

一种压缩感知的无线传感器网络分簇算法

曹欲晓, 钱 瑛, 韩 磊, 俞 翔

(南京工程学院 计算机工程学院, 江苏 南京 211167)

摘 要: 当前 WSN 分簇算法, 主要从优化节点间拓扑关系的角度出发进行协议设计, 没有考虑通过减少网络中传输的数据量来延长 WSN 的生存时间. 新兴的压缩感知算法可以通过在前端进行简单的运算来大幅度地压缩原始数据. 由此提出一种压缩感知的 WSN 分簇算法 CSCA. 首先根据最优簇数目对 WSN 进行区域划分, 然后由基站根据节点参数选出簇头并成簇. WSN 在进行数据传输时, 簇中节点对监测数据在感知矩阵上投影之后发送给簇头, 簇头对各节点的数据做差之后再进行一次投影之后发送给基站. 仿真实验证明, CSCA 算法由于减少了网络内的数据流量从而有效地降低了节点的能耗, 延长了 WSN 的生存时间.

关键词: 无线传感器网络; 分簇; 压缩感知; 稀疏; 投影

中图分类号: TP393

文献标识码: A

文章编号: 1000-7180(2015)11-0059-05

A Kind of Compressed Sensing Clustering Algorithm for Wireless Sensor Network

CAO Yu-xiao, QIAN Ying, HAN Lei, YU Xiang

(School of Computer Engineering, Nanjing Institute of Technology, Nanjing 211167, China)

Abstract: At present the emphasis of WSN's clustering algorithm is to optimize the topology of nodes in WSN, and it is not considered to prolong the WSN's lifetime by decreasing the data flow in network. New compressed sensing algorithm may compress original data in vast scale by simple operation at front-end. A kind of compressed sensing clustering algorithm named CSCA is put forward. Firstly the WSN's area is divided into several sections which number equals to the best cluster number. Secondly the base station selects cluster head according to related parameter and the cluster is made up. When transporting data node projects monitoring data on sensing matrix and sends projection data to cluster head. Then cluster head subtracts node's projection data and projects the difference. Lastly the cluster head's projection data is sent to base station. The emulation result proves that CSCA can decrease node's energy cost and prolong the WSN's lifetime obviously.

Key words: wireless sensor network; cluster; compressed sensing; sparse; projection

1 引言

无线传感器网络 WSN (Wireless Sensor Network) 由部署在监测区域的大量传感节点构成, 每个节点集测量、计算、存储、通信等功能于一身. WSN 的节点一般由电池供电, 大多数情况下不能更换, 因此如何减少 WSN 的能量消耗, 延长网络的生存时间始终是 WSN 领域的研究热点.

WSN 按节点之间的拓扑关系可以分成平面结

构和分簇结构两种. 分簇结构比平面结构具有更好的灵活性、可伸缩性、可扩展性, 能更有效地利用节点能量, 延长网络的生存期^[1]. LEACH、PEGASIS、TEEN 等都是经典的 WSN 分簇算法. LEACH^[2] 协议的每一轮随机产生簇头, 选定簇头后节点再根据簇头信号强度选择加入不同的簇. PEGASIS^[3] 协议把网络中的所有节点构成一条链, 每轮全网只选一个节点做簇头, 链中相邻的节点互相通信, 把信息汇集到簇头. TEEN^[4] 协议是在 LEACH 协议的基础

上改进得到的,分簇方法与 LEACH 相同,但是在发送数据时以阈值控制数据的发送时机.随着群体智能算法的发展,也有研究者把智能算法应用到 WSN 的分簇之中.文献[5-6]中应用粒子群算法对 WSN 进行了分簇,文献[7]在 WSN 的分簇过程中使用了遗传算法,文献[8]中应用了蚁群算法求解 WSN 的分簇.

上述分簇算法在降低节点能量消耗,延长网络生存时间方面均取得了良好的效果.但是目前的 WSN 分簇算法,核心思想都是优化网络的拓扑关系,选择节点间最优的通信路径,而没有考虑如何通过减少网络内的数据流量来降低 WSN 的能量消耗.随着压缩感知 CS(Compressed Sensing)技术的出现,使得采用少量的压缩感知测量数据来精确重构原始数据成为可能.因此本文从减少网络内的数据流量出发,提出了一种新的 WSN 分簇算法—压缩感知的分簇算法 CSCA(Compressed Sensing Clustering Algorithm).CSCA 算法首先根据 WSN 的能量消耗模型,计算出整个网络应该划分的最佳簇数,然后由基站选出每个簇中的簇头并令节点成簇.压缩感知的分簇算法的着重点在于在数据传输之前,节点首先运行压缩感知算法对传感数据进行压缩,从而大量减少网络中的数据流量,以达到减少节点能量消耗,延长网络生存时间的目的.

2 压缩感知理论

压缩感知(Compressed Sensing)^[9-10]是文献[5]中提出的一种新的信号获取与处理理论.压缩感知理论指出:如果原始信号是稀疏的(或者经过一个稀疏矩阵变换之后是稀疏的),则对信号的采样和压缩可以同时进行,使用一个和稀疏变换矩阵不相干的测量矩阵对稀疏信号进行投影,可以得到降维的测量值.接收端通过求解一个优化问题可以高概率重构出原始信号.压缩感知理论突破了传统的奈奎斯特采样定理对采样频率的限制,即在测量端可以使用低频率的采样压缩而获得完整的信息,信息的重构通过在接收端进行复杂的计算来实现.

压缩感知理论涉及信号的稀疏表示、测量矩阵的设计和重构算法三个方面.设有一个维度为 N 的离散信号 X ,把 X 看成 R^N 空间的一个 $N \times 1$ 维列向量,则可以用一组标准正交基的线性组合来表示 X ^[11],即

$$X = \sum_{i=1}^N \theta_i \Psi_i, \text{ 或 } X = \Psi \Theta \quad (1)$$

式中, $\Psi = [\Psi_1 \Psi_2 \cdots \Psi_N]$, Ψ_i 是 N 维的列向量, Θ 也是一个 N 维的列向量, Θ 是 X 的加权系数序列, Θ 中的元素 $\theta_i = \langle X, \Psi_i \rangle = \Psi_i^T X$.可见 Θ 是信号 X 的等价表示,如果 Θ 中只有 K 个元素非零 ($K \ll N$),则称信号 X 是 K 稀疏的^[9].

对信号 X 进行稀疏变换后得到稀疏系数向量 Θ ,再设计一个测量矩阵 Φ 对 Θ 进行投影就可以得到信号的压缩表示. Φ 是一个 $M \times N$ ($M \ll N$) 的矩阵, Φ 稀疏系数向量相乘后得到 M 个测量值.测量值的数量远远小于原始信号的元素数量,但是测量值中包含了原始信号的足够信息,只要测量矩阵设计得当,就可以保证以高概率重构出原始信号^[12].令测量值向量 $Y = (Y_1, Y_2, \cdots, Y_M)^T$,则测量过程可以表示为

$$Y = \Phi \Theta = \Phi \Psi^T X = A^{CS} X \quad (2)$$

这里 Φ 是非自适应的,即 Φ 不必随着信号 X 的变化而变化, A^{CS} 称为感知矩阵.利用测量矩阵得到的测量值 Y 不是传统信号获取中的点采样,而是信号的 K 线性泛函.

在已知测量矩阵 Φ 和测量值 Y 的情况下,通过式(2)求解原始信号 X 是一个欠定问题,因为方程组中未知数的个数远多于方程的个数,方程组有无穷多组解.然而如果原始信号经过稀疏变换后是 K 稀疏的,并且 Φ 和 Y 之间满足一定的条件,那么稀疏系数向量 Θ 可以通过求解最优 l_0 范数问题精确重构.

$$\hat{\Theta} = \operatorname{argmin} \|\Theta\|_0 \quad \text{s.t.} \quad \Phi \Theta = Y \quad (3)$$

式中, $\|g\|_0$ 表示向量的 l_0 范数.如果要精确重构 K 稀疏向量 Θ , Y 的维数 M 须满足条件 $M = KO(\log(N))$,并且测量矩阵 Φ 要满足限制等距条件 RIP^[12-13].由于求解 l_0 范数下的最小化是一个 NP 困难问题,现已证明可以利用 l_1 范数进行重构优化.

3 压缩感知的 WSN 分簇算法 CSCA

3.1 网络模型

对要研究的 WSN 的性质作某些假设.

(1) 网络中一共有 N 个节点,随机均匀分布在边长为 a 的正方形区域中,基站位于网络覆盖区域以外.所有节点能量有限且初始能量相等,基站能量

和计算能力不受限制。

(2) 节点部署后不再移动,每个节点知道自己的位置且有一个全网唯一的 ID。

(3) 基站可以和网络中的所有节点通信,网络内的节点可以根据到簇头或者基站的距离调整发射功率。

(4) 每隔一定时间,节点把监测数据传送给簇头,簇头收集簇内所有成员的数据并融合后发送给基站。

3.2 无线通信能量消耗模型

采用文献[14]中给出的无线通信能量消耗模型,两个欧式距离为 d 的节点传输 k 比特数据的能量消耗可用式(4)和式(5)表示。

$$E_{Tx}(k, d) = \begin{cases} kE_{elec} + k\epsilon_{fs}d^2 & d < d_0 \\ kE_{elec} + k\epsilon_{mp}d^4 & d \geq d_0 \end{cases} \quad (4)$$

$$E_{Rx}(k) = kE_{elec} \quad (5)$$

式中, E_{elec} 是节点电路处理 1 个比特数据所消耗的能量, ϵ_{fs} 和 ϵ_{mp} 是节点无线通信能量消耗系数, d_0 是节点通信距离阈值。

3.3 网络分簇

3.3.1 最优分簇数目

WSN 分簇时如果簇的数量过少,则簇内成员节点较多,成员离簇头的距离较远,成员向簇头传输数据时需要消耗更多的能量,这会缩短簇内成员节点的寿命。如果簇的数量过多,则有很多簇头向基站传输数据,簇头的能量消耗比例过大,这会缩短簇头的寿命。由文献[15]可知,对于一个 WSN,存在一个最优分簇数 H_{Best} ,可用式(6)表示。

$$H_{Best} = \sqrt{\frac{N\epsilon_{fs}}{2\pi\epsilon_{mp}}} \frac{a}{d_{H-B}^2} \quad (6)$$

式(6)中, N 是 WSN 的节点数量, a 是 WSN 覆盖区域的边长, ϵ_{fs} 和 ϵ_{mp} 就是式(4)中用到的能量消耗系数, d_{H-B}^2 是簇头和基站之间距离的平方。

3.3.2 分簇步骤

WSN 的分簇分为簇头选择和节点加入相应的簇两个阶段。CSCA 算法的簇头选择在能量和计算能力不受限制的基站进行。第一轮分簇开始前, WSN 的所有节点把自己的 ID、位置和能量信息报告给基站。根据式(6),基站计算出 WSN 的最优分簇数目 H_{Best} ,把 WSN 的覆盖区域划分成 H_{Best} 个小正方形,给每一个小正方形赋一个 ID,并计算出小正方形的中心位置。因为 WSN 的节点是随机均匀

分布的,因此每个小正方形中的节点数量近似相等。在每个簇(小正方形)中选择簇头时,按照以下步骤进行。

(1) 基站根据式(7)计算簇中每一个节点充当簇头时的适应值 f 。

$$f = \alpha \sum_{i=1}^{M-1} d_{H-i}^2 + \beta d_{H-B}^2 + \gamma \frac{\sum_{i=1}^M e_i}{e_H} \quad (\alpha + \beta + \gamma = 1) \quad (7)$$

式中, $\sum_{i=1}^{M-1} d_{H-i}^2$ 表示成员节点到簇头的距离平方之和,由式(4)可知,在一个簇内, $\sum_{i=1}^{M-1} d_{H-i}^2$ 的值越小,

则成员节点向簇头传送数据所消耗的能量就越小。 d_{H-B}^2 表示簇头到基站距离的平方, d_{H-B}^2 越小,则簇头向基站发送簇中融合数据的能量消耗就越小。

$\frac{\sum_{i=1}^M e_i}{e_H}$ 表示簇头能量占簇中所有节点能量百分比的

倒数, $\frac{\sum_{i=1}^M e_i}{e_H}$ 越小,则簇头的能量相对越大,选择能量

大的节点做簇头可以均衡节点的能量消耗。 α 、 β 、 γ 表示式(7)中三个因子对簇头选择适应值的贡献比例,分簇过程中可以动态调整。一般情况下可以取 $\alpha = 0.3$, $\beta = 0.3$, $\gamma = 0.4$ 。基站经过计算后,选择具有最小 f 值的节点作为所在簇的簇头,并向整个网络广播这条消息。

(2) 当选簇头接收到自己成为簇头的消息后,以一定的功率广播 ADV(Advertisement Message)消息,周围节点根据自己接收到的簇头广播消息的信号强弱,选择一个簇头加入其所在的簇,并把成为这个簇头成员的消息发送给簇头。所有的节点都加入到相应的簇之后,簇头给簇中的节点分配时隙表,并把簇中成员节点的 ID 报告给基站。

3.4 数据传输

分簇完成后,WSN 中的节点采集监测数据,并周期性的传送到簇头。CSCA 算法在数据传输阶段,采用压缩感知算法对监测到的数据进行采样压缩,降低网络中的数据流量,这是 CSCA 算法和传统分簇算法的最大不同之处。压缩感知的稀疏矩阵 Ψ 可以采用快速傅里叶变换基^[16],观测矩阵 Φ 可以采用高斯随机矩阵^[17-18],由式(2)知 $\Phi\Psi^T = A^{CS}$,基站计

算出感知矩阵 A^{CS} ,并向全网广播,WSN 的节点接收到 A^{CS} 之后存储在自身的存储器中,用来对监测数据进行投影。

设在一段时间 T 内簇内节点采集到的监测数据 X 可用一个 N 维列向量 $(X_1, X_2, \dots, X_N)^T$ 来表示,在把数据发送到簇头之前,节点利用感知矩阵 A^{CS} 和 X 相乘,得到 X 在感知矩阵上的投影即压缩测量数据 $Y = (Y_1, Y_2, \dots, Y_M)^T$,其中 $M \ll N$,节点把经过压缩后的测量值 Y 发送给簇头。

簇头在接收到各个成员节点的投影值之后,利用式(8)把自身监测值在感知矩阵 A^{CS} 上的投影值和各

$$\Delta Y_i = Y_H - Y_i \quad (8)$$

个成员节点的投影值进行一次做差的运算。式(8)中 Y_H 是簇头的投影值, Y_i 是簇内节点的投影值。因为一个簇内的成员节点在空间上距离较近,监测到的数据有较强的空间相关性,所以 ΔY_i 的值很多近似为 0,簇头做差得到的 ΔY 是一个稀疏信号,可以再一次应用压缩感知进行投影以减少簇头向基站传输的数据量。簇头利用一个高斯随机矩阵 Φ_G 和各个节点的投影差值组成的列向量 $(\Delta Y_1, \Delta Y_2, \dots, \Delta Y_J)^T$ (J 是簇中成员节点的数目)相乘得到簇头的投影值 Z ,然后簇头把 Z 和自身的投影值 Y_H 传输给基站。基站得到各个簇头的投影值以及簇头和节点的投影差值 ΔY 之后,可以根据需要进一步向更上层控制中心传输或者在基站运行压缩感知的重构算法恢复 WSN 中节点监测到的数据。簇头使用的高斯随机矩阵 Φ_G 可以在每轮分簇结束后由基站下发得到。

4 仿真结果及分析

为了验证 CSCA 算法的性能,应用 Matlab 软件对 CSCA、LEACH 以及文献[8]中的蚁群分簇算法进行仿真对比,评价指标有两个:一是网络中节点的平均能量消耗,二是整个网络的生存时间(以网络中 70% 的节点存活作为判断依据)。仿真环境的设置为:WSN 中一共有 200 个节点,随机均匀分布在 $100 \text{ m} \times 100 \text{ m}$ 的正方形区域中,参照文献[14]取 WSN 的通信参数分别为 $E_{\text{elec}} = 50 \text{ nJ/bit}$, $E_{\text{fs}} = 10 \text{ pJ/bit/m}^{-2}$, $E_{\text{mp}} = 0.0013 \text{ pJ/bit/m}^{-4}$, $d_0 = 87 \text{ m}$,每个节点的初始能量均为 0.5 J 。每次数据传输前节点监测到的原始数据大小为 1 kB 。为了能客观的评

价各种算法的性能,每种算法都运行 100 次,把运行结果的平均值作为最终的仿真结果。

图 1 所示为 CSCA、LEACH、蚁群三种分簇算法的总能量消耗仿真结果,横坐标是分簇轮数,纵坐标是网络总的能量消耗。从图 1 可以看出,随着分簇轮数的增加,三种算法的能量消耗都在增加,但是在相同的分簇轮数时 CSCA 算法消耗的能量较其他两种算法少得多,这是由于 CSCA 算法在数据传输阶段采用压缩感知把传输的数据量降低到了原始数据量的 23% 所致。

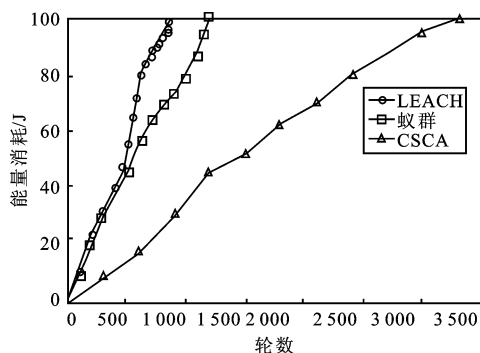


图 1 三种分簇算法的能量消耗仿真结果

图 2 所示为 CSCA、LEACH、蚁群三种分簇算法的网络生存时间仿真结果,横坐标是分簇轮数,纵坐标是网络中存活的节点数量,节点的死亡以能量耗尽作为判断依据。从图 2 可以看出,随着分簇轮数的不断增加,网络中的存活节点在不断减少。LEACH 算法分簇时网络的存活时间大约是 590 轮,蚁群算法分簇时网络的存活时间大约是 1 150 轮,而 CSCA 算法分簇时网络的存活时间大约是 3 400 轮。由图 2 可见,CSCA 算法较 LEACH 算法把 WSN 的生存期延长了 576%,较蚁群分簇算法把网络生存时间延长了 295%。

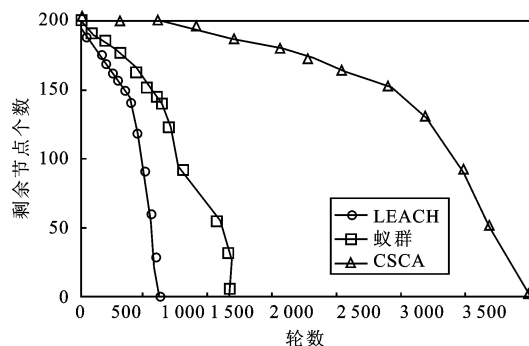


图 2 三种分簇算法的网络生存期仿真结果

5 结束语

压缩感知作为新的数据处理方法,大大减少了采样过程产生的数据量,并且压缩感知在前端数据采样压缩时计算简单,而在后端的数据重构计算复杂,这正好符合 WSN 节点计算能力有限的特点,因此,可以在 WSN 的分簇算法中使用压缩感知技术,把节点监测到的数据进行压缩后再在网络中传输。

压缩感知的分簇算法 CSCA 不同于传统分簇算法在数据传输之前首先应用压缩感知对原始数据进行压缩以减少网络中传输的数据量。仿真结果表明,压缩感知的 WSN 分簇算法 CSCA 较传统的分簇算法大幅度降低了 WSN 节点的能量消耗,延长了网络的生存时间。

参考文献:

- [1] Niculescu D, Amerig N L. Communication paradigms for sensor network[J]. IEEE Communications Maganize, 2005, 43(3): 116-122.
- [2] Heinzelman W R, Kul IK J, Bala Krishnan H. Adaptive protocols for information dissemination in wireless sensor networks[C]// Proc. of the 5th Ann. Int'l Conf. on mobile computing and networking. Italy, Rome, 2001: 174-185.
- [3] Lindsey S, Rag havendra C S. PEGASI S: power efficient gather ing in sensor information systems[C]// Proceedings of the IEEE Aero space Conference. USA, Montana, 2002: 1-6.
- [4] Arati Manjeshwar, Dharma Agrawal. TEEN: a routing protocol for enhanced efficiency in wireless sensor network[C]// Proceedings of the 15th Parrallel and Distributed Processing Symposium. California, San Francisco, 2001: 2009-2015.
- [5] 梁英,于海斌,曾鹏. 应用 PSO 优化基于分簇的无线传感器网络路由协议[J]. 控制与决策, 2006, 21(4): 453-456.
- [6] 蒋畅江,唐贤轮,向敏. 基于 PSO 的无线传感器网络非均匀分簇路由协议[J]. 计算机应用研究, 2012, 29(8): 3074-3077.
- [7] 周玉,景博,杨洲. 一种基于遗传算法的无线传感器网络 LEACH 协议改进算法[J]. 计算机研究与发展, 2010, 47(Supl): 175-179.
- [8] 苏兵,黄娟. 一种基于蚁群算法的 WSN 能量均衡路由[J]. 微电子学与计算机, 2012, 29(11): 50-52.
- [9] Donoho D L. Compressed sensing[J]. IEEE Transactions on Information Theory, 2006, 52(4): 1289-1306.
- [10] Candès E. Compressive sampling[C]// Proceedings of International Congress of Mathematicians. Madrid, Spain: European Mathematical Society Publishing House, 2006: 1433-1452.
- [11] Baraniuk R. A lecture on compressive sensing[J]. IEEE Signal Processing Magazine, 2007, 24(4): 118-121.
- [12] Candès E, Romberg J, Tao T. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information[J]. IEEE Transactions on Information Theory, 2006, 52(2): 489-509.
- [13] Candès E. The restricted isometry property and its implications for compressed sensing[J]. Académie des sciences, 2006, 346(I): 598-592.
- [14] Heinzelman W, Chandrakasan A, Balakrishnan H. Energy-efficient communication protocol for wireless microsensor networks[C]// Proc of the 33rd International Conference on System Science. Hawaii, 2000: 535-540.
- [15] Hesham A, Yang S H. Dynamic cluster head for life-time efficiency in WSN[J]. International Journal of Automation and Computing, 2009, 6(1): 48-54.
- [16] Mallat S. A wavelet tour of signal processing[M]. San Diego: Academic Press, 1996.
- [17] Candès E, Tao T. Decoding by linear programming [J]. IEEE Transactions on Information Theory, 2005, 51(12): 4203-4215.
- [18] Candès E, Romberg J, Tao T. Stable signal recovery from incomplete and inaccurate measurements [J]. Communications on Pure and Applied Mathematics, 2006, 59(8): 1207-1223.

作者简介:

曹欲晓 男, (1971-), 硕士, 讲师. 研究方向为智能算法、无线传感器网络. E-mail: caoyx@njt.edu.cn.

钱 瑛 女, (1982-), 博士研究生, 讲师. 研究方向为图像处理、模式识别.

韩 磊 男, (1982-), 博士研究生, 讲师. 研究方向为图像处理、无线传感器网络.

俞 翔 男, (1982-), 博士, 讲师. 研究方向为数字信号处理、嵌入式系统.