

不可靠链路下基于压缩感知的 WSN 数据收集算法

张策¹, 张霞¹, 李鸥¹, 梅关林¹, 韩哲¹, 张大龙², 刘广怡¹

(1. 解放军信息工程大学信息工程学院, 河南 郑州 450001; 2. 郑州大学信息工程学院, 河南 郑州 450001)

摘 要: 为了解决 WSN 中基于压缩感知 (CS, compressive sensing) 的数据收集方法会受不可靠链路影响的问题, 首先通过实验对基于 CS 的数据收集算法中数据重构信噪比与链路误码率的关系进行了定量研究, 根据 WSN 链路分组丢失特性将分组丢失问题分为轻负载和重负载 2 种情况。针对轻负载下的链路不可靠, 建立随机分组丢失模型, 并提出了基于邻居拓扑空间相关预测的 CS 数据收集算法, 利用数据空间相关性减小错传的影响。针对重负载下的链路不可靠, 建立节点伪失效模型, 并提出了基于稀疏调度的 CS 数据收集算法, 通过改变观测矩阵稀疏度, 避免观测出错数据, 弱化不可靠链路的影响。仿真分析表明, 在不增加能耗的前提下有效提高了数据重构质量, 降低了不可靠链路对 CS 数据收集的影响。

关键词: 无线传感网; 数据收集; 压缩感知; 不可靠链路; 空间相关性

中图分类号: TP393

文献标识码: A

Compressive sensing based data gathering algorithm over unreliable links in WSN

ZHANG Ce¹, ZHANG Xia¹, LI Ou¹, MEI Guan-lin¹, HAN Zhe¹, ZHANG Da-long², LIU Guang-yi¹

(1. School of Information Systems Engineering, PLA Information Engineering University, Zhengzhou 450001, China;
2. School of Information Engineering, Zhengzhou University, Zhengzhou 450001, China)

Abstract: To solve the problem that the ubiquitous unreliable links in the WSN influence the performance of the compressive sensing (CS) based data gathering, first the relationship between the reconstruction SNR of CS-based data gathering algorithm and the bit-error-ratio (BER) were simulated quantitatively. Then classify two cases were classified, namely light-payload and heavy-payload, relying on the analysis of wireless link packet loss characteristics. The random packet loss model was conceived to describe the packet loss under light-payload scenario. Further the neighbor topology spatial correlation prediction-based CS data gathering (CS-NTSC) algorithm was proposed, which utilized the nodes spatial correlation to reduce the impact of error. Additionally, the node pseudo-failure model was conceived to describe the packet loss occurred in network congestion, and then the sparse schedule-aided CS data gathering (CS-SSDG) algorithm were conceived, for the purpose of changing the sparsity of measurement matrix and avoiding measurements amongst the nodes affected by unreliable links, thus weakening the impact of error/loss on data reconstruction. Simulation analysis indicates that the proposed algorithms are not only capable of improving the accuracy of the data reconstruction without extra energy, but also effectively reducing the impact affected by the unreliable links imposed on CS-based data gathering.

Key words: WSN, data gather, compressive sensing, unreliable link, spatial correlation

1 引言

无线传感网 (WSN) 正在逐步从实验室走向实际应用, 在此过程中, 传感网仍受到诸多限制, 面

临许多挑战: 传感器节点通常使用电池作为能源, 能源受限; 所采集原始数据冗余度高, 传输能耗高; 节点能耗不均衡, 靠近 sink 的节点能耗快, 易形成能量空洞, 影响网络寿命; 传感网无线链路不可靠,

收稿日期: 2016-01-24; 修回日期: 2016-08-05

基金项目: 国家科技重大专项基金资助项目 (No.2014zx03006003)

Foundation Item: The National Science and Technology Major Projects of China (No.2014zx03006003)

存在较高误码率。

近年来,基于压缩感知的数据收集方法受到研究者的关注^[1-3]。将压缩感知理论^[4,5]应用于 WSN 数据收集中,具有如下优势:第一,有效利用传感器节点所采集的原始信息的空间相关性,减少数据传输能耗;第二,压缩感知数据收集方法具有天然的能耗均衡特性,能降低甚至克服能量洞问题,从而延长网络寿命;第三,相对于分布式数据压缩等传统的压缩方法,该方法具有压缩过程简单而数据重构过程复杂的特点,十分适合传感器节点信息处理能力和能源受限的要求。文献[6]将 CS 技术与 Pegasus 路由协议相结合,在路由链中压缩数据以均衡能耗、延长网络寿命,但是该方法的网络顽健性差。文献[2~7]将 CS 技术与最小路径树路由协议相结合,以降低整个数据传输路径上的能耗。文献[8]指出,简单地在树形路由中使用 CS,将增加叶节点和距离叶节点较近的中间节点的通信量,针对该问题提出了混合压缩感知(hybrid-CS)数据收集方法,仅对一部分通信量高的父节点使用压缩感知技术,以此减少网络数据通信量。当网络规模较大时,通过分簇构建层次化网络结构,更有助于提高网络数据传输和管理的效率。文献[9]借鉴 Leach 的思想,在簇内压缩感知模式下计算出全网最优簇首个数,并使簇首均匀分布在全网中,以减少网络能耗。

以上文献都假定传输链路是可靠的,在实际应用场景中,传感器节点被大量布置在森林、火山和战场等环境中^[10],受环境因素、障碍物阻挡、节点布置高度^[11]、天线方向^[12]、不对称链路^[13]和信道冲突堵塞等原因的影响,无线链路不可靠现象普遍存在。

传统的数据收集方法中,一个分组丢失或出错,仅影响单个传感器,然而在分布式压缩感知数据收集中,多个传感器节点采集的信息通过随机观测和求和操作,被合并为一个数据分组,一旦该分组丢失或错传,参与本次压缩感知的所有传感器节点都将受到影响,不可靠无线链路对基于压缩感知的数据收集方法的影响不容忽视,然而,现有研究中对此问题的讨论还很少。文献[14]考虑了树形路由中不可靠链路对压缩感知的影响,分析了分组丢失对数据重构精度的影响。首先根据预先设定的调度概率随机选择出本轮数据收集需要参与的节点,并通过最短路由将这些节点采集数据传输至 sink 节点,不可靠链路会使其中部分数据丢失。sink 根据收到的正确数据分组构造观测矩阵,令观测矩阵每列只有一

个非零值,使通过最短路由方式收到的数据即为压缩感知观测值,并构造出与此观测矩阵相对应的稀疏基。最后根据重构精度的高低,动态调整调度概率和观测数,以满足最小重构精度。由于 sink 根据最终接收到的数据构造观测矩阵,避免观测发生分组丢失的节点,能够抵抗由于不可靠链路对数据收集带来的影响,但其数据分组通过最短路由方式传输,仍然存在网络能耗不均衡问题,且没有考虑被监测网络中事件源对数据相关性的影响。

为此,本文在有事件源发生的传感网中、在不可靠无线链路的条件下,研究了基于压缩感知的数据收集算法。首先,通过实验的方法定量研究了在压缩感知数据收集中,数据重构信噪比,即重构精度,与无线链路误码率的关系;分别针对网络负载较轻、无线链路随机分组丢失和网络负载较重、无线链路由于拥塞连续分组丢失的 2 种情况,建立无线链路分组丢失模型,并提出了基于邻居拓扑空间相关性预测的压缩感知数据收集算法(CS-NTSC, neighbor topology spatial correlation prediction based CS data gathering)和基于稀疏调度的压缩感知数据收集算法(CS-SSDG, sparse schedule for CS data gathering),降低不可靠链路对数据收集的影响;最后验证了方法的有效性,实现对不可靠链路中有损数据高精度、低能耗的数据收集。

2 系统模型

本节首先介绍整个网络的数据空间相关性模型,考虑网络内有事件源发生时节点空间相关性并建立模型,介绍了传感网中基于压缩感知的分簇数据收集方法。

2.1 传感网数据空间相关性模型

本文假设节点按照网格部署。假设 WSN 的监测范围是一个 $a_i \times a_i$ 的方形区域,此区域被均匀分成了 N 块大小相等的子区域,假设每个子区域中都有且仅有一个传感器节点。用方阵 $\mathbf{F}_{\sqrt{N} \times \sqrt{N}}$ 记录传感器节点采集的信息,其中, f_{ij} 表示网络子区域 (i, j) 布设的传感器节点所采集到的信息^[15]。

假设监测区域中有突发事件源发生,如森林中的着火点。用方阵 \mathbf{G} 记录网络中的突发事件源的信号强度,其中, g_{ij} 表示子区域 (i, j) 中的事件源的信号强度, $g_{ij} = 0$ 表示该子区域中无突发事件源。

本文将方阵转化为向量的形式来表示整个区域中传感器节点和事件源,即

$$\mathbf{X}_{N \times 1} = (f_{11}, \dots, f_{1\sqrt{N}}, \dots, f_{\sqrt{N}1}, \dots, f_{\sqrt{N}\sqrt{N}})^T \quad (1)$$

$$\mathbf{V}_{N \times 1} = (g_{11}, \dots, g_{1\sqrt{N}}, \dots, g_{\sqrt{N}1}, \dots, g_{\sqrt{N}\sqrt{N}})^T \quad (2)$$

传感网中每个传感器节点采集的信号强度是由网络中 N_s 个事件源的信号叠加而成的^[15], 即

$$\mathbf{X}_{N \times 1} = \boldsymbol{\Psi}_{N \times N} \mathbf{V}_{N \times 1} = \begin{pmatrix} \psi_{11} & \dots & \psi_{1N} \\ \vdots & & \vdots \\ \psi_{N1} & \dots & \psi_{NN} \end{pmatrix} \begin{pmatrix} g_{11} \\ \vdots \\ g_{1\sqrt{N}} \\ \vdots \\ g_{\sqrt{N}1} \\ \vdots \\ g_{\sqrt{N}\sqrt{N}} \end{pmatrix} \quad (3)$$

其中, $\boldsymbol{\Psi}$ 为传感器感知数据随距离衰减系数矩阵。本文利用欧氏距离的空间相关性模型^[16], 假定传感器节点 i 和 j 的坐标为 (x_i, y_i) 和 (x_j, y_j) , 两节点之间距离为

$$d_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \quad (4)$$

若在节点 i 处有事件源发生, 节点 i 接收到信号功率为 P_i , 节点 j 接收到信号功率为 P_j , 信号按欧氏距离衰减, 则有

$$P_j = C P_i d_{ij}^{-n} = C P_i \left[\sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \right]^{-n} \quad (5)$$

其中, C 为常数, n 为信号的衰减系数。不同类型的事件源衰减系数不同。2 个区域中节点采集到的信息之间的相关性与其欧氏距离成反比, 距离越近, 衰减系数 n 越小, 节点采集到的信息越接近, 相关性越大。本文令

$$\psi_{ij} = \psi_{ji} = C \left[\sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \right]^{-n} \quad (6)$$

本文研究均基于本节的全网相关性数据模型, 提出相关算法。

2.2 基于压缩感知的分簇数据收集方法

假定网络采用分簇结构, 以距离事件源最近的节点为簇首, 若网络中有 N_s 个事件源, 将整个网络划分为 N_s 个簇^[16], 在每个簇内, 使用压缩感知技术进行数据收集, 如图 1 所示, 重点研究簇内不可靠链路对基于压缩感知的数据收集方法的影响及相应对策。

假设簇内有 N_1 个成员节点, 随机观测矩阵为 $\boldsymbol{\Phi} = (\phi_{ij})_{M \times N_1}$, 其中, $M \ll N_1$, 使用文献[17]中给出的观测矩阵。

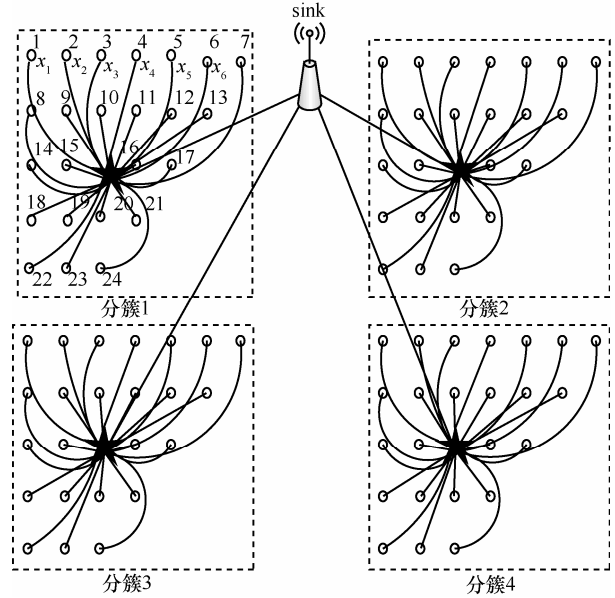


图 1 网内分簇数据收集示意

$$\phi_{ij} = \sqrt{s} \begin{cases} +1, & p = \frac{1}{2s} \\ 0, & p = 1 - \frac{1}{s} \\ -1, & p = \frac{1}{2s} \end{cases} \quad (7)$$

其中, s 控制观测矩阵的稀疏程度, p 为 3 种情况出现的概率。若 $\frac{1}{s} = \frac{\lg N_1}{N_1}$, 则 $\boldsymbol{\Phi}$ 中每一行有 $\lg N_1$ 个非零元素, 本文令 $\frac{1}{s} = 1$, 即 $\boldsymbol{\Phi}$ 为稠密观测矩阵。簇首随机生成 $M \times N_1$ 维的观测矩阵 $\boldsymbol{\Phi}$ 并存储, 每个簇内成员节点将感知到的数据 $\mathbf{X} = (x_1, \dots, x_{N_1})^T$ 发送至簇首, 则簇内的观测矢量 \mathbf{Y} 可以表示为

$$\mathbf{Y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_M \end{pmatrix} = \boldsymbol{\Phi} \mathbf{X} = \begin{pmatrix} \phi_{11} & \dots & \phi_{1N_1} \\ \vdots & & \vdots \\ \phi_{M1} & \dots & \phi_{MN_1} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{N_1} \end{pmatrix} \quad (8)$$

簇首将本簇的观测向量发送至 sink, sink 根据观测矩阵重构簇内数据, 其计算过程可转化为一个求解凸优化问题。

$$\hat{\mathbf{X}} = \arg \min \|\mathbf{X}\|, \text{ s.t. } \mathbf{Y} = \boldsymbol{\Phi} \mathbf{X} \quad (9)$$

3 不可靠链路对簇内数据收集影响分析

在无线传感网中, 链路质量并不完全可靠, 因系统噪声和随机噪声等环境因素干扰导致节点间传输

数据发生错误。本文通过实验的方法，定量地研究不可靠链路对压缩感知数据收集方法的影响。在有一定误码率的信道中，当节点发送数据分组发生错传时，若接收节点通过信道译码无法恢复发生的错传位，此时接收节点将丢弃错传数据分组，即分组丢失。

在一个簇中使用压缩感知技术进行数据收集，成员节点通过单跳的方式将收集到数据传输到簇首节点，如图 1 所示，在簇首节点处进行数据观测。假设簇中有 N_1 个节点，在可靠链路下收集数据，没有发生错传时，簇首收到的数据为 $\mathbf{X}_{N_1 \times 1} = (x_1, x_2, \dots, x_i, \dots, x_{N_1})^T$ ，则观测矢量 \mathbf{Y} 为

$$\mathbf{Y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_M \end{pmatrix} = \begin{pmatrix} \phi_{11} & \dots & \phi_{1N_1} \\ \vdots & & \vdots \\ \phi_{M1} & \dots & \phi_{MN_1} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_i \\ \vdots \\ x_{N_1} \end{pmatrix} \quad (10)$$

为了评估感知数据恢复质量，本文用重构数据信噪比来衡量数据重构精度，定义为

$$SNR = 10 \lg \frac{\|\mathbf{X}\|_2}{\|\hat{\mathbf{X}} - \mathbf{X}\|_2} = 10 \lg \frac{(\sum |\mathbf{X}|^2)^{\frac{1}{2}}}{(\sum |\hat{\mathbf{X}} - \mathbf{X}|^2)^{\frac{1}{2}}} \quad (11)$$

由式(9)可知， \mathbf{X} 为原数据， $\hat{\mathbf{X}}$ 为重构数据，重构信噪比越高，算法性能越好。图 2 给出了观测次数 M 与数据重构信噪比的关系，仿真中假定簇内有 900 个节点，一个事件源，衰减系数 $n = 0.01$ 且误码率 $P_b = 0$ 。随着 M 的增加，重构数据信噪比也在增加，当 $M > 450$ 后，信噪比趋于平稳。由于不同的用户对重构数据信噪比要求不尽相同，相应的 M 也不同。为了便于比较，在后续的仿真实验中，本文选取 $SNR = 36.16$ dB 作为比较基准，由图 2 可得此时 $M = 450$ 。

若节点 i 因链路问题发生错传，传输数据变为 x'_i ，簇首收到的数据为 $\mathbf{X}'_{N_1 \times 1} = (x_1, x_2, \dots, x'_i, \dots, x_{N_1})^T$ ，若直接使用错传数据，则观测向量 \mathbf{Y}' 为

$$\mathbf{Y}' = \begin{pmatrix} y'_1 \\ y'_2 \\ \vdots \\ y'_M \end{pmatrix} = \begin{pmatrix} \phi_{11} & \dots & \phi_{1N_1} \\ \vdots & & \vdots \\ \phi_{M1} & \dots & \phi_{MN_1} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x'_i \\ \vdots \\ x_{N_1} \end{pmatrix} \quad (12)$$

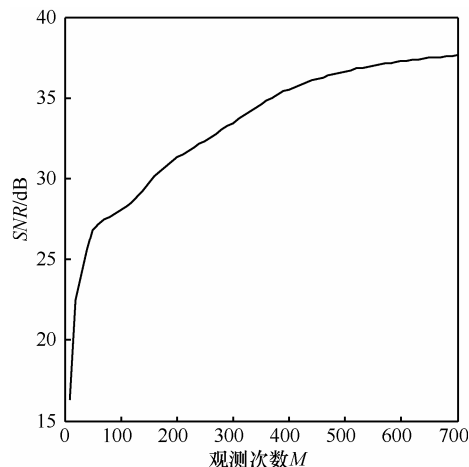


图 2 观测次数与信噪比关系

此时，由于一个节点的发送数据发生错误，导致观测向量 \mathbf{Y}' 中每一个观测值都受到了影响。图 3 所示为当观测次数 $M = 450$ 时，不同误码率对簇内压缩感知数据重构精度的影响，在高误码率环境中，重构数据信噪比远远低于无误码情况下的重构信噪比 36.16 dB；随着误码率的减小，重构精度也在增加。在实际的无线传感网中，链路质量十分不可靠，且误码率较高，此时利用压缩感知收集到的数据进行重构，会得到低精度甚至无用的数据。

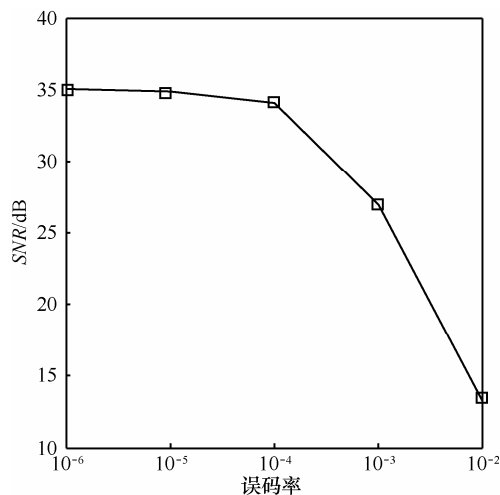


图 3 误码率对重构精度影响

进一步分析给定误码率条件下不同比特发生错误对重构信噪比的影响。若传感器节点将采集到的数据转换成 8 bit 二进制符号进行传输，即第 i 个节点采集到数据为

$$x_i = \sum_{b=1}^8 k_b^i \cdot 2^{8-b} \quad (13)$$

其中, $i \in \{1, \dots, N_1\}, b \in \{1, \dots, 8\}$, k'_b 表示第 i 个节点的第 b 位, 值为 0 或 1。 $b=1$ 表示高位, $b=8$ 表示低位。误码会使 k'_b 反转, 当传输链路不可靠时, 这种错误会在任意比特位发生。图 4 给出了当误码率 $Pb=10^{-3}$ 时, 错误发生的位数 b 与重构信噪比的关系。由图可知, 在高误码率的情况下, 错误发生在高位, 严重影响了重构数据的精度; 若错误发生在最低位, 即错误程度较小, 即便在误码率较高的环境中, 重构精度仍然较高。

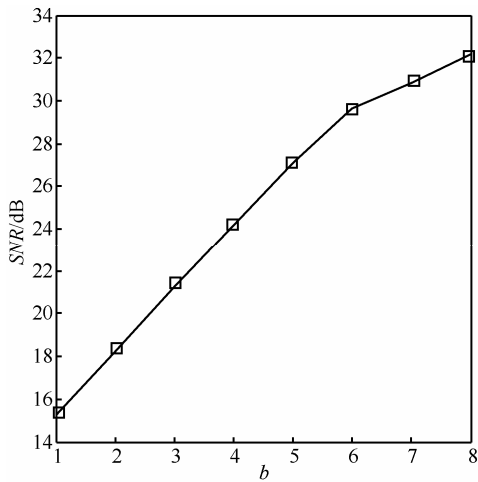


图4 错误发生位数对重构精度影响

由以上仿真实验可知, 不可靠链路对分布式压缩感知数据收集有一定影响, 当误码率较低, 错误位发生在低位时, 重构精度受影响程度小; 在误码率较高的信道环境和错误位发生在高位时, 数据重构精度会大幅降低, 甚至无法使用。下面针对受不可靠链路影响较大的情况开展研究, 建立相应的分组丢失模型和数据收集算法, 给出对策。

4 分组丢失模型建立

传感网中当节点采集数据和回传频率低时, 单个节点数据分组发送任务减少, 全网负载轻; 当节点采集数据和回传的频率高, 单个节点数据分组发送任务增多, 全网负载加重, 网络易发生拥塞。基于传感网负载轻重与文献[18]中多种分组丢失模型, 本文将分组丢失分成如下 2 种情形。第一种情况考虑网络轻负载、有随机干扰或噪声, 此时节点在每个数据传输周期随机分组丢失; 第二种情况考虑网络局部突然产生大量负载、链路存在噪声和拥塞, 此时节点会在多个相邻数据传

输周期连续分组丢失。分别针对以上 2 种情形, 总结出 2 种分组丢失模型: 随机分组丢失模型和节点伪失效模型。

在误码率为 Pb 的无线信道中, 假设节点 A 发送一个 L byte 的数据分组, 则数据分组成功发送的概率为

$$p_0 = (1 - Pb)^{8L} \quad (14)$$

引入变量 $R_i(t)$ 来表示节点 i 在 t 时刻数据分组发送状态, 数据分组被成功接收, 则 $R_i(t)=1$, 数据分组发送失败, 则 $R_i(t)=0$, 如式(15)所示。

$$R_i(t) = \begin{cases} 1, & \text{发送成功} \\ 0, & \text{发送失败} \end{cases} \quad (15)$$

4.1 随机分组丢失模型

若网络没有发生拥塞, 且每个节点在每个时刻发生分组丢失是独立随机的, 即在一轮数据收集发生错传后, 下一轮数据收集并不受影响, 本文将这种分组丢失模型称为随机分组丢失模型, 如图 5(a) 所示, 该模型一般由环境噪声和信道冲突造成的, 单个节点在第 t 轮数据收集时数据分组成功发送概率为

$$p(t) = p_0 = (1 - Pb)^{8L} \quad (16)$$

根据节点随机丢失发送数据的特点, 可以利用邻居节点之间的空间相关性“合成” A 节点采集数据的预测值。

4.2 节点伪失效模型

若网络发生拥塞, 节点 A 在 t 轮数据收集分组丢失后, 在 $t+1$ 轮数据收集时则有更大概率分组丢失, 即在一段时间内节点如同失效, 不能成功收发数据, 会影响多轮的数据收集, 本文将这种分组丢失模型称为节点伪失效模型, 如图 5(b)所示, 该模型下单个节点在第 t 轮数据收集时数据分组成功发送概率为

$$p(t) = \begin{cases} p_0, & R_A(t-1) = 1 \\ p_1, & R_A(t-1) = 0 \end{cases} \quad (17)$$

其中, $p_1 < p_0$, 一旦节点 A 在 $t-1$ 时刻发生了分组丢失, 则该节点在此后连续的多个发送周期内, 都会以更高的概率发生分组丢失, 此时考虑减小丢失数据分组对数据重构的影响。且在节点部署密集传感网中, 网络一旦发生拥塞, 一般在一定范围内的多个节点均发生伪失效。

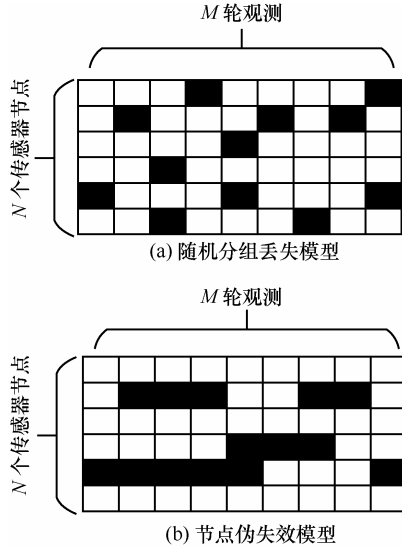


图 5 分组丢失模型示意

5 算法设计

本节分别针对以上 2 种分组丢失模型，提出 CS-NTSC 算法和 CS-SSDG 算法。由于在链路可靠性较低的情况下，重传不仅会带来极大的通信开销与传输时延，而且在有些情况下可能恶化系统性能^[19]，本文所提算法均设定在不重传的前提下来提高数据重构精度，减小误差。

5.1 CS-NTSC 算法

针对随机分组丢失模型，根据对数据空间相关性特性的分析，提出基于邻居拓扑空间相关性预测的压缩感知数据收集算法。

在传感网中，传感器节点通常可以根据距离来判断节点是否在自己邻居范围内，假设此范围为 R_{Thr} ，引入矩阵 $\mathbf{A} = (\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_{N_1})^T$ ，即令

$$\mathbf{A} = (\mathbf{a}_{ij})_{N_1 \times N_1} = \begin{cases} 1, & R(i, j) \leq R_{\text{Thr}} \\ 0, & \text{其他} \end{cases} \quad (18)$$

其中， N_1 为簇中节点数， $R(i, j)$ 为节点 i 与节点 j 的距离， $i \in \{1, 2, \dots, N_1\}$ ， $j \in \{1, 2, \dots, N_1\}$ 。若第 k_c 个节点发生错传，则令矩阵 \mathbf{A} 的第 k_c 列为 $\mathbf{0}$

$$(\mathbf{a}_{1k_c}, \mathbf{a}_{2k_c}, \dots, \mathbf{a}_{N_1k_c})^T = \mathbf{0}, k_c \in \{1, 2, \dots, N_1\} \quad (19)$$

矩阵 \mathbf{A} 表示节点与其邻居节点的相关性，利用节点之间的距离表示相关性强弱。若节点 z_0 与邻居节点 z_1 距离较近，则其相关性强， $a_{z_0z_1}$ 较大；若节点 z_0 与邻居节点 z_1 距离较远，则其相关性弱， $a_{z_0z_1}$ 较

小；若节点 z_0 的邻居节点 z_1 发生错传，则其相关性为 0，即 $a_{z_0z_1} = 0$ 。

将矩阵 \mathbf{A} 每一行归一化后，引入邻居拓扑矩阵 $\mathbf{H} = (\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_{N_1})^T$ 。

$$\mathbf{h}_i = \frac{\mathbf{a}_i}{\sum_{j=1}^{N_1} \mathbf{a}_{ij}}, i \in [1, N_1] \quad (20)$$

若第 i 个节点发生错传，则 $\mathbf{X}'_{N_1 \times 1} = (x_1, x_2, \dots, x'_i, \dots, x_{N_1})^T$ ，令

$$\mathbf{x}_i^* = \mathbf{X}^*(i, 1) = \mathbf{h}_i \times \mathbf{X}' = (h_{i1}, h_{i2}, \dots, h_{iN_1}) \times \begin{pmatrix} x_1 \\ \vdots \\ x'_i \\ \vdots \\ x_{N_1} \end{pmatrix} \quad (21)$$

其中， $\mathbf{X}^*(i, 1)$ 为经过邻居拓扑矩阵处理后的节点 i 的采集数据，丢弃节点 i 的错传数据，利用邻居节点的空间相关性，矩阵 \mathbf{H} 保证了相关性强的节点权值大，相关性弱的节点权值小，合成估计值 x_i^* ，使 $|x_i^* - x_i| \ll |x'_i - x_i|$ 。由式(19)可知，错传节点 i 对应的矩阵 \mathbf{H} 中第 i 列为 $\mathbf{0}$ ，所以错传节点 i 并不参与其需要估计的邻居节点的估计。由 \mathbf{X}^* 与观测矩阵 Φ 可得观测向量 \mathbf{Y}^* 。

$$\mathbf{Y}^* = \Phi \mathbf{X}^* \quad (22)$$

通过 \mathbf{Y}^* 重构 \mathbf{X} ，如式(23)，具体数据算法如算法 1 所示。

$$\hat{\mathbf{V}} = \arg \min \|\mathbf{V}\|_1, \text{ s.t. } \mathbf{Y}^* = \Phi \mathbf{X}^* = \Phi \Psi \mathbf{V} \quad (23)$$

算法 1 CS-NTSC 算法

输入 子节点感知数据 $\mathbf{X}'_{N_1 \times 1}(t)$ ，观测矩阵 Φ

输出 观测向量 $\mathbf{Y}^*_{M \times 1} = (y'_1, y'_2, \dots, y'_M)^T$

- 1) 簇首收到整个簇的邻居矩阵 \mathbf{A} ;
- 2) 当簇首从其子成员节点收到 $\mathbf{X}'_{N_1 \times 1}(t)$
- 3) for $i = 1: N_1$ 检测是否有错误数据分组
- 4) if 第 x'_i 个数据分组有错
- 5) for $j = 1: N_1$
- 6) $\mathbf{A}(j, i) = 0$;
- 7) end for
- 8) $\Omega \leftarrow i$ /* 记录错误节点号的索引 */;
- 9) end if
- 10) end for
- 11) for $i = 1: N_1$

$$12) \quad \mathbf{h}_i = \frac{\mathbf{a}_i}{\sum_{j=1}^{N_1} \mathbf{a}_{ij}} \quad /* \text{ 获取整个簇的邻居拓扑}$$

矩阵 \mathbf{H} */;

13) end for

14) for $i=1:N_1$

15) if $i \in \Omega$ then

$$16) \quad \mathbf{x}_i^* = \mathbf{X}^*(i,1) = \mathbf{h}_i \times \mathbf{X}';$$

17) end if

18) end for

19) for $i=1:M$

$$20) \quad \mathbf{y}_i^* = \sum_{j=1}^{N_1} \phi_{ij} \mathbf{X}_{N_1 \times 1}^* ;$$

21) end for

22) 发送 $\mathbf{Y}_{M \times 1}^*$ 至 sink;

23) end if

该算法的复杂度为 $O(N_1)$ ，簇首只需在首次数据传输前，根据网络中节点布设拓扑结构获取矩阵 \mathbf{A} ，在每轮数据传输时根据错传节点号和矩阵 \mathbf{A} 得到邻居拓扑矩阵 \mathbf{H} ；在数据处理时，簇首只需做简单的乘、加线性运算，无大量复杂运算，最终快速得到本簇观测向量，具有实时性。

5.2 CS-SSDG 算法

当网络发生拥塞时，网络一定范围内的多个节点均会发生伪失效，即伪失效节点与邻居节点均会传输数据失败，此时无法利用其空间相关性。针对节点伪失效模型，提出基于稀疏调度的压缩感知数据收集算法，通过改变观测矩阵的稀疏度，避免观测错误或丢失的数据，弱化错传和分组丢失对整体信息采集的影响，利用信息之间的相关性，重构原始数据。

当簇首收集到各簇成员发来带错数据 $\mathbf{X}'_{N_1 \times 1} = (x_1, x_2, \dots, x'_r, \dots, x_{N_1})^T$ 时，其中， x'_r 为丢失或错误的分组，簇首首先根据式(7)生成观测矩阵 $\Phi = (\phi)_{ij}$ ，再根据接收到数据 \mathbf{X}' 重新定义观测矩阵 Φ^* 。

假设观测矩阵 Φ 为

$$\Phi = \begin{pmatrix} \phi_{11} & \phi_{12} & \dots & \phi_{1r} & \dots & \phi_{1N_1} \\ \phi_{21} & \phi_{22} & \dots & \phi_{2r} & \dots & \phi_{2N_1} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \phi_{M1} & \phi_{M2} & \dots & \phi_{Mr} & \dots & \phi_{MN_1} \end{pmatrix} \quad (24)$$

当第 r 个节点发生伪失效时，令 $\Phi^{(r)} = \mathbf{0}$ ，即第 r 列置为 0，改变观测矩阵的稀疏度，则

$$\Phi^* = \begin{pmatrix} \phi_{11} & \phi_{12} & \dots & 0 & \dots & \phi_{1N_1} \\ \phi_{21} & \phi_{22} & \dots & 0 & \dots & \phi_{2N_1} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \phi_{M1} & \phi_{M2} & \dots & 0 & \dots & \phi_{MN_1} \end{pmatrix} \quad (25)$$

每轮数据收集时，并没有对节点 r 进行观测，对数据重构起作用的是其他没有发生分组丢失的节点，第 i 个观测值 y_i 如式(26)，此时将节点 r 对数据重构的影响降到了最低。当 sink 节点用矩阵 Φ^* 进行数据重构时，即可重构出原始数据，具体数据收集算法如算法 2 所示。

$$\mathbf{y}_i^* = \phi_{i1}x_1 + \phi_{i2}x_2 + \dots + 0 \cdot x'_r + \dots + \phi_{iN_1}x_{N_1} \quad (26)$$

算法 2 CS-SSDG 算法

输入 子节点感知数据 $\mathbf{X}'_{N_1 \times 1}(t)$ ，观测矩阵 Φ

输出 观测向量 $\mathbf{Y}_{M \times 1}^* = (y'_1, y'_2, \dots, y'_M)^T$

- 1) 当簇首从其子成员节点收到 $\mathbf{X}'_{N_1 \times 1}(t)$ ，then
- 2) for $i=1:N_1$
- 3) if 节点 r 发生伪失效
- 4) $\Phi^{*(r)} = \Phi(:,r) = \mathbf{0}$;
- 5) end if
- 6) end for
- 7) for $i=1:M$
- 8) $\mathbf{y}_i^* = \sum_{j=1}^{N_1} \phi_{ij}^* \mathbf{X}'_{N_1 \times 1}$;
- 9) end for
- 10) 发送 $\mathbf{Y}_{M \times 1}^*$ 至 sink;
- 11) end if

由算法 2 可知，该算法的复杂度为 $O(N_1)$ ，且簇首可根据收到的数据实时计算观测矩阵 Φ^* ，相关运算均为简单的线性运算，具有实时性与可使用性。

5.3 网络负载判断机制

簇首根据节点传输数据的历史情况，来判断网络负载情况，进而确定当前采用 CS-SSDG 还是 CS-NTSC 算法。具体判断策略如下：在首轮数据收集中，假设网络没有发生拥塞，即网络运行在轻负载下，节点的分组丢失模型为随机分组丢失模型。簇首在收集数据的同时，记录并存储数据传送失败的成员节点号；若在连续 3 个数据收集周期内均传输失败的节点个数超过本簇内成员节点总数的 $p_c\%$ ，此时簇首即认为网络内负载变重，节点分组丢失模型为节点伪失效模型。簇首判断网络负载具体流程如下。

输入 子节点感知数据 $X'_{N_i \times 1}(t)$

- 1) if 簇首从其子节点收到 $X'_{N_i \times 1}(t)$, then
- 2) for $i=1:N_1$ 校验是否有错误数据分组
- 3) if 第 x'_i 个数据分组有错
- 4) $\Omega_t \leftarrow i$ /* 记录第 t 轮错误节点号的索引*/;
- 5) end if
- 6) end for
- 7) $j=0$;
- 8) end if
- 9) for $i=1:N_1$
- 10) if $i \in \Omega_{t-2} \cap \Omega_{t-1} \cap \Omega_t$,
- 11) $j=j+1$;
- 12) end if
- 13) end for
- 14) if $j \geq p_c \% \cdot N_1$;
- 15) 簇首使用 CS-SSDG 算法;
- 16) else 簇首使用 CS-NTSC 算法;
- 17) end if

当簇首判断网络是重负载时，则簇首采用 CS-SSDG 算法；若在连续 3 个数据收集周期内均传输失败的节点个数少于本簇内成员节点总数的 $p_c\%$ ，则重新采用 CS-NTSC 算法。网络算法流程图如图 6 所示。

6 仿真与分析

为了验证算法的有效性，在 MATLAB 平台下进行仿真分析，仿真环境设置如下：在一个簇内，簇成员节点通过单跳将数据传输给簇首，由簇首进行数据压缩，将观测矩阵发送给 sink 节点，在 sink 处进行数据重构，发送数据分组长度为 10 byte；设定簇内有 900 个节点随机均匀分布在 30×30 区域内，簇内有一个事件源，其衰减系数 $n=0.01$ ，观

测次数 $M=450$ ；采用正交匹配追踪算法(OMP, orthogonal matching pursuit)作为重构算法。

本文假设在有一定误比特率的无线信道中，簇首接收到有错误的数据分组时不重传，利用一轮重构数据精度作为算法性能指标；同时，还将所提方法与基本的压缩感知数据收集(CDG, compressive data gathering)算法^[2]和 SRS-DG (sparsest random scheduling based CDG scheme)算法^[14]对比。文献[2]中传统 CS 数据收集算法利用压缩感知技术收集并重构数据，并不考虑误比特率对算法的影响，文献[14]中 SRS-DG 算法考虑到链路中存在分组丢失，利用构造的极稀疏矩阵 (sparsest measurement matrix) 观测无错数据。

6.1 CS-NTSC 算法仿真与分析

6.1.1 CS-NTSC 算法性能分析

当事件源衰减系数 $n=0.01$ ，邻居范围 $R_{thr}=2$ 时，CS-NTSC 算法性能如图 7 所示。在误码率较小 $Pb=10^{-5}$ 时，3 种算法性能接近；当误码率较高为 $Pb=10^{-3}$ 时，CS-NTSC 算法具有较好的性能，由图 7 可知，此时 CDG 算法的重构数据信噪比为 27.33 dB，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 29.72 dB，由于 SRS-DG 算法是针对分组丢失设计的算法，数据分组一旦有错传即丢弃，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，因此，本轮数据重构信噪比并不高；CS-NTSC 算法得到重构数据信噪比为 35.91 dB，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-NTSC 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。

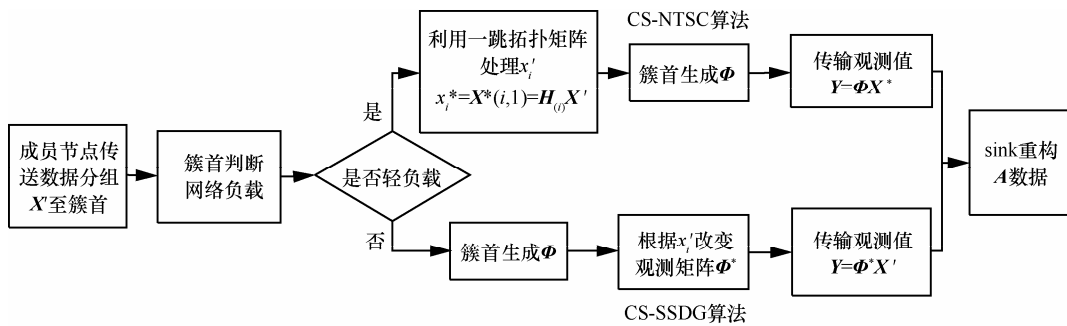


图 6 网络算法流程

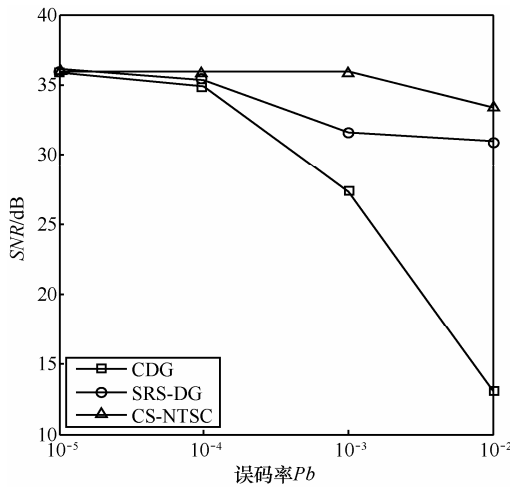


图 7 CS-NTSC 算法性能

6.1.2 影响 CS-NTSC 算法因素分析

CS-NTSC 算法利用了节点之间的空间相关性以降低不可靠链路对压缩感知数据收集的影响，所以事件的衰减系数 n 、邻居范围 R_{thr} 和数据分组长度都会影响 CS-NTSC 的性能，下面通过仿真讨论上述因素对算法性能的影响。

由于事件源的衰减系数 n 会影响全网数据的空间相关性，所以衰减系数会影响 CS-NTSC 算法性能，如图 8 所示，为 $R_{thr} = 3$, $P_b = 10^{-3}$, 数据分组为 10 byte 时，衰减系数 n 与 SNR 的关系。当衰减系数较小时，即事件源的影响范围大，此时全簇的节点均较大程度地受到事件源影响，节点空间相关性强，在邻居范围内，节点与节点采集数据接近，CS-NTSC 算法性能好；当衰减系数较大时，即事件源的影响范围小，距离事件源较远节点受到影响小，全簇节点的空间相关性弱，CS-NTSC 算法性能较差。当 $n > 0.5$ ，此算法的数据重构精度变差。

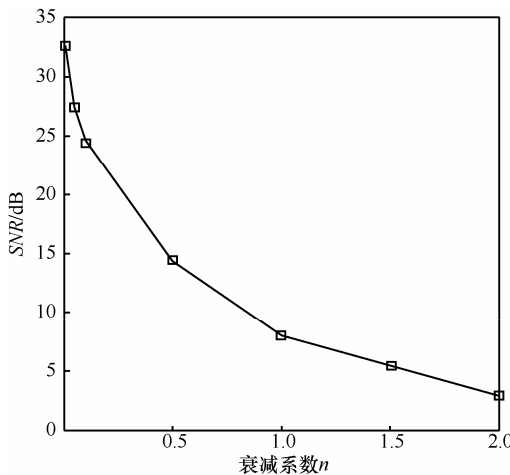


图 8 衰减系数与重构信噪比的关系

邻居范围 R_{thr} 也会影响算法性能，如图 9 所示，为信噪比和衰减系数不变时， R_{thr} 与 SNR 的关系。随着 R_{thr} 的增大，节点的邻居节点数会随之增加，在事件源影响范围不变的情况下， R_{thr} 过大会使邻居节点之间的空间相关性减小，CS-NTSC 算法性能变差，算法不再适用。

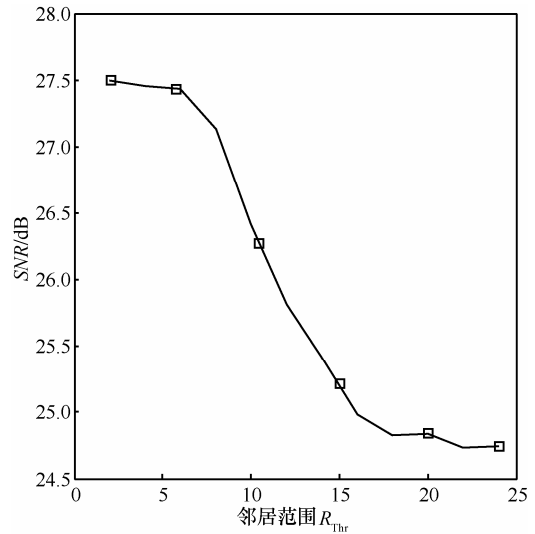


图 9 邻居范围与重构信噪比关系

信道误码率一定时，数据分组越长，会使网络的分组丢失率越高。图 10 反映了 $P_b = 10^{-4}$, $R_{thr} = 3$ 时，分组长与数据重构精度的关系。由图可知，在 n 较小，分组长较短的情况下，由于网络节点空间相关性强，分组丢失率低，此时 CS-NTSC 算法的数据重构精度高；当 n 增大，分组长变长时，节点空间相关性弱，分组丢失率高，此情况下 CS-NTSC 算法性能变差，不再适用。

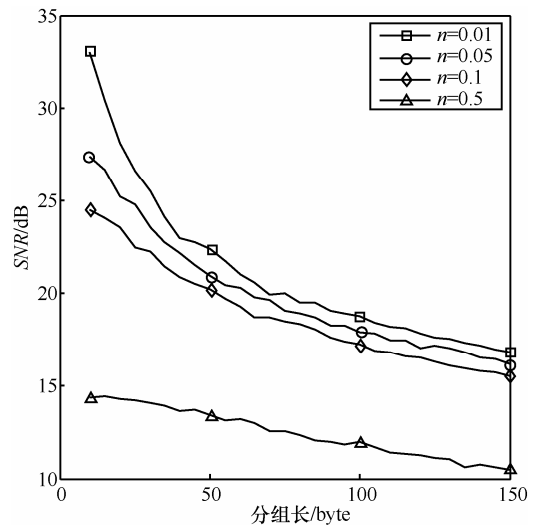


图 10 分组长与重构信噪比的关系 (CS-NTSC)

由以上仿真结果可以得出，当网络中事件源衰减系数较小，邻居范围小且数据分组短时，CS-NTSC 算法能够在不分组丢失也不重传的情况下，抵抗不可靠链路对数据重构的影响，以较高精度重构出数据，性能优势明显。

6.2 CS-SSDG 算法仿真与分析

CS-SSDG 算法性能如图 11 所示。当误码率 $P_b = 10^{-3}$ 较高时，由图 11 可知，CDG 算法性能很差，重构信噪比为 27.63 dB，即在高误码率的情况下，CDG 算法重构精度低，不具备抵抗不可靠链路的能力；SRS-DG 算法重构信噪比为 31.6 dB，此算法在本轮不调整观测次数的情况下，重构精度随着误码率的增加而降低且高于 CDG 算法；CS-SSDG 算法重构信噪比为 35.68 dB，无论误码率高还是低，CS-SSDG 算法都保持了较高的重构精度，该算法改变观测矩阵，避免观测发生失效节点的采集数据，减小其对数据重构的影响，利用正常节点采集数据高精度重构全网数据。由仿真结果可以看出，CS-SSDG 算法具有先进性与有效性。

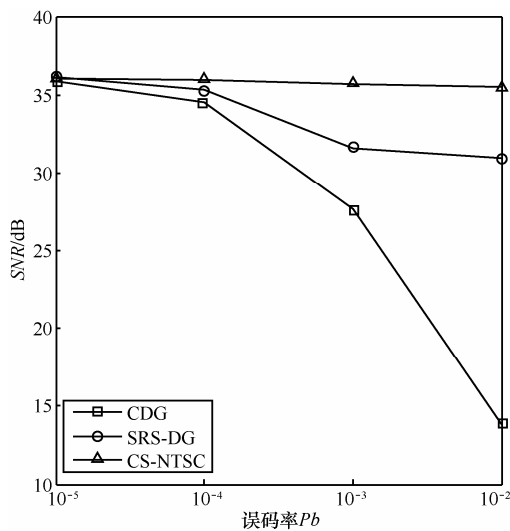


图 11 CS-SSDG 算法性能

同样，本文在 $P_b = 10^{-4}$ 和不同的 n 下，研究了分组长对算法 CS-SSDG 性能的影响，如图 12 所示。与 CS-NTSC 算法相同，随着 n 的增大，节点空间相关性变弱，CS-SSDG 重构信噪比也在减小。与 CS-NTSC 算法不同，当 n 不变时，算法 CS-SSDG 的重构信噪比对分组长并不敏感，分组长增长，会使分组丢失率变大，发生分组丢失的邻居节点数变少，估计误差变大，而算法 CS-SSDG 并不需要邻

居节点来估计分组丢失节点，所以在较高的分组丢失率下亦能高精度重构数据。

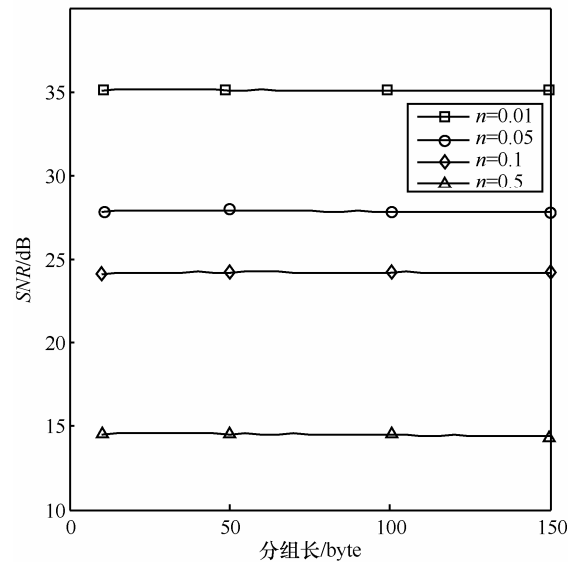


图 12 分组长与重构信噪比的关系 (CS-SSDG)

7 结束语

本文针对链路不可靠的无线传感网，首先利用仿真的方法，详细分析了不可靠链路对簇内压缩感知数据收集的影响，误码较高的信道环境对压缩感知数据收集具有很大影响。根据真实网络提出 2 种分组丢失模型，针对随机丢失模型，提出了 CS-NTSC 算法，分析了网络中的空间相关性，并利用空间相关性和邻居拓扑矩阵对发生错传的数据进行估计，减小错误幅度。讨论了衰减系数和邻居范围对 CS-NTSC 算法的影响，算法在衰减系数小和邻居范围小时，重构精度更高；针对节点伪失效模型，提出了 CS-SSDG 算法，将发生失效节点的相关观测列向量置为 $\mathbf{0}$ ，来避免收集发生失效的节点数据，只收集完整正确的数据，以减小错传或分组丢失对整个数据收集和重构的影响。仿真结果表明，CS-NTSC 算法与 CS-SSDG 算法能够在高误码率的环境中高精度重构出数据，具有有效性。

本文分别针对轻负载、随机分组丢失和重负载、伪随机失效 2 种情况研究了不可靠链路下的压缩感知数据收集算法，并简单讨论了 2 种算法的切换条件。然而在实际传感网中，节点分组丢失模型可能会出现 2 种分组丢失模型的混合，如何将算法有效地应用于混合分组丢失模型下，是下一步研究的重点。

参考文献:

- [1] RABBAT M, HAUPT J, SINGH A, et al. Decentralized compression and predistribution via randomized gossiping[C]//The 5th Int Conf on Information Processing in Sensor Networks. New York: ACM, 2006: 51-59.
- [2] LUO C, WU F, SUN J, et al. Compressive data gathering for large-scale wireless sensor networks[C]//The 15th Annual Int Conf on Mobile Computing and Networking. New York: ACM, 2009: 145-156.
- [3] WANG J, TANG S, YIN B, et al. Data gathering in wireless sensor networks through intelligent compressive sensing[C]// IEEE INFOCOM 2012. Piscataway, NJ: IEEE, 2012: 603-611.
- [4] DONOHO D L. Compressed sensing[J]. IEEE Trans on Information Theory, 2006, 52(4): 1289-1306.
- [5] BARANIUK R. Compressive sensing[J]. IEEE Signal Processing Magazine, 2007, 56(4): 4-5.
- [6] OSAMY W, SALIM A, AZIZ A. Efficient compressive sensing based technique for routing in wireless sensor networks[J]. Infocomp Journal of Computer Science, 2013, 12(1): 1-9.
- [7] LUO C, WU F, SUN J, et al. Efficient measurement generation and pervasive sparsity for compressive data gathering[J]. IEEE Trans on Wireless Communications, 2010, 9(12): 3728-3738.
- [8] LUO J, XIANG L, ROSENBERG C. Does compressed sensing improve the throughput of wireless sensor networks?[C]// IEEE Int Conf on Communications (ICC 2010). New York: IEEE Communications Society, 2010: 1-6.
- [9] WU X, XIONG Y, HUANG W, et al. An efficient compressive data gathering routing scheme for large-scale wireless sensor networks[J]. Computers & Electrical Engineering, 2013, 39(6): 1935-1946.
- [10] AKYILDIZ I F, SU W, SANKARASUBRAMANIAM Y, et al. Wireless sensor networks: a survey[J]. Computer Networks, 2002, 38(4): 393-422.
- [11] NDZI D L, ARIF M A M, SHAKAFF A Y M, et al. Signal propagation analysis for low data rate wireless sensor network applications in sport grounds and on roads[J]. Progress in Electromagnetics Research, 2012, 125(17):1-19.
- [12] AHMED N, KANHERE S S, JHA S. Utilizing link characterization for improving the performance of aerial wireless sensor networks[J]. IEEE Journal on Selected Areas in Communications, 2013, 31(8): 1639-1649.
- [13] BACCOUR N, KOUBAA A, MOTTOLA L, et al. Radio link quality estimation in wireless sensor networks: a survey[J]. ACM Transactions on Sensor Networks, 2012, 8(4):688.
- [14] WU X, YANG P, JUNG T, et al. Compressive sensing meets unreliable link: sparsest random scheduling for compressive data gathering in lossy WSN[C]//The 15th ACM Int Symposium on Mobile Ad Hoc Networking and Computing. New York: ACM, 2014: 13-22.
- [15] 唐亮, 周正, 石磊, 等. 基于 LEACH 和压缩感知的无线传感网目标探测[J]. 北京邮电大学学报, 2011, 34(3): 8-11.
TANG L, ZHOU Z, SHI L, et al. Source detection in wireless sensor network by leach and compressive sensing[J]. Journal of Beijing University of Posts & Telecommunications, 2011, 34(3): 8-11.
- [16] 张策, 张霞, 李鸥, 等. 基于 CS 的无线传感网动态分簇数据收集算法[J/OL]. <http://crad.ict.ac/cn/CN/abstract/abstract3059.shtml>.
ZHANG C, ZHANG X, LI O, et al. Data gathering using dynamic clustering based on WSN compressive sensing algorithm[J/OL]. <http://crad.ict.ac/cn/CN/abstract/abstract3059.shtml>.
- [17] WANG W, GAROFALAKIS M, RAMCHANDRAN K. Distributed sparse random projections for refinable approximation[C]//The 6th Int Conf on Information Processing in Sensor Networks. New York: ACM, 2007: 331-339.
- [18] KONG L, XIA M, LIU X Y, et al. Data loss and reconstruction in sensor networks[C]// IEEE INFOCOM 2012. Piscataway, NJ: IEEE, 2013: 1654-1662.
- [19] WU L, YU K, DU T, et al. Efficient information transmission under lossy WSNs link using compressive sensing[C]//2014 IEEE 9th Conference on Industrial Electronics and Applications (ICIEA). NJ: IEEE, 2014:493-498.

作者简介:



张策 (1991-), 男, 四川南充人, 解放军信息工程大学博士生, 主要研究方向为无线自组织网络、无线传感网与路由协议。



张霞 (1979-), 女, 山东济南人, 博士, 解放军信息工程大学讲师, 主要研究方向为无线传感网、信息处理与流量识别。



李鸥 (1961-), 男, 陕西宝鸡人, 博士, 解放军信息工程大学教授、博士生导师, 主要研究方向为无线传感网、认知无线网络与无线自组织网络。

梅关林 (1989-), 男, 四川泸州人, 解放军信息工程大学硕士生, 主要研究方向为无线通信、卫星调度。

韩哲 (1991-), 男, 河南洛阳人, 解放军信息工程大学硕士生, 主要研究方向为无线通信、无线传感器网络。

张大龙 (1976-), 男, 河南郑州人, 博士, 郑州大学讲师, 主要研究方向为无线通信、无线传感网与 MAC 协议。

刘广怡 (1982-), 男, 河南郑州人, 博士, 解放军信息工程大学讲师, 主要研究方向为传感网、智能算法、网络数据分析与物联网。