

基于压缩感知的无线传感器网络数据融合算法^①

史久根, 张加广

(合肥工业大学 计算机与信息学院, 合肥 230009)

摘要: 无线传感器网络中存在大量的数据冗余, 数据融合技术通过对采样数据进行压缩, 消除冗余, 有效的减少了节点发送的数据量, 延长传感器网络的寿命. 提出了压缩感知与数据转发相结合的数据融合算法, 在网络采样数据收集的过程中根据节点的子节点个数选择利用压缩感知对数据进行压缩还是直接对数据进行数据转发. 仿真结果表明, 和基于压缩感知的数据融合算法相比, 数据转发与压缩感知相结合的数据融合算法, 有效地在平衡节点间负载的同时减少节点的发送量.

关键词: 无线传感器网络; 压缩感知; 数据融合; 网络数据收集协议; 网络负载均衡

Data Fusion Based on Compressed Sensing in Wireless Sensor Networks

SHI Jiu-Gen, ZHANG Jia-Guang

(School of Computer and Information, Hefei University of Technology, Hefei 230009, China)

Abstract: There are a lot of data redundancy in wireless sensor networks. By compressing the original sampling data, the data fusion technology eliminates redundancies in data, reduces the amount of data sent by nodes effectively and prolongs lifetime of sensor networks. This paper proposed a data fusion algorithm that combined data forwarding and compressed sensing. During the process of collecting sampling data in sensor networks, the algorithm selects using compressed sensing to compress original sampling data or simply storing and forwarding sampling data according to the amount of nodes' child nodes. Simulations indicate that compared with the data fusion algorithm based on compressed sensing, the data fusion algorithm that combined data forwarding and compressed sensing achieved both network load balance and data compression effectively.

Key words: Wireless Sensor Networks(WSNs); compressed sensing; data fusion; network data collection protocol; network load balance

随着传感器节点成本以及功耗不断降低, 无线传感器网络在军事、环境监测等领域的应用日益广泛^[1]. 但由于无线传感器网络中资源受限、能量受限的本质没有发生改变, 因此, 如何在无线传感器网络中减少节点能量消耗以及网络能量均衡仍是在设计网络数据收集协议中很重要的考量因素.

无线传感器网络是指将大量传感器节点密集地部署到监控区域进行数据采集. 节点的密集分布会导致相邻节点采集的原始数据之间存在大量的空间相关性, 使得大量冗余数据的产生, 造成了节点能量的浪

费, 减少了网络寿命. 传感器节点能量消耗主要体现在数据的发送以及数据处理方面. 文献[2]指出, 传感器节点在处理数据进行指令运算所消耗的能量要远远小于消耗在无线发送数据所消耗的能量上, 因此数据融合技术通过牺牲节点存储资源、计算资源为代价来实现数据的压缩. 数据融合技术作为无线传感器网络研究的热点, 主要作用是将传感器节点采集到的大量原始数据在数据收集过程中进行网内处理, 消除冗余^[3].

本文针对无线传感器网络中存在的大量的数据冗余, 提出压缩感知与数据转发相结合的数据融合算法,

^① 基金项目:国家重大仪器设备开发专项(2013YQ030595)

收稿时间:2014-02-22;收到修改稿时间:2014-03-28

减少节点的数据发送量。

1 压缩感知介绍

对于长度为 N 的向量 $x=(x_1, x_2, \dots, x_N)^T$, 可逆正交变换矩阵 $\psi_{N \times N}=[\varphi_1, \dots, \varphi_N]$, 若信号 x 可以由 ψ 中 k 列向量线性组合表示, 即:

$$x = \psi\theta = \sum_{n=1}^N \theta_n \varphi_n = \sum_{l=1}^k \theta_{n_l} \varphi_{n_l} \quad (1)$$

则称 x 在基 ψ 下为 k 稀疏信号^[4], 称 $\psi_{N \times N}$ 为稀疏变换基。

压缩感知技术指出若信号 x 在基 ψ 下为 k 稀疏信号, 则可以通过与稀疏变换基 $\psi_{N \times N}$ 不相关的测量矩阵 $\Phi_{M \times N}$ 将高维信号 x 投影到低维空间 y_M 上, 并保证数据的高概率重构^[4]. 即: $y_M = \Phi_{M \times N} x_N$, 将公式(1)代入得:

$$y_M = \Phi_{M \times N} x_{N \times 1} = \Phi_{M \times N} \psi_{N \times N} \theta_{N \times 1} = A_{M \times N} \theta_{N \times 1} \quad (2)$$

$\Phi_{M \times N}$ 中 M 为测量矩阵的测量次数, 其值与稀疏度 k 有关, $M = ck \lg(N/k)$, 其中 c 为常数^[4].

测量矩阵和稀疏变换基的不相关性保证了可以通过求解公式(3)最优化问题来高概率重构出原始信号 x ^[5]. 通常选择测量矩阵中的每个元素均为随机变量, 来保证测量矩阵和稀疏变换基的不相关性. 常用的测量矩阵有高斯随机矩阵、二值随机矩阵、局部傅里叶矩阵等^[6]. 本文采用二值随机矩阵来进行测量, 这是由于每个元素均为 0 或 1, 可用一位存储, 节省资源。

$$\min \|\theta\|_1 \quad st. \quad y_M = A_{M \times N} \theta_N \quad (3)$$

常用的信号重构算法有基追踪算法、匹配追踪算法、正交匹配追踪算法等^[7].

2 基于压缩感知的数据融合算法设计

无线传感器网络中, 定义节点 i 采集到的信号为 x_i , 假设单次数据采集得到的数据大小为 $Size$, 单位比特. 数据收集的工作是将每个节点采集到的数据发送到汇聚节点同时减少能量的消耗以及保证负载的均衡. 传统的数据转发算法并不对采样数据进行数据融合, 原理是每个节点均发送采集到的数据并转发接收到的数据到上一跳节点. 记节点 i 的子树节点数为 m_i , 则根据数据转发算法节点 i 发送的数据为 $packet_i = (m_i + 1) * Size$. 数据转发算法虽然简单

于实现, 但节点 i 发送的数据量 $packet_i$ 与 m_i 成正比, 会造成靠近汇聚节点的传感器节点负载较大, 网络拥塞等。

为了优化数据收集方案, 文献[8]提出了基于压缩感知的数据融合算法. 定义向量 $x = (x_1, x_2, \dots, x_N)^T$. 根据压缩感知理论假设 x 为 k 稀疏信号, 则测量次数 $M = ck \lg(N/k)$. 得到的压缩结果为:

$$y = \Phi_{M \times N} x = (\varphi_1, \varphi_2, \dots, \varphi_N)(x_1, x_2, \dots, x_N)^T = \sum_{i=1}^N \varphi_i x_i = \sum_{i=1}^N \delta_i$$

其中 φ_i 为 $M \times 1$ 的列向量, 称 φ_i 为随机向量, δ_i 为测量向量。

基于压缩感知的数据融合算法要求每个节点 i 在部署到监测区域之前, 存储自身的节点号和对应的随机数向量 φ_i , 向量 φ_i 大小为 M 比特. Sink 节点记录所有节点 ID 号和 φ_i 的对应关系. 数据收集阶段, 每个节点 i 在发送数据之前首先计算测量向量 $\delta_i = \varphi_i x_i$, 并等待其子节点发送来的数据包, 最后将这些数据包中的测量向量 δ 和 δ_i 直接相加, 得到 $Send_i = \sum_{i=1}^{m_i} \varphi_i x_i + \varphi_i x_i$,

封装成以下数据包格式后发送给上一跳节点。

表 1 数据包封装格式

节点	节点	节点	节点	$Send_i$
l_1	l_2		l_m	l_i	

$Send_i$ 为各子节点测量向量之和, 其长度和单个子节点测量向量长度相同. 由于这些数据包相加得到的结果长度不变(忽略节点号所占空间不计), 因此传感器网络中每个节点发送的数据量均为 $packet_i = M * Size$. 最终汇聚节点收到所有节点的向量积之和。

$$y_M = \sum_{i=1}^N \varphi_i x_i$$

通过汇聚节点将测量值传输到 Sink 节点后, Sink 节点通过运行压缩感知重构算法来恢复原始信号 x_N . 和传统的数据转发比较, 对于网络中子节点数 $m_i > M - 1$ 的传感器节点, 该算法能有效的减少数据传输量, 但对于子节点数 $m_i < M - 1$ 的节点, 增加了数据量的传输. 这是由于算法中对每一个传感器节点均采用对采样数据求测量向量后发送的方式, 对于网络中子节点数 $m_i = 0$ 的叶子节点, 需要发送的信息大小为

Size 比特, 然而经过压缩采样处理后发送的数据量增加到 $M * Size$ 比特, 增加了节点的负荷. 为了解决上述问题, 本文提出压缩感知和数据转发相结合的数据融合算法, 该算法同样要求节点在部署之前, 存储自身的节点号和对应的随机数向量 φ_i , Sink 节点记录所有节点 ID 号和 φ_i 的对应关系. 该算法可以实现对于子节点数 $m_i < M-1$ 的节点使用数据转发算法, 而 $m_i \geq M-1$ 的节点使用压缩感知算法进行数据融合. 具体实现包含两个阶段: 网络配置阶段和数据收集阶段.

网络配置阶段发生在路由树构造完毕之后, 路由树构造过程中各节点统计自身子节点数 m_i , 当路由树构造完毕后, 执行以下操作:

- ①若 $m_i < M-1$, 将自身节点号 i 以及对应 φ_i 封装成融合参数消息, 发送给节点的上一跳. 同时转发来自其子节点的融合参数消息至上一跳节点.
- ②若 $m_i \geq M-1$, 接收来自子节点的融合参数消息, 并存储融合参数消息中包含的子节点 ID 号与 φ_i 的对应关系到子节点配置列表中, 不再向上一跳节点转发融合参数消息.

经过网络配置阶段后, $m_i \geq M-1$ 的节点中的子节点配置列表将包含所有进行压缩感知时所必须的子节点号对应的随机向量 φ .

在数据采集阶段, 当采集到数据 x_i 后, 按照以下算法将数据发送到上一跳节点.

(1)若 $m_i < M-1$, 将节点号 i 以及采集到的数据 x_i 封装成数据包发送给上一跳节点, 并转发来自其子节点的数据包.

(2)若 $m_i = M-1$, 则读取收到的 $M-1$ 个数据包, 对收到的每一个数据包执行以下操作:

- ①根据子节点配置列表, 查找数据包中节点 n 对应的随机向量 φ_n .
- ②将数据包中采集到的数据 x_n 乘以 φ_n 得到测量向量 $\delta_n = \varphi_n x_n$.

计算节点本身的 δ_i 并对得到的所有测量向量 δ 求和, 记为

$$Send_i = \sum_{n=1}^{M-1} \delta_n + \delta_i = \sum_{n=1}^{M-1} \varphi_n x_n + \varphi_i x_i$$

$Send_i$ 为要发送的数据. 将 $Send_i$ 封装成以下格式的数据包发送到上一跳节点.

表 2 数据包封装格式

节点 n_1	节点 n_2	节点 n_M	$Send_i$
----------	----------	-------	----------	----------

(3)若 $m_i > M-1$, 假设节点 i 收到 r 个仅包含单个节点的数据包, s 个包含大于等于 M 个节点的数据包, 则

$$Send_i = \sum_{t=1}^r \varphi_{n_t} x_{n_t} + \sum_{l=1}^s Send_l + \varphi_i x_i$$

最终在汇聚节点收到所有节点的测量向量 δ 之和, 通过在 Sink 节点进行数据重构恢复采集到的数据.

该协议在网络配置阶段每个节点发送的数据为

$$Packet_i^1 = \begin{cases} (m_i + 1) * M, & m_i < M - 1 \\ 0, & m_i \geq M - 1 \end{cases}$$

在数据收集节点每个采集周期发送的数据为

$$Packet_i^2 = \begin{cases} (m_i + 1) * Size, & m_i < M - 1 \\ M * Size, & m_i \geq M - 1 \end{cases}$$

相比于基于压缩感知的数据融合算法, 对于 $m_i > M-1$ 的节点两种数据收集方案没有任何区别. 但对于 $m_i < M-1$ 的节点, 虽然压缩感知与数据转发相结合的融合算法在网络配置阶段多发送了 $NEG_i = (m_i + 1) * M, m_i < M - 1$ 的网络配置参数, 但在数据收集阶段, 节点在每一次数据采集减少发送了 $POS_i = M * Size - (m_i + 1) * Size = (M - m_i - 1) * Size$ 的数据量. 随着数据收集次数的增加, 该方案将优于基于压缩感知的数据融合方案. 在数据收集阶段, 对于 $m_i < M-1$ 的节点, m_i 越小, 发送的数据量越少, 趋近于数据转发方案, 而对于 $m_i > M-1$ 的节点, 发送的数据量统一为 $M * Size$, 和基于压缩感知的数据收集方案相同.

3 仿真实验及分析

为了验证本文提出的网络数据融合算法对网络中各节点发送数据量的影响, 本文以 Matlab 2010b 为工具进行了仿真实验. 以无线传感器网络监控周围温度信号为仿真对象, 节点采集到的温度值用 2 字节空间存储 $Size = 16$ bit, 监控区域中布置了 100 个传感器节点. 每次发送的数据为 $x = (x_1, x_2, \dots, x_{100})$, 通过压缩感知进行线性测量后得到测量值 $y = (y_1, y_2, \dots, y_M)$, 其中 M 的值与信号稀疏度有关, 信号相关性越大, M 的值越小, 数据融合的效率越高. 图 1 和图 2 分别展示了在 $M = 50$ 的情况下经过 Count 次数据采集后, 节点具有的子节点数与节点所发送的平均数据量大小的关系. 从图 1 可知, 压缩感知与数据转发相结合的数据融合算法在子节点数小于 $M-1$ 时, 平均发送的数据量

与子节点数成正比。但斜率要比基于数据转发的网络收集协议大。而对于子节点数大于等于 $M-1$ 时, 平均发送的数据量与基于压缩感知的数据融合算法相同, 不再随着子节点数线性增长。因此, 基于压缩感知和数据转发相结合的数据融合算法综合了两种算法的优点, 有效的减少了网络中存在的数据冗余量。

从图 1 可以看到, 对于采集次数 $Count = 10$ 时, 部分节点平均发送的数据量要大于基于压缩感知数据融合算法发送的平均数据量, 而从图 2 中可以看到, 当采集次数 $Count = 200$ 时, 传感器网络中所有节点的平均数据发送量均低于或等于基于压缩感知的数据融合算法。这是由于在网络配置阶段 $m_i < M-1$ 的节点发送了大量的融合参数数据, 且 m_i 越大, 发送的数据越多。对于网络配置阶段发送数据量最大的 $m_i = M-2$ 的节点 i , $PCS_i = Size$, $NEG_i = (M-1)*M$, 因此最少经过 $Min = \frac{(M-1)*M}{Size}$ 次数据采集周期后, 整个无线传感器网络的所有节点的数据传输量将小于或等于基于压缩感知的数据采集方案。因此如果需要更换路由树, 要确保数据采集超过了 Min 次。

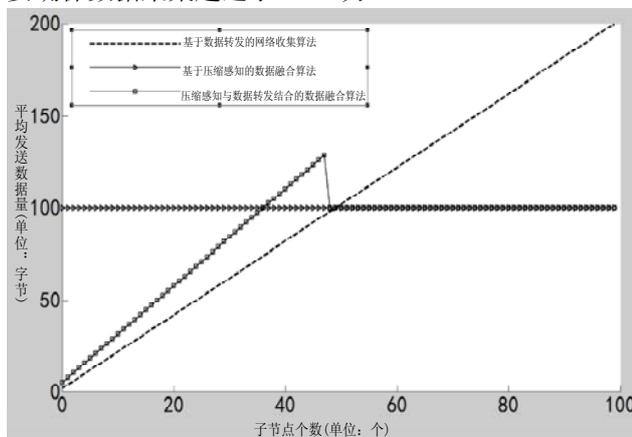


图 1 Count=10 下子节点数与平均发送数据量的关系

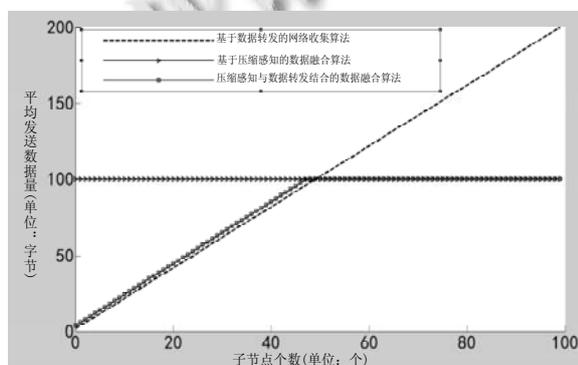


图 2 Count=200 下子节点数与平均发送数据量的关系

从图 2 可以看到, 压缩感知与数据转发相结合的数据融合算法要优于基于压缩感知的数据融合算法, 但数据融合算法是建立在牺牲网络时延、节点存储资源以及计算资源的基础上的。

在网络时延方面, 基于数据转发的网络数据收集算法由于没有数据融合, 因此节点收到子节点发送来的数据直接转发给上一跳节点, 不存在网络时延。基于压缩感知的数据融合算法由于需要将所有子节点发送回来的测量向量求和得到结果后才能发送上一跳节点, 因此该数据融合算法要求节点收到所有节点发送来的数据之后才能相上一跳发送数据, 增加了网络时延。基于压缩感知和数据转发相结合的数据融合算法, 对于 $m_i < M-1$ 的节点采用数据转发协议不进行数据融合, 因此不存在网络延迟, 而对于 $m_i \geq M-1$ 的节点需要采用压缩感知进行数据融合, 因此同样需要等待成功接收所有子节点的数据, 存在网络时延。

在消耗传感器节点存储资源以及计算资源方面, 基于数据转发的网络收集算法由于每个节点只进行存储转发操作, 因此需要消耗的计算量以及存储量比较少。基于压缩感知的数据融合算法要求节点存储自身 ID 对应的随机向量, 大小为 M bit, 同时需要计算所有测量向量之和。压缩感知与数据转发相结合的数据融合算法不仅要求存储自身 ID 对应的随机向量而且要求存储部分子节点 ID 对应的随机向量, 通过算法中网络配置阶段的协议可以看出节点最多存储 M 个这样的对应记录, 因此该算法要求节点的存储能力要达到 $M * M$ bit (未考虑节点 ID 号所占字节)。在计算资源消耗方面和基于压缩感知数据融合算法相同, 需要计算所有子节点发送过来的测量向量之和。

通过以上分析可以看到, 基于压缩感知与数据转发相结合的数据融合算法由于对传感器采集节点的硬件存储空间大小有要求, 同时存在一定的网络时延, 因此该算法主要适用于采样节点具有一定的硬件存储空间且对网络延迟具有一定的容忍性的传感器网络中。从图 1 和图 2 的比较可以看到, 在确定路由树后, 应进行至少 Min 次数据采集才能保证该算法优于基于压缩感知的数据收集算法, 因此该算法不适用于路由树经常发生变化的传感器网络中。

4 结语

数据融合是无线传感器网络规模不断扩大过程中

需要解决的关键技术,数据融合通常是以牺牲节点硬件资源和计算资源的代价下减少数据的发送量,主要解决在数据收集过程中如何融合子节点中发送来的数据,消除数据冗余,从而减少数据的传输量.无线传感器网络中引入压缩感知进行数据融合能够有效地提高网络寿命,具有巨大的发展前景.本文提出了数据转发与压缩感知相结合的数据融合算法,集成了数据转发和压缩感知各自的优点,有效地在减少数据传输量的同时,平衡节点之间的网络负载.

参考文献

- 1 任丰原,黄海宁,林闯.无线传感器网络.软件学报,2003,14(7):1282-1291.
- 2 Boulis A, Ganeriwal S, Srivastava MB. Aggregation in sensor networks: an energy-accuracy trade-off. Ad Hoc Networks, 2003, 1(2): 317-331.
- 3 康健,左宪章,唐力伟,等.无线传感器网络数据融合技术.计算机科学,2010,37(4):31-35.
- 4 Baraniuk RG. Compressive sensing. Signal Processing Magazine, 2007, 24(4): 118-121.
- 5 Donoho DL, Elad M, Temlyakov VN. Stable recovery of sparse overcomplete representations in the presence of noise. IEEE Trans. on Information Theory, 2006, 52(1): 6-18.
- 6 Candes EJ, Tao T. Near-optimal signal recovery from random projections. Universal encoding strategies. IEEE Trans. on Information Theory, 2006, 52(12): 5406-5425.
- 7 Haupt J, Bajwa WU, Rabbat M, et al. Compressed sensing for networked data. Signal Processing Magazine, IEEE, 2008, 25(2): 92-101.
- 8 Xie D, Zhou Q, Liu J, et al. A Chain-Based Data Gathering Protocol under Compressive Sensing Framework for Wireless Sensor Networks. Computational and Information Sciences (ICCIS), 2013 Fifth International Conference on. IEEE. 2013. 1479-1482.
- 9 Caione C, Brunelli D, Benini L. Distributed compressive sampling for lifetime optimization in dense wireless sensor networks. IEEE Trans. on Industrial Informatics, 2012, 8(1): 30-40.
- 10 Donoho DL. Compressed sensing. IEEE Trans. on Information Theory, 2006, 52(4): 1289-1306.
- 11 Chen F, Chandrakasan AP, Stojanovic VM. Design and analysis of a hardware-efficient compressed sensing architecture for data compression in wireless sensors. IEEE Journal of Solid-State Circuits, 2012, 47(3): 744-756.
- 12 李树涛,魏丹.压缩传感综述.自动化学报,2009,35(11):1369-1377.
- 13 张明,朱俊平,蔡骋.WSN中基于压缩感知的数据收集案.计算机工程,2012,38(20):68-71.
- 14 Plan Y, Vershynin R. One-bit compressed sensing by linear programming. Communications on Pure and Applied Mathematics, 2013, 66(8): 1275-1297.