

文章编号:1005-0523(2018)02-0113-07

# WSN 中改进的基于压缩感知的分簇数据采集算法

谢 昕<sup>1</sup>,汪加楠<sup>1</sup>,姜 楠<sup>1</sup>,黄晓生<sup>2</sup>,葛松林<sup>1</sup>

(华东交通大学 1.信息工程学院;2.软件工程学院,江西 南昌 330013)

**摘要:** WSNs 中节点的能量受限,为节约网络节点的功耗,提出了一种将 LEACH 模型和 CS 理论相结合的基于节点时空相关性的压缩感知算法。LEACH 算法对网络节点进行分簇并选择簇首,然后,采用 CS 理论对簇首节点进行采样,将采样后的数据通过多跳路由的方法传至远端的汇聚节点,最后,汇聚节点对来自簇首的少量数据采用重构算法以对原始信号进行精确恢复。仿真结果表明,该方法有效减小了数据传输量,降低了网络节点的功耗,延缓了网络寿命。

**关键词:** 无线传感器网络;压缩感知;LEACH 协议;时空相关性

**中图分类号:** TP393.4

**文献标志码:** A

无线传感器网络(wireless sensor networks, WSNs)是由众多传感器节点组成的一种无线自组织网络,被看作连接人类社会与物理世界的纽带。但是 WSNs 中的网络节点分布十分密集,而且节点的能量、计算与存储能力也受到限制。如何确保监测信息不丢失,尽量减少数据传输量,同时尽可能地延缓网络寿命,就变成无线传感器网络设计中的焦点。

众多学者从不同方面提出并设计了降低网络能耗的算法。如 Carol Habib 等提出了 Modified LED 算法<sup>[1]</sup>,将测量数据离散化为一些有限的等间隔区间,再由等间隔区间编号变化来决定是否传输数据,算法虽然可以在一定程度上减少数据传输量,但极大地削弱了数据的精度,不适宜数据精度要求高的应用场合。压缩感知<sup>[2]</sup>(compressed sensing, CS)能够利用远低于奈奎斯特采样速率采集的少量数据,完成对信号的获取,这为降低无线传感器节点的通信能耗提供了很好的解决思路。赵贻玖等利用压缩感知原理<sup>[3]</sup>,将奈奎斯特采样得到的数据进行压缩传输,降低了通信能耗,但这并不是真正意义上的压缩感知。胡海峰等进行了压缩采样的研究<sup>[4]</sup>,由于传统压缩感知技术的实现对硬件资源要求比较高<sup>[5]</sup>,无法直接将其应用于节点资源受限的无线传感器网络。Osamy W 等将 CS 技术与 Pegasus 路由协议相结合<sup>[6]</sup>,构建路由链,在链中压缩数据以改善网络寿命,在 Sink 处统一重构数据,与树形路由中最小生成树相比,虽然链路路由中节点之间每一跳能耗最小,但整个链路能耗不一定达到最小,且网络鲁棒性差。为了减少网络节点的通信量以节省能耗, Li Shi 等提出多种网内数据融合方法如使用 BP 神经网络<sup>[7]</sup>、多区域能量感知等<sup>[8]</sup>,能在某一特定应用场合起到很好的效果,但是算法复杂且不具有通用性。

针对以上局限性,本文综合考虑了实际采集过程所获取数据的时空相关性,提出了一种将 LEACH(low energy adaptive clustering hierarchy)模型与 CS 理论相结合的方法。该方法在分簇路由的基础上,采用 CS

收稿日期:2017-11-06

基金项目:国家自然科学基金项目(61762037,61763011,41402290,61462028);江西省科技厅重点研发计划(20151BBE50055);江西省教育厅科技项目(GJJ150541);南昌市科技局知识创新团队项目(2016T75)

作者简介:谢昕(1969—),男,教授,研究方向为机器视觉和网络与信息安全。

理论对簇首节点进行采样,将采样后的数据通过多跳路由的方法传至远端的汇聚节点,汇聚节点对来自簇首的少量数据采用重构算法以对源数据进行精确恢复<sup>[9]</sup>。对节省网络节点的功耗十分有效,延缓了网络的有效工作时间。

## 1 LEACH 算法

LEACH 是由 MIT 的 Heinzelman 等人提出的一种低功耗自适应分簇路由协议,它将 WSNs 中的节点分成多个簇,在节点划分完簇后,寻找 1 个节点当选相应簇的簇首。首先,簇内节点把自身信息发送至所属的簇首,其次,由各自的簇首直接和汇聚节点进行通信,最后,将采集到的信息发送出去<sup>[10-11]</sup>。

WSNs 中的每个节点将自身信息乘一个相应的权值系数,之后传递到簇首,簇首将传来的节点信息和本地信息相加,把得出的信息传输到汇聚节点,最后,汇聚节点对来自簇首的少量数据采用重构算法以对原始信号进行精确恢复。LEACH 算法的过程为:

$$y_i = \sum_{k=1}^{K_i} w_{i,k} x_{i,k} + n_i \quad (1)$$

式中, $y_i$  是第  $i$  个簇的簇首发送的数据; $x_{i,k}$  是第  $i$  个簇中第  $k$  个节点上的数据; $n_i$  是第  $i$  个簇中各节点的数据发送至簇首再由簇首传递至汇聚节点的过程的总噪声。假定  $n_i$  符合高斯分布  $N(0, \beta^{-1})$ ;  $w_{i,k}$  是第  $i$  个簇内第  $k$  个网络节点的信息权值系数; $K_i$  是第  $i$  个簇中的节点个数。令权值矩阵  $\mathbf{W}$  的第  $i$  行  $k$  列元素是  $w_{i,k}$ , 根据式(1)得到:

$$y = \mathbf{W}\mathbf{x} + \mathbf{n} = \mathbf{W}\boldsymbol{\psi}\mathbf{v} + \mathbf{n} = \boldsymbol{\Gamma}\mathbf{v} + \mathbf{n} \quad (2)$$

其中  $\mathbf{W}$  为  $M \times N$  矩阵。每个簇对应  $\mathbf{W}$  的一行,每个节点对应  $\mathbf{W}$  的一列。所以,若  $\mathbf{W}$  的第  $i$  行  $k$  列元素相应的节点不归属第  $i$  个簇,那么它的元素值为 0。 $\mathbf{x}$  为节点信息向量; $\mathbf{n}$  为观测噪声向量; $\boldsymbol{\psi}$  为稀疏基矩阵; $\boldsymbol{\Gamma}$  为  $\mathbf{W}$  和  $\boldsymbol{\psi}$  的乘积; $\mathbf{v}$  是稀疏系数组成的向量, $\mathbf{v}$  中非零元素的个数远少于维数,所以原始信号有稀疏性。

## 2 压缩感知算法的描述

### 2.1 压缩感知算法

如果有一有限长度的离散信号  $X \in R^n$ , 使此元素为  $X[n], n=1, 2, \dots, N$ 。对  $R^n$  空间的任意信号而言,都能够通过  $N \times 1$  维的基向量  $\{\psi_i\}_{i=1}^N$  的线性组合来表达。则任一信号  $X$  可表示为:

$$X = \sum_{i=1}^N \theta_i \psi_i \text{ 或 } X = \boldsymbol{\psi}\boldsymbol{\theta} \quad (3)$$

式中: $\theta_i, (i=1, 2, \dots, n)$  为投影系数;其组成的  $N \times 1$  的列向量是  $\boldsymbol{\theta} = [\theta_i]_{i=1}^N$ 。

信号在时域中的表示为  $X$ , 在  $\boldsymbol{\psi}$  域中的表示为  $\boldsymbol{\theta}$ 。若  $\boldsymbol{\theta}$  的非零个数不多,那么说明  $X$  是能够压缩的,即  $X$  在  $\boldsymbol{\psi}$  域下能够用  $K$  个系数表达。压缩感知理论用新的采集数据方式以及仅要少量的测量数据替代了以往的压缩方法,而且实现了压缩的同时完成采样,之后再传递或存储压缩后的数据,极大地节省了网络节点的功耗。压缩感知理论规定  $M \times N$  维的测量矩阵  $\boldsymbol{\Phi}$  与基矩阵  $\boldsymbol{\psi}$  不相关 ( $K < M \leq N$ )<sup>[12-13]</sup>, 那么压缩后的信号  $Y$  如下所示:

$$Y = \boldsymbol{\Phi}X = \boldsymbol{\Phi}\boldsymbol{\psi}\boldsymbol{\theta} = A^{\text{CS}}\boldsymbol{\theta} \quad (4)$$

式中:CS 理论的信息算子是  $A^{\text{CS}}$ ; 独立同分布的高斯随机矩阵是  $\boldsymbol{\Phi}$ , 确保了  $A^{\text{CS}}$  符合约束等距性 RIP。在符合此要求下,重构原始信号只要知道  $Y$  就可。此时,求  $\boldsymbol{\theta}$  的近似解就转化为  $l_0$  范数的最优化问题:

$$\hat{\boldsymbol{\theta}} = \arg \min \|\boldsymbol{\theta}\|_0 \quad \text{s.t.} \quad A^{\text{CS}}\boldsymbol{\theta} = Y \quad (5)$$

最后再利用式(3)求解出  $X$ 。

把压缩感知理论应用在无线传感器网络中能够减少能量消耗, 景博等阐述了当前应用在 WSNs 中的 Berkeley Motes 传感器的通讯模块在休眠状态的能耗是 0.03 mW, 发送状态的能耗是 60 mW, 空闲和接收状态下的能耗是 12 mW<sup>[14]</sup>, 显然压缩感知理论通过降低节点间的数据传输量的方式, 以节省网络节点的能耗

具有较为理想的效果。

假若 WSNs 中采集到的数据不必实时传递,而要存储后定时传输至簇首节点,簇首节点再通过多跳路由的方法传至远端的汇聚节点。为降低 WSNs 中数据的重复及交错传递,假定每个簇首节点至多能够与邻近的 8 个节点进行通讯,如果簇首节点的能量小于确定的能量阈值,那么将向此区域传递相应指令,选定比此簇首节点的能量大许多的普通节点当选新的簇首节点,以实现能量均衡,从而减小极少数节点失效的概率。CS 理论应用于无线传感器网络的数据传输过程如图 1。

## 2.2 基于时空相关性的压缩感知算法

WSNs 中的节点完成数据采集后,LEACH 算法对各节点进行分簇并选择簇首,然后,采用 CS 理论对簇首节点进行采样,将采样后的数据通过多跳路由的方法传至远端的汇聚节点,最后,汇聚节点对来自簇首的少量数据采用重构算法以对原始信号进行精确恢复。在能耗和均衡方面该方法拥有很好的优越性。通常算法一般运用 WSNs 中数据间的空间相关性来传递数据,在 WSNs 对实时性的需求不特别高的场合,将 CS 应用于数据间的时空相关性,可以更理想地压缩数据,以减少网络节点的功耗,延缓无线传感器网络的生存时间。具体过程如下:

假定在选定的监测区域中,有  $N$  个无线传感器布置在某一个簇首节点周围,每个无线传感器节点在相同时间内自动地采集  $M$  个数据。那么第  $i$  个节点采集的数据 ( $i=1,2,\dots,N$ ) 为:

$$x_i=[x_{i1},x_{i2},\dots,x_{im}]^T \quad (6)$$

故在簇首节点处,接收到的  $N$  个节点数据如下:

$$X=\begin{bmatrix} x_{11} & x_{21} & \cdots & x_{n1} \\ x_{12} & x_{22} & \cdots & x_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ x_{1m} & x_{2m} & \cdots & x_{nm} \end{bmatrix} \quad (7)$$

在簇首节点,联合数据之间的时空相关性,运用矩阵向量量化函数,把得到的  $M \times N$  矩阵转化为  $(M \times N, 1)$  向量:

$$\text{rvec}(X)=(x_{11},\dots,x_{n1},x_{1,2},x_{2,2}\dots,x_{n-1,m-1},x_{n,m})^T \quad (8)$$

令此数据下的稀疏基为  $\psi_1$  和  $\psi_2$ ,测量矩阵为  $\Phi$ ,运用 Kronecker 积把式(4)转换成:

$$Y=\Phi \text{rvec}(X)=\Phi \text{rvec}(\psi_1 \theta \psi_2)=\Phi (\psi_2^T \otimes \psi_1) \text{rvec}(\theta) \quad (9)$$

式中:

$$\psi_1=\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & -1 & 1 \end{bmatrix}_{M \times M} \quad (10)$$

$$\psi_2=\begin{bmatrix} 1 & -1 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix}_{N \times N} \quad (11)$$

式(7)中某一系列数据的相关性,表现为 WSNs 中节点间的时间相关性;同理,某一行数据的相关性,表现为 WSNs 中节点间的空间相关性,采用此方法能用数学语言来巧妙地表达 WSNs 中相邻节点间的时空相关

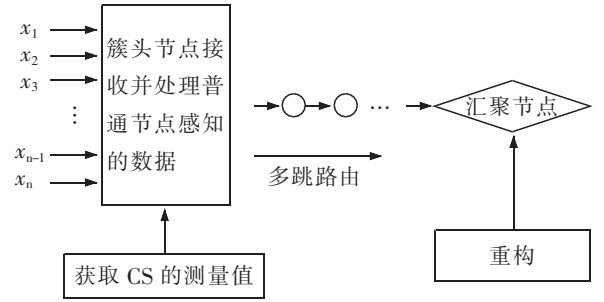


图 1 压缩感知算法

Fig.1 Compressed sensing algorithm

性。在分析实际采集过程所获取的数据时,观察到式(8)中邻近元素间的数值区分很小,稀疏分解完,稀疏性更理想。通过差分矩阵稀疏分解后,在择取测量矩阵时,为了实现尽可能理想地恢复原始信号,故规定测量矩阵和稀疏基两者间要符合约束等距性<sup>[15]</sup>。随机高斯矩阵被选为本文的测量矩阵,为将计算复杂度减小,故把随机高斯矩阵  $\Phi$  作以下变换:

$$\Phi = \begin{bmatrix} \Phi_1 & 0 & \cdots & 0 \\ 0 & \Phi_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \Phi_n \end{bmatrix} \quad (12)$$

式中  $\Phi_i(i=1,2,\dots,n)$ 指的是各节点所选取的测量矩阵。接下来更深入研究的内容则为最优测量矩阵的择取。

采用 CS 理论对簇首节点采样,把采样后的数据传至远端的汇聚节点,汇聚节点对来自簇首的少量数据采用重构算法以对原始数据进行精确恢复。本文选择 OMP 迭代算法作为恢复原始信号时的重构算法<sup>[16-17]</sup>。在详细阐述该 OMP 迭代算法前作下述规定:

相关系数  $u$ :

$$u = \{u_j | u_j = |r, \Phi_j|, j=1, 2, \dots, M\} \quad (13)$$

对于信号的逼近,则采用最小二乘法:

$$\hat{x} = \underset{x \in \Lambda}{\operatorname{argm}} \operatorname{in} \|y - \Phi_{\Lambda} x\|_2 \quad (14)$$

余量的更新定义为:

$$r_{\text{new}} = y - \Phi_{\Lambda} \hat{x} \quad (15)$$

OMP 迭代算法的框图如图 2 所示。

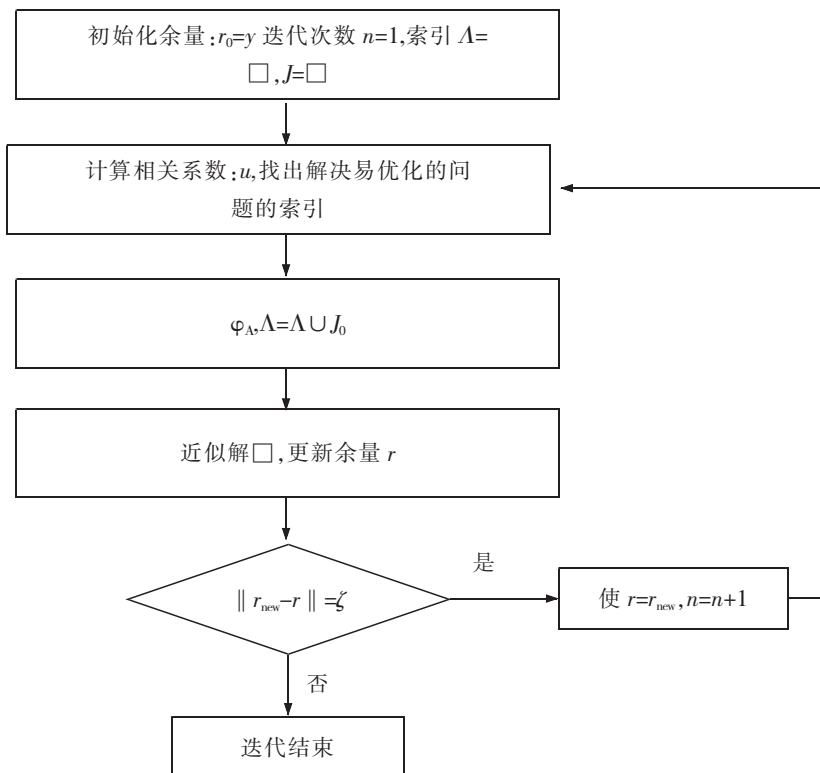


图 2 OMP 算法框图

Fig.2 OMP algorithm diagram

### 3 仿真结果与分析

本文采用 EPFL SensorScopeWSN 内实际测量所得的环境温度数据,为方便研究,只选择 8 个 WSNs 节点,每个节点采集 32 个数据。然后由式(8)可得到(8×32,1),根据式(9)测量数据,最后利用 OMP 迭代算法只用少量的数据就可以对源数据进行精确恢复。对于重构误差而言,选取相对重构误差:

$$\varepsilon = \frac{\|rvec(x) - rvec(x^*)\|_2}{\|rvec(x)\|_2} \quad (16)$$

式中:原始信号变换为向量后的结果为  $rvec(x)$ ;重构之后的结果为  $rvec(x^*)$ 。

实验结果如图 3。为方便对比,使其和只结合数据间空间相关性的重构效果对比,由仿真结果知,运用数据间的时空相关性以对原始信号进行精确恢复时(相对重构误差  $\varepsilon=O(10^{-15})$ ),测量值个数至少为 43 个,然则仅结合数据间空间相关性对原始信号进行精确恢复时需要的测量值个数至少为 95 个,对于数据的压缩效果而言,本文提出的算法更为理想。接下来以簇首节点的能耗方面来检验该模型的优越性。

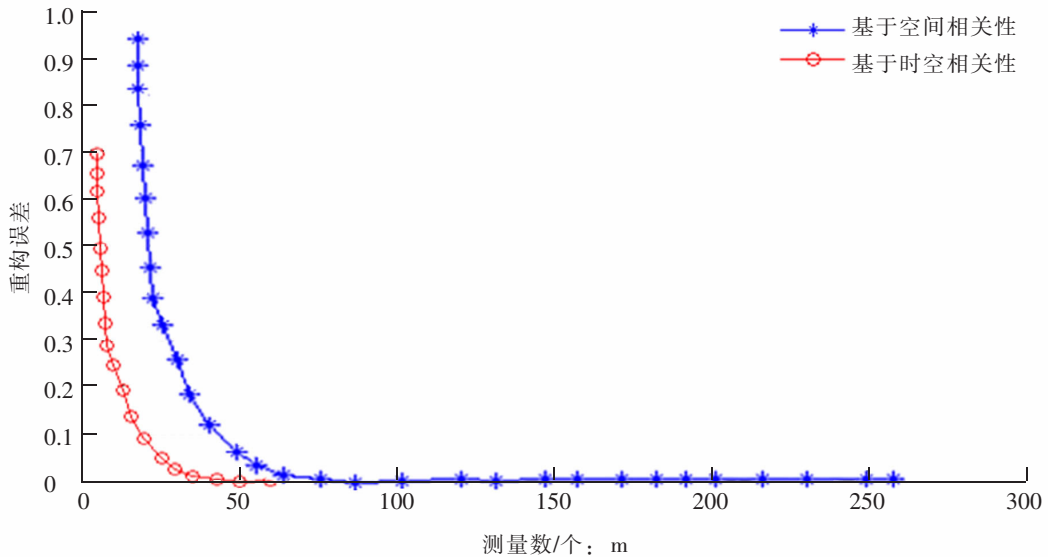


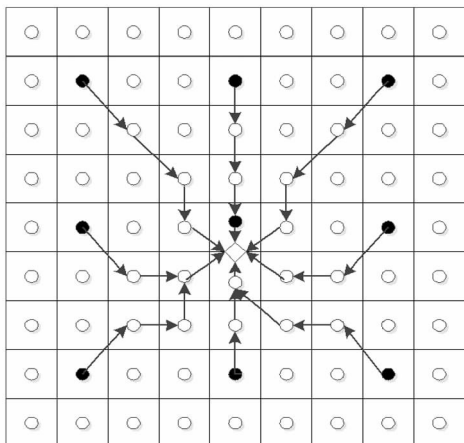
图 3 基于时空相关性和空间相关性的重构效果

Fig.3 Reconstruct comparison based on spatial-temporal correlation and spatial correlation

运用图 4 描述的分簇栅格网络,网络的详细参数如表 1。

表 1 网络参数

Tab.1 Network parameters



○ 簇头节点    ◇ 汇聚节点    ○ 感知节点

图 4 基于分簇的栅格网络模型

参数	值
节点数量	81
区域大小/m <sup>2</sup>	81
汇聚节点位置	(50,48)
簇首节点数量	9
初始能量/J	2
发送功耗/(nJ·bit <sup>-1</sup> )	50
放大功耗/(pJ/bit·s <sup>-1</sup> ·m <sup>-2</sup> )	10
广播数据包/byte	25
时分多址数据包/byte	43
数据包/byte	25
数据/byte	300
解码功耗/(pJ·bit <sup>-1</sup> )	20

Fig.4 Grid network model based on clustering



不同算法分簇时的能耗大致相等<sup>[18]</sup>,故先不研究分簇时的能耗因素。对比分簇后簇首节点的功耗,改进的算法使簇首节点和汇聚节点之间的信息量降低到了未采用任何算法的17%即本要发送的数据为255个,但应用该算法只要发送的数据为43个,把改进的算法与LEACH算法及改进的HEED算法对簇首节点的能量消耗作比较,对于簇首节点而言,另2个算法需采集的数据个数要多于改进后的算法(对比结果如图5)。

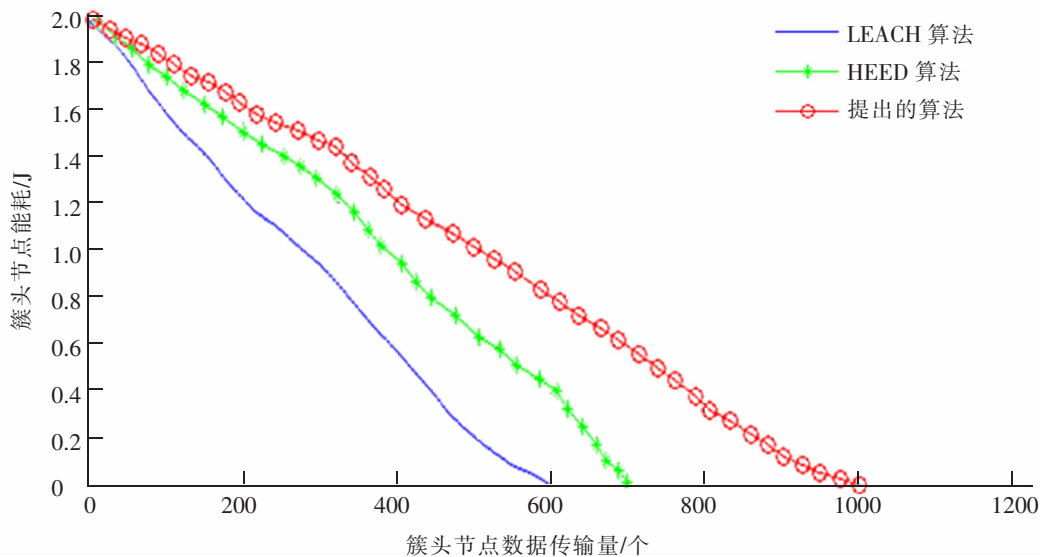


图5 簇首节点能量消耗对比

Fig.5 Comparison of cluster head energy consumption

## 4 总结

本文提出了一种基于节点时空相关性的压缩感知算法,用LEACH算法对网络节点进行分簇并选择簇首,再采用CS理论对来自簇首的少量数据采用重构算法以对原始信号进行精确恢复,有效减少了向汇聚节点传递的数据量。仿真结果显示,在满足相对重构误差条件下,测量值 $M$ 对应减少,节省了能量消耗。对于WSNs节点分布十分密集的应用场合,普通节点和簇首节点相距不远时,本算法比较理想。然而普通节点和簇首节点相距较远时,单个节点通讯的能量消耗没有减少,效果并非很理想,这是因为WSNs中普通节点和簇首间的数据传输量并未降低,距离的平方越大,则通信成本越高。这也需要在后续作进一步的研究。

## 参考文献:

- [1] HABIB C, MAKHOUL A, DARAZI R, et al. Self-Adaptive data collection and fusion for health monitoring based on body sensor networks[J]. IEEE Transactions on Industrial Informatics, 2016, 12(6): 2342-2352.
- [2] DONOHO D L. Compressed sensing[J]. IEEE Transactions on Information Theory, 2006, 52(4): 1289-1306.
- [3] 赵贻玖, 戴志坚, 王厚军. 基于压缩感知理论的随机等效采样信号的重构[J]. 仪器仪表学报, 2011, 32(2): 247-251.
- [4] 胡海峰, 杨震. 无线传感器网络中基于空间相关性的分布式压缩感知[J]. 南京邮电大学学报: 自然科学版, 2009, 29(6): 12-16.
- [5] BAJWA W, HAUPT J, SAYEED A, et al. Compressive wireless sensing[C]//International Conference on Information Processing in Sensor Networks. ACM, 2006: 134-142.
- [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] YADAV SGS, CHITRA A. An energy aware framework for multiple-zone data fusion technique in WSN[J]. International Journal of Applied Engineering Research, 2016, 11(4): 2263-2270.

- [8] LI S, LIU M, XIA L. WSN data fusion approach based on improved BP algorithm and clustering protocol[C]// Control and Decision Conference. IEEE, 2015: 1450–1454.
- [9] XIE X, XU Y, LIU Q, et al. A study on fast SIFT image mosaic algorithm based on compressed sensing and wavelet transform[J]. Journal of Ambient Intelligence and Humanized Computing, 2015, 6(6): 835–843.
- [10] 严斌亨, 刘军, 刘广斌, 等. 基于改进蚁群算法的 LEACH 协议研究[J]. 计算机测量与控制, 2016, 24(12): 136–138.
- [11] 彭蕾, 吕敬祥, 刘秋平. 大规模无线传感网络的混合 LEACH 协议研究[J]. 传感技术学报, 2016, 29(11): 1737–1741.
- [12] 党小超, 刘颜星, 郝占军. 无线传感网中一种用于压缩感知的测量矩阵优化方法[J]. 小型微型计算机系统, 2017, 38(9): 1966–1970.
- [13] 李哲涛, 臧浪, 田淑娟, 等. 基于混合压缩感知的分簇式网络数据收集方法[J]. 计算机研究与发展, 2017, 54(3): 493–501.
- [14] 景博, 张劫, 孙勇. 智能网络传感器与无线传感器网络[M]. 北京: 国防工业出版社, 2011: 39–40
- [15] 石光明, 刘丹华, 高大化, 等. 压缩感知理论及其研究进展[J]. 电子学报, 2009, 37(5): 1070–1081.
- [16] 马博珩, 彭艺. 基于 OMP 算法的快速压缩感知图像重构[J]. 云南大学学报: 自然科学版, 2017, 39(2): 207–211.
- [17] 邵炫, 孙超, 孔德智, 等. 一种基于正交匹配追踪的水下目标相关检测方法[J]. 西北工业大学学报, 2017(4): 622–628.
- [18] 谢昕, 吴颖, 张磊, 等. 基于无线传感器网络节点的 RFID 系统节能研究[J]. 传感器与微系统, 2012, 31(6): 66–68.

## Improved Clustering Algorithm for Data Acquisition Based on Compressed Sensing in WSN

Xie Xin<sup>1</sup>, Wang Jianan<sup>1</sup>, Jiang Nan<sup>1</sup>, Huang Xiaosheng<sup>2</sup>, Ge Songlin<sup>1</sup>

(1. School of Information Engineering, East China Jiaotong University, Nanchang 330013, China; 2. School of Software Engineering, East China Jiaotong University, Nanchang 330013, China)

**Abstract:** The energy of the sensor network nodes is limited, and in order to save the energy consumption of the sensor nodes, an improved compressed sensing algorithm combining the LEACH model with the CS theory based on the spatial-temporal correlation of nodes was proposed. The LEACH algorithm was used to cluster the network nodes and select the cluster head. Then, the cluster head node was sampled by the compressed sensing theory. The sampled data was passed to the remote sink node through multi-hop routing. Finally, a small amount of data from the cluster head used a reconstruction algorithm to accurately recover the original signal. The simulation results show that this method can effectively reduce the amount of data transmission, save the energy consumption of nodes and prolong the lifetime of the wireless sensor network.

**Key words:** wireless sensor network; compressed sensing; leach protocol; spatial-temporal correlation