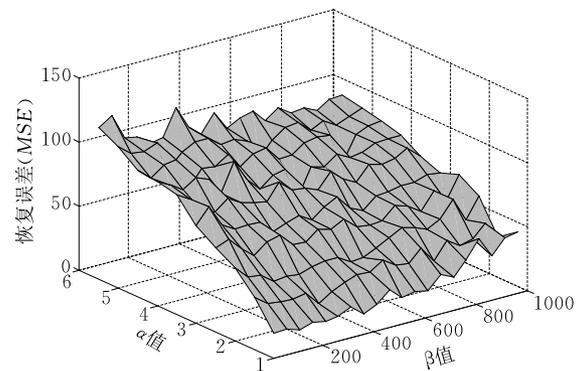


图 6 不同  $\alpha, \beta$  值下的稀疏表示基对数据恢复性能对比

$\alpha$  表示感知数据压缩误差权值,  $\beta$  表示投影矩阵与稀疏表示基之间的相干性权值. 从图 6 中可以看出, 无论是  $\alpha$  还是  $\beta$  值的增大都会导致数据恢复性能的下降. 这是因为问题(P)的限制条件  $\|\Phi(\mathbf{X}-\Psi\mathbf{S})\|_F^2$  和  $\|\Psi^T\Phi^T\Phi\Psi-\mathbf{I}\|_F^2$  都会影响数据的压缩性能.

### 7.3 可靠网络环境下数据恢复性能评估

如果本文提出的数据收集方案仅在不可靠网络情况下具备有效的感知数据重构功能, 而在可靠网络情况下数据收集性能低下, 则本数据收集方法也不能称为一种好的数据收集方法. 所以, 我们首先进行链路可靠情况下的性能评价, 即每跳链路均无丢包现象产生. 为了便于表示, 我们用  $\mathbf{G}$  表示高斯随机矩阵并作为稠密投影矩阵,  $\mathbf{S}$  表示元素为  $\pm 1$  的随机稀疏随机投影矩阵,  $\mathbf{R}$  表示我们提出的联合路由投影矩阵.

图 7(a) 显示的是采样稠密投影和稀疏投影两种方法在 DCT、K-SVD 和我们提出的表示基下的数据收集性能对比. 从图 7(a) 数据可以看出, 在相同的测量值数目下, 无论是稠密投影还是稀疏投影均具有相同的感知数据恢复精度. 同时, K-SVD 所

得到的稀疏表示基数据恢复性能最好, 其原因有两个方面: (1) K-SVD 所得到的稀疏表示基对数据的压缩性能最好; (2) 随机投影矩阵与任何稀疏表示基均具有良好的不相干性, 这也与压缩感知理论结果相一致. 图 7(b) 显示的是本文提出的稀疏表示基在稠密投影矩阵、稀疏投影矩阵和联合路由投影矩阵下的感知数据恢复性能. 从图 7(b) 可以看出, 随着测量值数目的增加感知数据的恢复精度不断提高. 同时, 联合路由的稀疏投影矩阵同样具有良好的数据恢复性能, 其性能与其它两种投影矩阵性能比较接近.

图 8 显示的是原始数据与本文提出的方法在 80 个测量值下的恢复数据, 恢复数据的误差  $\epsilon_{mse} = 1.5564$ . 从图 8 可以看出, 本文提出的方法具有精确恢复原始数据的能力.

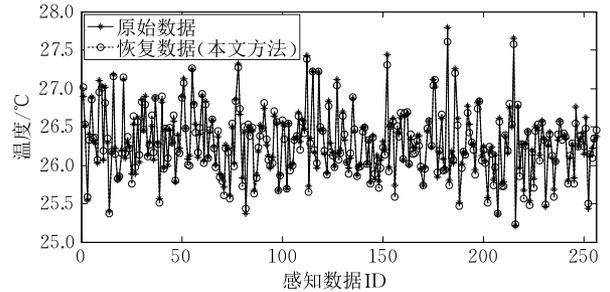
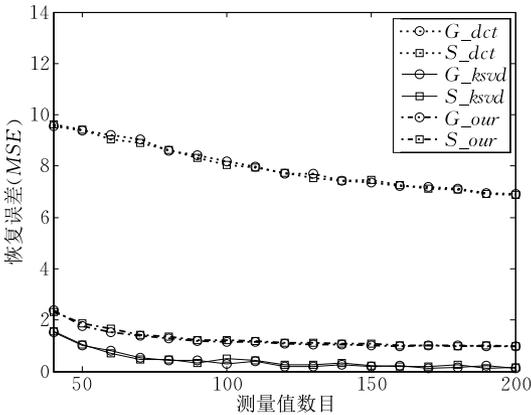


图 8 原始数据与本文提出的重构数据对比

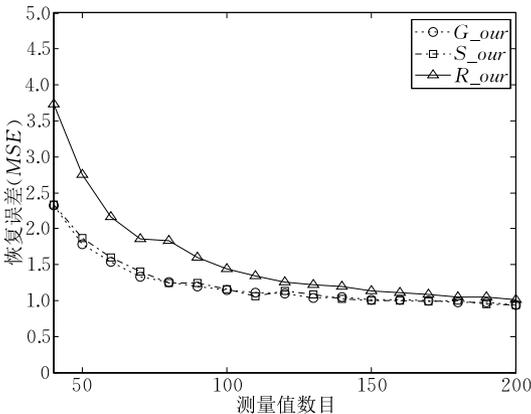
### 7.4 不可靠网络环境下数据恢复性能评估

在链路不可靠网络中, 我们分别针对丢包不重传和丢包重传两种情况下, 不同数据收集方案的感知数据恢复性能进行对比. 在本文实验中, 假设每条链路上丢包概率是相同的.

首先, 我们考虑链路丢包不重传的情况. 图 9 显示的是在链路丢包率为 5% 的情况下, 稠密随机投



(a) 不同表示基与稠密投影和稀疏投影



(b) 不同投影矩阵与本文提出的表示基

图 7 感知数据的恢复性能对比

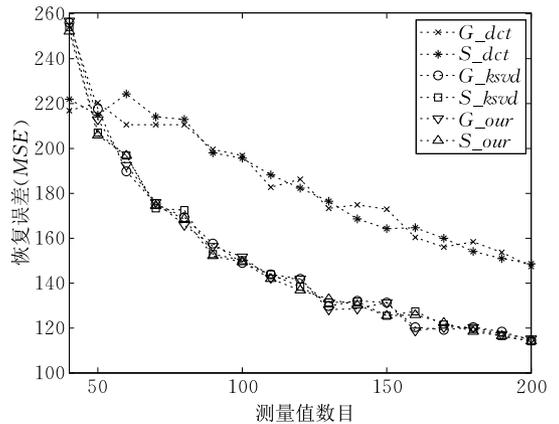
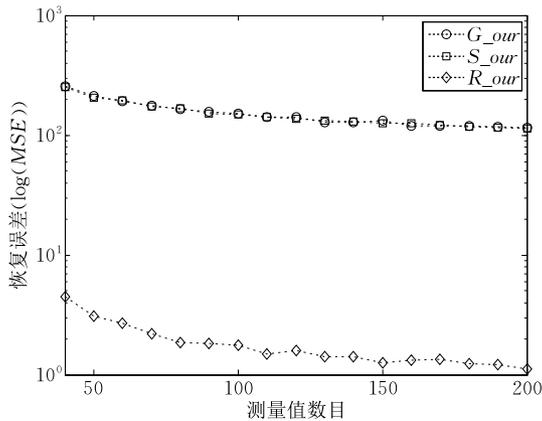


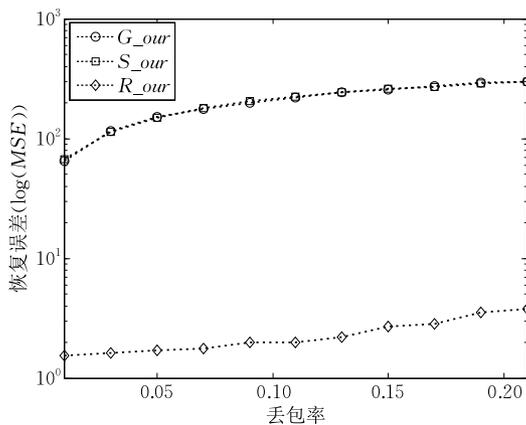
图 9 不同稀疏表示基下稠密投影和稀疏投影数据恢复性能对比(5%丢包率, 不重传)

影和稀疏随机投影对感知数据的恢复性能对比,从实验结果可以看出随着收集到的测量值数目增多,感知数据的恢复误差也在不断降低.但即便测量值数目为 200 时,不同类型的数据投影矩阵和不同稀疏表示基下的感知数据恢复误差都在 110 之上,这也意味着被恢复的感知数据基本都不能用,因为其测试的感知数据的数值在 20~35 之间.

图 10(a)显示的是在链路丢包率为 5% 的情况下,不同投影矩阵在本文提出的稀疏表示基下感知数据恢复误差对比结果.实验结果显示联合路由稀疏投影与本文学习到的稀疏表示基可以对感知数据进行有效恢复.例如,在测量值数目为 80 时,感知数据就可以高质量地恢复,而稠密投影和随机稀疏投影则无法实现感知数据的有效恢复.图 10(b)显示的是在测量值数目为 80 的情况下,不同丢包率下的感知数据恢复误差对比结果.从图中可以看出随着丢包率的增加,不同数据收集方案下的数据恢复误差都在不断增加.但本文提出的联合路由稀疏投影数据收集方法在丢包率为 20% 情况下,感知数据的



(a) 不同测量值下感知数据恢复性能对比(5%丢包率)

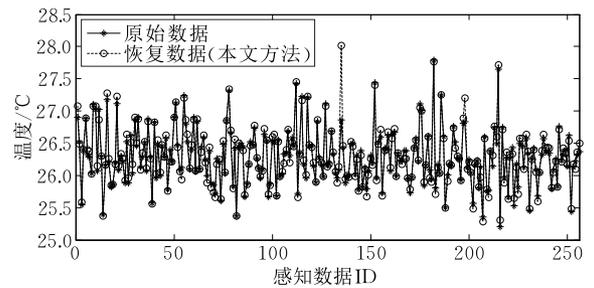


(b) 不同丢包率下感知数据恢复性能对比(M=80)

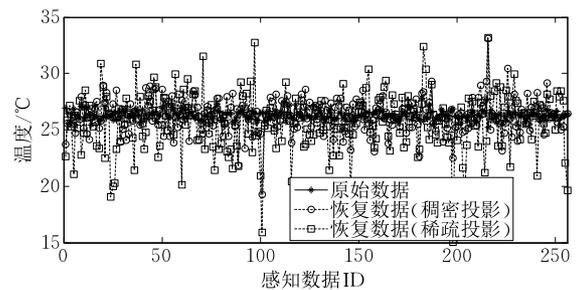
图 10 不同投影矩阵与本文稀疏表示基对感知数据恢复误差对比(丢包不重传)

重构评价误差  $\epsilon_{mse}$  仍然小于 10. 然而,稠密投影和稀疏投影在丢包率为 2% 的情况下,其误差  $\epsilon_{mse}$  就高达 200.

图 11 显示的是网络中某一时刻感知数据在本文学习到的稀疏表示基与稠密投影、稀疏投影和本文投影方法下数据恢复对比,其中测量值数目为 80,网络丢包率为 5%.图 11(a)显示的是原始数据与本文提出方法所恢复的数据,从对比结果可以看出本文提出的方法能比较精确恢复原始数据.图 11(b)显示的是原始数据与稠密投影和稀疏投影下恢复的数据.从图 11(b)中可以看出,两种数据收集下的恢复数据基本完全失真.因此,实验结果表明,即便网络存在丢包情况,本文提出的稀疏路由随机投影矩阵与学习得到的稀疏表示基仍具有良好的数据恢复功能.



(a) 原始数据与本文方法恢复数据

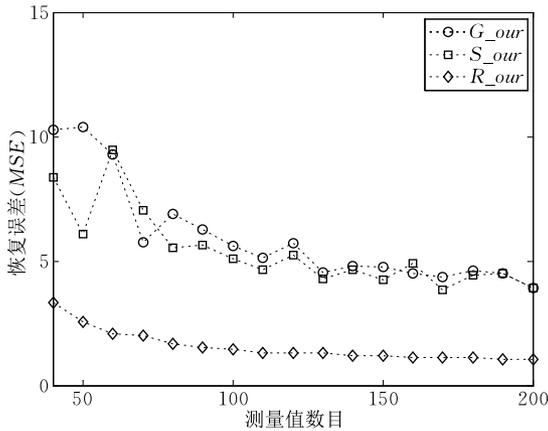


(b) 原始数据与稠密投影和稀疏投影恢复的数据

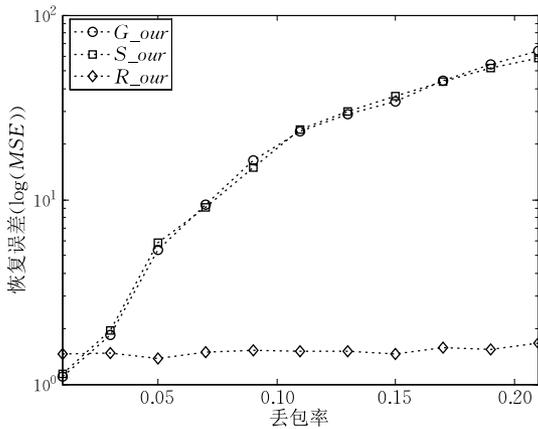
图 11 恢复数据与原始数据对比结果(丢包不重传)

其次,我们考虑网络丢包重传的情况.在本文实验中,丢包重传的最多次数为 30.图 12(a)显示的是在丢包率为 5% 和不同测量值数目下,高斯随机矩阵、随机稀疏矩阵和联合路由稀疏投影矩阵与本文提出的稀疏表示基对感知数目的恢复性能对比结果.尽管在可靠网络环境下,高斯随机投影矩阵和随机稀疏投影矩阵在本文提出的稀疏表示基下对感知数据的恢复性能优于联合路由稀疏投影矩阵,如图 7(b)所示.但从图 12(a)中可以明显看出联合路由稀疏投影矩阵具有更好的感知数据恢复性能.同时图 12(a)显示,随着收集到的测量值数目增加,感

知数据的恢复误差明显降低且具有较高的恢复精度 ( $\epsilon_{mse} < 4$ ). 图 12(b) 显示随着丢包率的增大, 稠密投影和随机稀疏投影下的感知数据恢复性能急剧降低, 而本文提出的方法对数据的恢复性能几乎不受影响. 从图 12 中可以看出, 在丢包率比较小的情况下, 丢包重传技术可以一定程度上改善感知数据的恢复性能. 但随着丢包率的增大, 感知数目的恢复性能急剧增加, 而本文提出的数据收集策略受丢包影响很小. 例如即便丢包率高达 20% 的情况, 感知数据的恢复误差仍然很小. 所以对比另外两种投影方法, 本文提出的联合路由稀疏投影与稀疏表示基学习方法具有最优的数据收集性能.



(a) 不同测量值下感知数据恢复性能对比(5%丢包率)

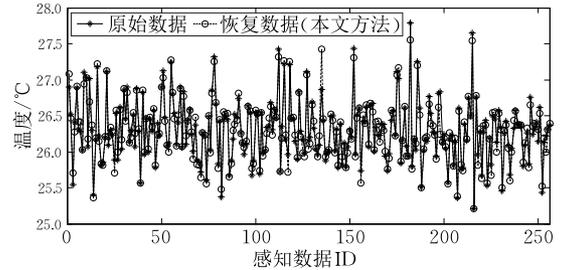


(b) 不同丢包率下感知数据恢复性能对比( $M=80$ )

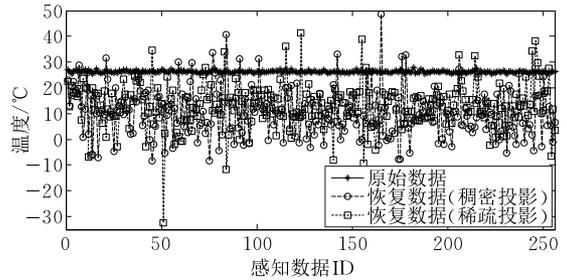
图 12 不同投影矩阵与本文稀疏表示基对感知数据恢复误差对比(丢包重传)

图 13 显示的是网络中某一时刻感知数据在本文学习到的稀疏表示基与稠密投影、稀疏投影和本文投影方法下数据恢复对比, 其中测量值数目为 100, 网络丢包率为 15%, 丢包重传最大数目为 30. 图 13(a) 显示的是原始数据与本文提出方法所恢复的数据, 从对比结果可以看出本文提出的方法能比

较精确恢复原始数据. 图 13(b) 显示的是原始数据与稠密投影和稀疏投影下恢复的数据. 从图 13(b) 中可以看出, 两种数据收集下的恢复数据完全偏离原始数据. 因此, 实验结果表明, 即便在网络丢包率高达 15%~20% 的情况下, 本文提出的稀疏路由由随机投影矩阵与学习得到的稀疏表示基仍具有良好的数据压缩与恢复功能.



(a) 原始数据与本文方法恢复数据



(b) 原始数据与稠密投影和稀疏投影恢复的数据

图 13 恢复数据与原始数据对比结果(丢包重传)

## 7.5 能量消耗评估

针对网络通信代价导致的能量消耗评估, 本文利用网络仿真平台 OMNeT++<sup>①</sup> 进行网络能量消耗实验. 为了体现本文提出联合路由稀疏随机投影数据收集方案的能量消耗性能, 本部分实验对比数据收集方法为稠密投影和稀疏随机投影数据收集策略.

在本文实验中, 我们假设 500 个节点随机的部署在  $1000\text{ m} \times 1000\text{ m}$  的二维平面内, sink 节点位于监控区域的边缘. 在网络能量消耗方面, 我们采用比特-跳能量消耗模型<sup>[30]</sup>, 其表示如下:

$$E_{Tx}(l, d) = E_{dec} \times l + \epsilon_{amp} \times l \times d^2$$

$$E_{Rx}(l) = E_{dec} \times l$$

其中,  $E_{Tx}(l, d)$  表示  $l$  比特的数据传输距离为  $d$  所消耗的能量,  $E_{Rx}(l)$  表示接收  $l$  比特的数据所消耗的能量,  $\epsilon_{amp}$  为传输放大器. 在本文实验中,  $E_{dec} = 50\text{ nJ/bit}$ ,  $\epsilon_{amp} = 100\text{ pJ/(bit} \cdot \text{m}^2)$ , 网络传输的数据包长度设

① OMNeT++: [http://www.omnetpp.org/omnetpp/cat\\_view/17-downloads/1-omnet](http://www.omnetpp.org/omnetpp/cat_view/17-downloads/1-omnet)

置为 1024 bits, 每个节点初始能量为 5000 J. 为了便于对比数据收集方案对能量的消耗, 我们忽略数据采样与数据处理的能量消耗, 其原因包括两个方面. 一方面是与数据传输对比, 数据采样与数据处理消耗的能量较少. 另一方面是不同的数据收集方案都需要节点进行数据采样与数据处理, 忽略采样与数据处理能量不影响数据传输过程中的能量对比.

图 14 显示的是不同数据收集策略下的网络节点能量消耗对比结果. X-轴表示收集的测量值数目, Y-轴表示死亡节点的数目. 从图 14 实验结果可以看出, 在稠密投影数据收集策略下, 网络能量消耗比较均衡, 几乎所有节点都在同一时间段内因能量消耗完而死亡, 但其只能收集到 5000 多个压缩感知测量值. 稀疏随机投影数据收集策略可以有效增加网络中收集的测量值数目. 当网络节点全部死亡时, 收集到的测量值数目为近 34000 个, 但对比本文提出的数据收集策略, 本文提出的方案能量消耗更少. 在收集到 40000 多个测量值数目时, 网络中还有近 300 个传感器节点存活. 实验结果表明, 对比现有数据收集策略, 本文提出的方案可以有效减少网络中节点能量消耗.

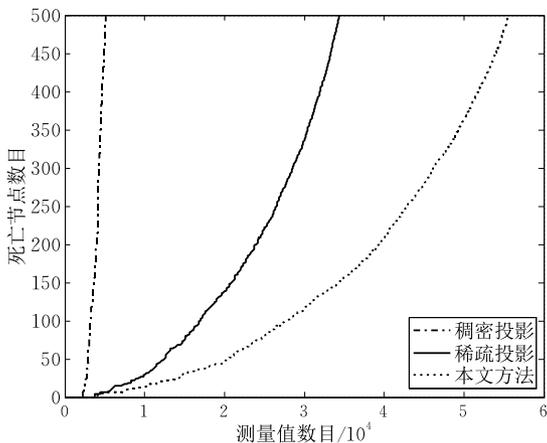


图 14 不同数据收集策略下的网络能量消耗对比

## 8 结 论

本文针对链路不可靠网络环境下, 现有压缩感知的大规模无线传感器网络数据收集方案存在感知数据恢复性能易受干扰的问题, 提出一种联合路由的稀疏随机投影数据收集方案, 该方案是根据网络路由来设计稀疏投影矩阵. 因此, 投影矩阵可以与路由主动匹配, 从而减少网络丢包对测量值的破坏. 而现有方法是直接将投影矩阵与网络路由被动匹配,

导致参与单个测量值节点过多, 造成测量值易受网络丢包影响. 为了实现联合路由的稀疏投影矩阵能符合压缩感知数据恢复条件, 本文给出一种限制性条件下的低相干性稀疏表示基学习算法, 使得学习到的稀疏表示基既具有很好的数据压缩性能又能与稀疏路由投影矩阵满足低相干性. 实验结果表明, 即便在丢包率为 20% 情况下, 本文提出的随机稀疏投影数据收集方案仍能对感知数据进行有效重构. 同时能量仿真结果表明本文提出的方法可以减少网络能量消耗.

## 参 考 文 献

- [1] Akyildiz I F, Su W, Sankarasubramaniam Y, et al. Wireless sensor networks: A survey. *Computer Networks the International Journal of Computer & Telecommunications Networking*, 2002, 38(4): 393-422
- [2] Akyildiz I F, Wang X, Wang W. Wireless mesh networks: A survey. *Computer Networks*, 2005, 47(4): 445-487
- [3] Anastasi G, Conti M, Francesco M D, et al. Energy conservation in wireless sensor networks: A survey. *Ad Hoc Networks*, 2009, 7(3): 537-568
- [4] Kimura N, Latifi S. A survey on data compression in wireless sensor networks//*Proceedings of the International Conference on Information Technology: Coding and Computing*. Las Vegas, USA, 2005: 8-13
- [5] Donoho D L. Compressed sensing. *IEEE Transactions on Information Theory*, 2006, 52(4): 1289-1306
- [6] Wang J, Tang S, Yin B, et al. Data gathering in wireless sensor networks through intelligent compressive sensing//*Proceedings of the International Conference on Computer Communications*. Orlando, USA, 2012: 603-611
- [7] Luo C, Wu F, Sun J, et al. Compressive data gathering for large-scale wireless sensor networks//*Proceedings of the International Conference on Mobile Computing and Networking*. Beijing, China, 2009: 145-156
- [8] Luo J, Xiang L, Rosenberg C. Does compressed sensing improve the throughput of wireless sensor networks?//*Proceedings of the International Conference on Communications*. Cape Town, South Africa, 2010: 1-6
- [9] Wang W, Garofalakis M, Ramchandran K. Distributed sparse random projections for refinable approximation//*Proceedings of the International Symposium on Information Processing in Sensor Networks*. Cambridge, USA, 2007: 331-339
- [10] Wu X, Liu M. In-situ soil moisture sensing: Measurement scheduling and estimation using compressive sensing//*Proceedings of the International Conference on Information Processing in Sensor Networks*. Beijing, China, 2014: 1-12

- [11] Zheng H, Yang F, Tian X, et al. Data gathering with compressive sensing in wireless sensor networks: A random walk based approach. *IEEE Transactions on Parallel & Distributed Systems*, 2015, 26(1): 35-44
- [12] Liu X Y, Zhu Y, Kong L, et al. CDC: Compressive data collection for wireless sensor networks. *IEEE Transactions on Parallel & Distributed Systems*, 2015, 26(8): 2188-2197
- [13] Wu X, Xiong Y, Yang P, et al. Sparsest random scheduling for compressive data gathering in wireless sensor networks. *IEEE Transactions on Wireless Communications*, 2014, 13(10): 5867-5877
- [14] Liu Y, He Y, Li M, et al. Does wireless sensor network scale? A measurement study on GreenOrbs//*Proceedings of the International Conference on Computer Communications*. Shanghai, China, 2012: 873-881
- [15] Wu X, Yang P, Jung T, et al. Compressive sensing meets unreliable link: Sparsest random scheduling for compressive data gathering in lossy WSNs//*Proceedings of the International Symposium on Mobile Ad Hoc Networking and Computing*. Philadelphia, USA, 2014: 13-22
- [16] Xiang L, Luo J, Rosenberg C. Compressed data aggregation: Energy-efficient and high-fidelity data collection. *IEEE/ACM Transactions on Networking*, 2013, 21(6): 1722-1735
- [17] Leinonen M, Codreanu M, Juntti M. Sequential compressed sensing with progressive signal reconstruction in wireless sensor networks. *IEEE Transactions on Wireless Communications*, 2015, 14(3): 1622-1635
- [18] Fei Z, Lin F, Zhang Q, et al. Improved sparsest random scheduling for compressive data gathering under unreliable link in wireless sensor networks//*Proceedings of the International Conference on Internet of Things*. Chengdu, China, 2017: 205-210
- [19] Kreutz-Delgado K, Murray J F, Rao B D, et al. Dictionary learning algorithms for sparse representation. *Neural Computation*, 2014, 15(2): 349-396
- [20] Aharon M, Elad M, Bruckstein A. K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation. *IEEE Transactions on Signal Processing*, 2006, 54(11): 4311-4322
- [21] Jenatton R, Mairal J, Obozinski G, et al. Proximal methods for sparse hierarchical dictionary Learning//*Proceedings of the International Conference on Machine Learning*. Haifa, Israel, 2010: 487-494
- [22] Mailhé B, Barchiesi D, Plumbley M D. INK-SVD: Learning incoherent dictionaries for sparse representations//*Proceedings of the International Conference on Acoustics, Speech and Signal Processing*. Kyoto, Japan, 2012: 3573-3576
- [23] Barchiesi D, Plumbley M D. Learning incoherent dictionaries for sparse approximation using iterative projections and rotations. *IEEE Transactions on Signal Processing*, 2013, 61(8): 2055-2065
- [24] Bai H, Li G, Li S, et al. Alternating optimization of sensing matrix and sparsifying dictionary for compressed sensing. *IEEE Transactions on Signal Processing*, 2015, 63(6): 1581-1594
- [25] Ding X, Chen W, Wassell I J. Joint sensing matrix and sparsifying dictionary optimization for tensor compressive sensing. *IEEE Transactions on Signal Processing*, 2017, 65(14): 3632-3646
- [26] Tropp J A, Gilbert A C. Signal recovery from random measurements via orthogonal matching pursuit. *IEEE Transactions on Information Theory*, 2007, 53(12): 4655-4666
- [27] Chen S S, Donoho D L, Saunders M A. Atomic decomposition by basis pursuit. *Journal on Scientific Computing*, 1998, 20(1): 33-61
- [28] Needell D, Tropp J A. CoSaMP: Iterative signal recovery from incomplete and inaccurate samples. *Applied and Computational Harmonic Analysis*, 2009, 26(3): 301-321
- [29] Miao X, He Y, et al. CitySee: Urban CO<sub>2</sub> monitoring with sensors//*Proceedings of the International Conference on Computer Communications*. Orlando, USA, 2012: 1611-1619
- [30] Heinzelman W R, Chandrakasan A, Balakrishnan H. Energy-efficient communication protocol for wireless microsensor networks//*Proceedings of the Hawaii International Conference on System Sciences*. Maui, USA, 2000: 10-18



**WU Xuan-Gou**, born in 1979, Ph. D., associate professor. His research interests include wireless networks, and compressive sensing.

**CHU Zhao-Bin**, born in 1993, M. S. candidate. His research interest is wireless networks.

**ZHENG Xiao**, born in 1975, Ph. D., professor. His research interests include computer networks, service computing.

**WANG Xiu-Jun**, born in 1982, Ph. D., lecturer. His research interests include computer algorithm, algorithm optimizing.

**YANG Pan-Long**, born in 1976, Ph. D., professor. His research interests include wireless networks, and crowdsensing networks.

## Background

Compressive sensing (CS) based data gathering is considered as a promising approach to reduce data transmission cost and balance network energy consumption for large-scale wireless sensor networks (WSNs). Existing work mainly focuses on applying CS on reliable WSNs, which can be grouped into dense and sparse projection based data gathering. However, existing reliable CS based data gathering techniques cannot meet the requirement of error-prone sensor networks, because the packet loss would damage the CS measurements which could result in arbitrarily bad recovery performance of sensory data.

Due to the fact that existing efforts make the projection matrix match routing matrix passively, which leads to too many sensors to be participated in one CS measurement gathering. This is reason the why packet loss would damage the CS measurement severely. In our work, we first design the sparse projection matrix with routing consideration, and reduce the impact of packet loss on CS measurements.

Therefore, we design the representation basis of sensory data, and make it satisfy both data compression and low coherence with our proposed sparse projection matrix. Our sparse projection data gathering scheme can meet the packet loss during data gathering.

Our work is a partly research of project “The Study on Adaptively Compressive Sampling Data Gathering for Wireless Sensor Networks”, the National Natural Science Foundation of China under Grant No. 61402009. This project mainly researches adaptive compressive sensing based data gathering in lossy wireless sensor networks.

This research was supported by the National Natural Science Foundation of China under Grant Nos. 61672038, 61502520, 61502010, 61402008, the Advanced Outstanding Talents Cultivation Project of Anhui Province (No. gxgwx2018017), the NSF of Jiangsu for Distinguished Young Scientists (No. BK20150030), and the Major Technologies R&D Special Program of Anhui China (No. 16030901060).