



WSN 中基于多分辨率和压缩感知的数据融合方案^{*}

赵建军¹, 王怀宇¹, 赵泽阳¹, 陈生昌²

(1. 保定学院信息技术系 保定 071000; 2. 浙江大学理学院 杭州 310058)

摘要: 当前基于压缩感知的传感器网络数据融合方案中, 不论数据字段有何特征, 均假设网络具有固定而均匀的压缩阈值, 从而导致数据通信量过高, 能耗浪费较大。提出一种基于多分辨率和压缩感知的数据融合方案。首先, 对传感器网络进行配置, 以生成多个层次类型不同的簇结构, 用于过渡式数据收集, 在该结构上, 最低层的叶节点只传输原始数据, 其他层的数据收集簇进行压缩采样; 然后将其测量值向上发送, 当母数据收集簇收到测量值时, 利用基于反向 DCT 和 DCT 模型的 CoSaMP 算法恢复原始数据; 最后, 在 SIDnet-SWANS 平台上部署了该方案, 并在不同的二维随机部署传感器网络规模下进行了测试。实验结果表明, 随着分层位置的变化, 大部分节点的能耗均显著降低, 与 NCS 方案相比, 能耗下降 50%~77%, 与 HCS 方案相比, 能耗下降 37%~70%。

关键词: 无线传感器网络; 数据融合; 多分辨率; 压缩感知; 簇; 能耗

doi: 10.3969/j.issn.1000-0801.2014.09.013

Data Aggregation Scheme Based on Multi-Resolution and Compressive Sensing in Wireless Sensor Network

Zhao Jianjun¹, Wang Huaiyu¹, Zhao Zeyang¹, Chen Shengchang²

(1. Department of Information Technology, Baoding College, Baoding 071000, China;

2. The Science College, Zhejiang University, Hangzhou 310058, China)

Abstract: A data aggregation scheme based on multi-resolution with compressed sensing was proposed. Firstly, the network was configured to achieve the multiple-level and the different types of cluster structure for intermediate data collection, on this structure, the leaf nodes in the lowest level only transmit the raw data. The collecting clusters in other levels perform the compressed sampling and then transmit them to their parent cluster heads. When parent collecting clusters receive random measurements, they use inverse DCT and DCT model based CoSaMP algorithm to recover the original data. The proposed scheme was implemented on a SIDnet-SWANS simulation platform and test different sizes of two-dimensional randomly deployed sensor network. The experiment results show that the substantial energy savings are reported for a large portion of sensors on the different hierarchical positions, ranging from 50% to 77% when compared with NCS, and from 37% to 70% when compared with HCS.

Key words: wireless sensor network, data aggregation, multi-resolution, compressive sensing, cluster, energy consumption

1 引言

数据融合^[1]是无线传感器网络(wireless sensor network,

WSN)中一项非常重要的技术,也是目前的研究热点。该技术可以通过一定的算法对传感器节点采集到的大量原始数据进行各种网内处理,去除其中的冗余信息,只将少量

^{*} 国家自然科学基金面上项目(No.41374001)

有意义的处理结果传输给汇聚节点(sink节点)。采用数据融合技术能够大大减少WSN中需要传输的数据量,降低数据冲突,减轻网络拥塞,从而有效地节省能源开销,延长网络寿命^[2]。由于受到物理环境和感知器节点自身能量或容量的限制,现有数据融合方案的能耗较大,在节点发生故障时,数据融合的结果并不准确。本文主要针对无线传感器网络中的数据融合方案进行研究,方案的成功实现对于保证无线传感器网络的可靠性具有重要的理论意义和应用价值。

2 相关工作

数据融合问题一直是WSN中的一大研究热点。有众多学者相继提出了一系列方法用于无线传感器网络中的数据融合^[3-8]。

杨庚等人^[3]在数据融合的同时加入多类优化因子,提出了一种面向隐私保护的高精确度数据融合算法,仿真结果表明,该算法在保证数据融合精确度和安全性的前提下,有效地降低了通信量,延长了网络存活时间。

乐俊等^[4]提出了一种基于非均匀划分的分簇数据融合算法,将网络划分为非均匀的栅格,距离基站越远,栅格则越大,每轮从每个栅格中选择一个剩余能量最大的节点担任簇首,节点根据簇首广播消息的信号衰减强度选择簇首。仿真实验结果表明,该算法能够节约能量和均衡能耗,显著延长网络的生命周期。

余阳等^[5]提出一种小波支持向量机(SVM)和证据理论相结合的层级式数据融合算法,该算法利用小波变换方法对原始信号进行数据级消噪处理,建立SVM多分类器模型,以特征参数作为输入向量在普通节点处进行泄漏检测特征级融合。采用改进的证据组合规则,在汇聚节点处进行决策级证据组合,得到管网状态的最终决策。仿真结果表明,该方法可有效提高泄漏源位置检测的正确率,降低检测错误。

林蔚等^[6]针对含有缺失数据的无线传感器网络数据融合问题,提出了一种简单易行的二次数据融合算法(TFA),考虑到感知数据的时空相关性,对缺失数据进行线性插值插补和回归分析插补,综合考虑各节点的信息,利用自适应加权融合算法进行融合。仿真结果表明,该算法在含有缺失数据的前提下,能够以较低的计算开销和较高的估计精度实现数据融合。

马守明等^[7]提出了一种基于传感器节点信誉度集对安全数据进行融合的方法。在节点分簇阶段,利用基于密度函数的减法聚类方法进行分簇,确保分簇规模更加符合节点的实际布设情况;在数据传输阶段,将簇头节点选择下一跳数据转发节点建模为多属性决策过程,综合考虑备选转发节点的信誉度、能量等属性信息,从中选择综合评价最优的簇头节点转发融合数据,不仅使网络中的数据流量分布更加均衡,而且保证了数据的安全性。仿真结果表明,该方法在融合精度、安全性及簇头节点能耗等方面与同类方法相比具有明显的优势。

李小亚等^[8]提出了一种能量平衡的数据汇聚路由算法——EBGP。该协议依据邻居节点的剩余能量、邻居节点到汇聚节点的距离以及邻居节点到源节点的距离,计算概率值,然后选择邻居节点中概率最大的一个作为数据转发节点。仿真结果显示,EBGP能够实现能量平衡和较高的能量效率。

参考文献[9]提出了一种基于压缩感知的数据收集(NCS)方法,该方法可以将传输需求量降低至 $O(N \times M)$,其中 M 表示随机测量的数量,且 $M \ll N$ 。然而,当应用于大规模网络时, M 的值可能很大。此外,在数据融合初始阶段,叶子节点毫无必要地传输了 M 个测量值,超过了采集数据,进而在数据融合时引入了冗余。鉴于此,参考文献[10,11]提出了一种混合压缩感知融合(HCS)方法,设置一个阈值 M ,当传感器采集的数据大于或等于 M 时进行压缩感知融合,以对数据融合成本进行优化,从而降低了数据传输成本和能耗。然而只有少部分感知器使用了压缩感知融合方法,且使用压缩感知融合方法的节点的传输测量值 $M = K \log N$ 仍然过大。本文在现有研究工作的基础上,基于压缩感知理论提出了一种改进的数据融合方案,并通过仿真实验验证了本文方案的有效性。

3 数据融合模型

3.1 相关定义

本文方案的主要思路是:所有传感器节点不必再将数据发往单个汇聚节点。相反,将多个数据收集簇进行串联,形成不同层次不同类型的收集簇。数据从源节点出发,沿着簇结构到达汇聚节点。假设有 N 个传感器随机均匀部署于面积为 S 的二维正方形空间。设最低层次的单位面积为 S ,簇具有多种分辨率,且最高级为 T 。在每一级 i ,本文做如表1所示的定义。



表 1 相关符号含义

符号	含义
$s_i^{(l)}$	第 l 个簇的面积
$c_i^{(l)}$	网络中第 l 个簇的簇头
$N_i^{(l)}$	一个簇中的传感器节点数目
$M_i^{(l)}$	第 l 个簇测量值的传输次数
$d_i^{(l)}$	簇头 $c_i^{(l)}$ 和其子节点间的距离和
$\gamma_i^{(l)}$	传输数据规模和接收数据规模之比
$E_i^{(l)}$	第 l 个簇的传输能量成本
C_i	簇头集合
$ C_i $	簇头数量
M_i	网络中需要传输的测量值之和
E_i	所有簇头的传输能量成本

如表 1 所示, C_i 定义为所有簇头的集合, 即 $C_i = \{c_i^{(1)}, c_i^{(2)},$

$\dots, c_i^{(|C_i|)}\}$, 传输测量值之和为 $M_i = \sum_{l=1}^{|C_i|} M_i^{(l)}$, 总的能量成本为

$$E_i = \sum_{l=1}^{|C_i|} E_i^{(l)}.$$

3.2 网络部署和分析

为了简化模型分析并与先前研究展开定量分析对比, 对本文压缩感知数据融合架构施加一些约束。图 1 给出了簇的逻辑树, 该树由 i 层上相同的 n 个节点和随机叶节点 $N_i^{(0)} \geq n$ 构成且 $i \geq 2$ 。考虑一个 N 节点网络, 且 $N = N' + n^T$, 则有如下等式:

$$N_1^{(1)} + N_1^{(2)} + N_1^{(3)} + \dots + N_i^{(|C_i|)} = N$$

$$N_1^{(1)} + N_1^{(2)} + N_1^{(3)} + \dots + N_1^{(|C_i|)} = N' \quad (1)$$

如图 1 所示, 在层次 i , $d_i^{(l)}$ 是簇头 $c_i^{(l)}$ 及其 $i-1$ 层子簇头 $c_{i-1}^{((l-1) \cdot n+1)}, c_{i-1}^{((l-1) \cdot n+2)}, \dots, c_{i-1}^{((l-1) \cdot n+n)}$ 之间的距离和。簇头数据为 $|C_i| = n^{T-i}$ 。第 l 个簇 $s_i^{(l)}$ 的面积与该层上的所有其他簇相同, 表示为 s_i , 既结合了 $i-1$ 层的 n 个子区域, 又满足 $s_i = n s_{i-1}$ 。此外还对随机部署的传感器网络施加了

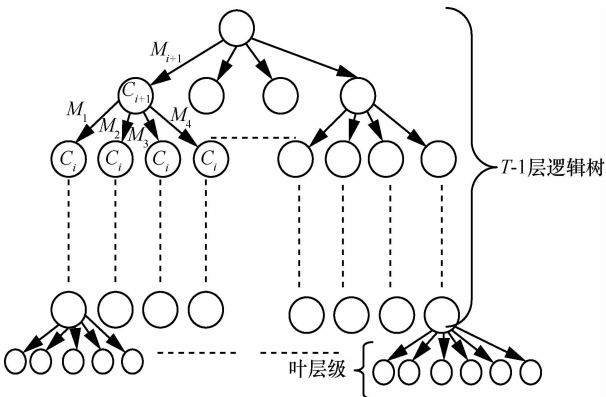


图 1 基于压缩感知的数据融合架构

如下约束。

(1) 第 1 层每个簇至少有 n 个节点。这一属性要求必须要实现 n 个节点在同一簇的概率最大化:

$$\max P_1 = \frac{n|C_1|}{N} = \frac{n^T}{N} = \frac{1}{1 + N'/n^T} \quad (2)$$

它要求 N'/n^T 最小化。根据历史经验, 倾向于设置 $n=4^a$ 且 $a \in \mathbb{N}^+$, 以保证每个正方形簇实现全区域覆盖, 且不会导致两个相邻簇交叉。因此, $T = \lceil \log_n N \rceil$, 得到 N'/n^T 的最小值。

(2) 剩余传感器随机均匀部署于 $|C_1|$ 个簇中, 需要实现 N' 个节点位于 $|C_1|$ 个簇中的概率最大化。

$$\max P_2 = \frac{|C_1|}{N'} = \frac{n^{T-1}}{N'} \quad (3)$$

这在约束(1)中已经实现。这一网络部署的主要优点在于, 基于二维随机部署网络拓扑, 与实际传感器部署情况一致, 也可以处理 $N=n^T$ 这一条件未被满足时的相关问题。此外, 簇头数量最大值为 n^{T-1} , 叶节点为 $N - n^{T-1} \geq n^T - n^{T-1} = (n-1)n^{T-1}$ 。如果 $n > 2$, 则 $N - n^{T-1} > n^{T-1}$ 。这一结果表明, 只有少部分节点参与了多层次数据处理和融合, 其余节点的唯一工作就是将其数据直接发往簇头。通过在每个工作周期内随机选择不同的簇头实现负载分配均衡。

4 数据融合过程

在初始阶段, 每个区域的 $N_1^{(0)} - 1$ 个传感器只将它们的初始数据发往簇头 $c_1^{(0)}$, 采用的策略与参考文献[10,11]相同, 以降低数据融合冗余度。簇头 $c_1^{(0)}$ 将这些数据压缩为 $M_1 = K \log N_1^{(0)}$ 个随机测量值。在层次 $i (i \geq 2)$, 第 l 个簇头 $c_i^{(l)}$ 从其子簇头 $c_{i-1}^{(j)}$ 接收 $M_{i-1}^{(j)}$ 个随机测量值, 且 $j \in [(l-1) \times n + 1, (l+1) \times n + n]$, 然后运行基于压缩感知的数据恢复算法以重构冗余数据。在从子节点收集到所有数据后, 簇头选择 $M_i^{(l)}$ 个信号随机测量值, 并将其发往层次 $i+1$ 上的母簇头。根据压缩感知理论, 对稀疏度为 K 的信号, 随机测量值 $O(K \log N)$ 足以完全表示基数为 N 的原始信号。采用参考文献[10,11]中的方法设置阈值 $M = K \log N$ 优化数据传输量。提出设置多个阈值 $M_i^{(l)} = K \log N_i^{(l)}$, 最高层 T 的上限为 $O(K \log N)$ 。当 i 较小时 M_i 也较小, 这一特性可以降低数据传输次数, 进而显著降低能耗。

5 理论分析

将上述的数据融合方案命名为 HDACS, 从测量值数目、压缩比和能耗 3 方面对其进行性能分析。

5.1 测量值数目

对于 $n^{T+1} \geq N = n^T + N' \geq n^T$, 有 $n^{i+1} \geq N_i^{(i)} \geq n^i$, 并且测量值 $M_i^{(i)}$ 取值范围为 $[iK \log n, (i+1)K \log n]$. 整个数据融合任务传输的总测量数目为:

$$M = \sum_{i=1}^T M_i = \sum_{i=1}^T \sum_{l=1}^{|C_l|} M_i^{(l)} = \sum_{l=1}^{|C_l|} (N_1^{(l)} - 1) + \sum_{i=2}^T \sum_{l=1}^{|C_l|} (n-1) M_{i-1}^{(l)} \quad (4)$$

设 $S_1 = \sum_{i=1}^{T-1} \frac{i}{n^i}$, $S_2 = \sum_{i=1}^{T-1} \frac{1}{n^i}$, 得到 S_1 的闭型为:

$$S_1 = \frac{1/n(1-1/n^{T-1})}{(1-1/n)^2} - \frac{T-1}{n^T(1-1/n)} \quad (5)$$

S_2 的闭型为:

$$S_2 = \frac{1/n \times (1-1/n^{T-1})}{1-1/n} \quad (6)$$

于是, 数据传输次数 M 的下限为:

$$\Omega(M) = N - n^{T-1} + K(n-1)n^{T-1} \log n S_1 \quad (7)$$

上限为:

$$O(M) = N - n^{T-1} + K(n-1)n^{T-1} \log n (S_1 + S_2) \quad (8)$$

另外, 如果使用相同的数据架构传递数据, 则基于 NCS 方案^[9]的总测量值为 $M_{NCS} = N \times K \log N$. 在参考文献[10,11]中, 基于 HCS 方案的总测量数目为:

$$\begin{aligned} M_{HCS} &= \sum_{l=1}^{|C_l|} (N_1^{(l)} - 1) + \sum_{i=2}^T \sum_{l=1}^{|C_l|} (n-1) K \log N \\ &= N - n^{T-1} + K(n-1)n^{T-1} \log n S_2 \end{aligned} \quad (9)$$

假设稀疏度 K 为单位稀疏度, 以去除数据字段对数据融合比较的影响. 图 2 给出了本文 HDACS 算法、参考文献[9]的 NCS 数据融合算法和参考文献[10,11]的 HCS 算法总体数据

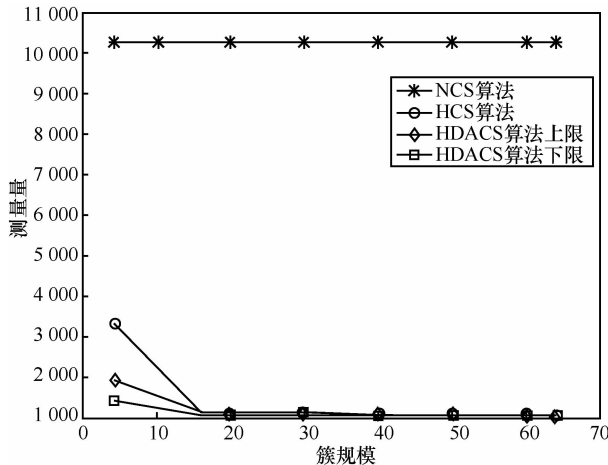


图 2 网络有 1024 个节点时的测量数目比较

传输测量值的定量比较结果。比较条件是 1024 个传感器节点, 簇大小为 $n=4, 16, 64$. 从图 2 可以看出, 簇规模越大, 数据传输所需的测量量越小. 然而, 这一理论分析没有考虑底层网络架构的路由协议, 仅仅提高局部簇内的簇规模, 所有节点将其采集数据直接发往簇头, 定将导致严重的数据溢出和数据丢失, 因此将簇头大小固定为 4 和 16, 如图 3 所示. 图 3 表明, 在两种固定簇头大小的条件下, 当传感器节点数量增加时, 总体数据传输测量量将会变化, NCS 算法具有严重的数据冗余, HCS 算法需要的测量量也要多于本文算法。

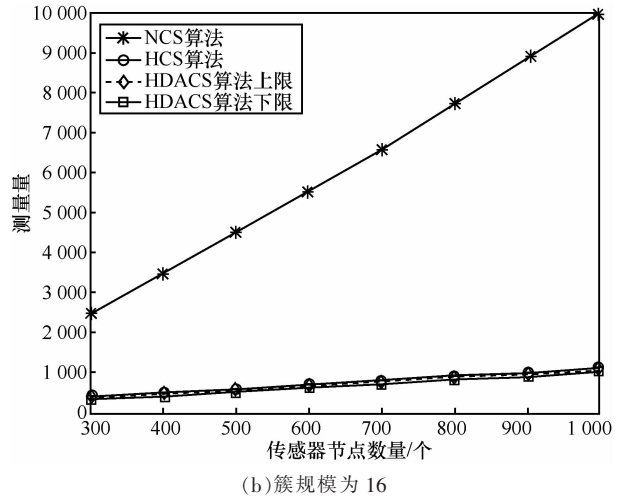
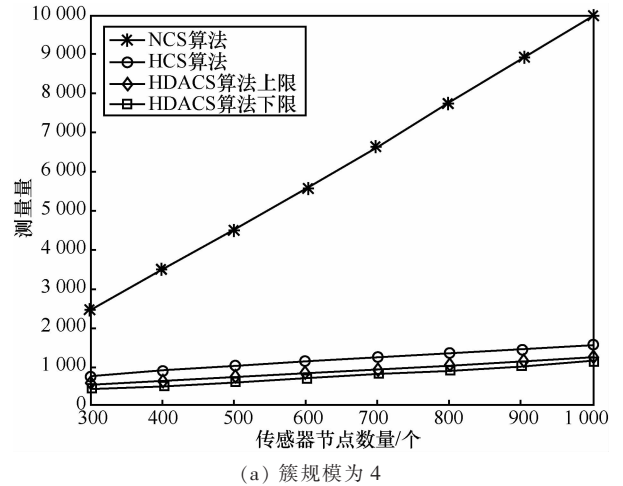


图 3 不同规模传感器网络时的测量值比较

5.2 压缩率

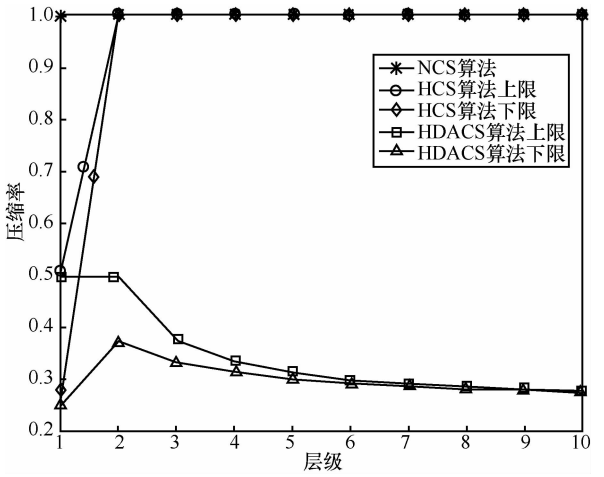
数据压缩率 γ_i 计算如下:

$$\gamma_i = \begin{cases} \frac{M_1^{(i)}}{N_1^{(i)}} = \frac{K \log N_1^{(i)}}{N_1^{(i)}}, & i = 1 \\ \frac{M_i^{(i)}}{\sum_{j=(i-1)n+1}^{(i-1)n+n} M_{i-1}^{(j)}}, & i \geq 2 \end{cases} \quad (10)$$

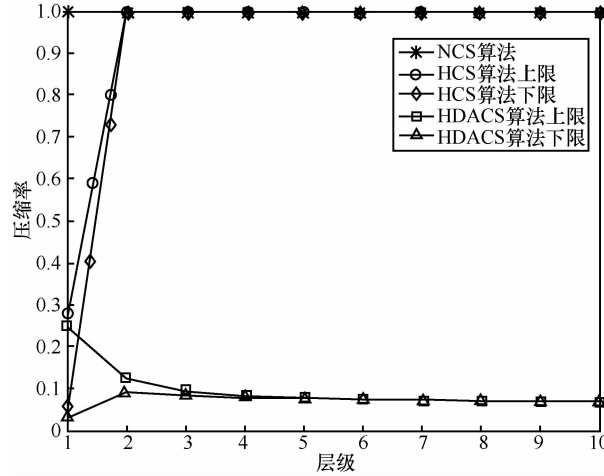


当 $i \geq 2$ 时,压缩率范围为 $\gamma_i \in \left[\frac{1}{n} \left(1 + \frac{1}{i-1} \right), \frac{1}{n} \left(1 + \frac{1}{i} \right) \right]$ 。

图4给出了簇规模分别为4和16时,3种算法每一层的压缩率变化情况。从图4中可以看出,NCS算法完全没有压缩;HCS算法是当数据量到达全局阈值时使用压缩感知算法压缩数据,因此在第一层具有压缩效果;本文分层算法的压缩率最为明显,每一层的压缩率均低于0.5。因为与中央汇聚节点距离较近的节点为中间节点,其能耗往往高于其他节点,因此本文压缩特性非常有用。这些节点在应用层任务中节省下来的能量可以均衡路由层的能耗,进而提升网络寿命。



(a)簇规模为4



(b)簇规模为16

图4 不同规模传感器网络下层数上升时的变化情况

5.3 能耗

传感器节点的能耗主要集中在数据传输上,传输能量成本 $E_i^{(0)}$ 往往是传输距离 $d_i^{(0)}$ 和数据量 $M_i^{(0)}$ 的函数。因此,可以将 $E_i^{(0)}$ 建模为 $E_i^{(0)} = c_s + c d_i^{(0)} M_i^{(0)}$,于是每一层传输能量总成

本为 $E_i = \sum_{l=1}^{|C_i|} E_i^{(l)}$ 。其中 c_s 为常数,表示每个数据传输任务的启动能耗; c 也是常数,表示单位距离单位数据量的传输成本。

设 $d_i^{(l)} = \sum_{t=1, (x_t, y_t) \in c_i} \left[(x_t - x_{c_i})^2 + (y_t - y_{c_i})^2 \right]^{1/2}$, 其中 (x_{c_i}, y_{c_i})

和 (x_i, y_i) 分别是 c_i^l 及其子节点的位置坐标。在一个大型随机均匀密集部署的传感器网络中,如果 $i \geq 1$,则:

$$d_i^{(l)} = 4 \frac{(n-1)}{s_i} \int_0^{\phi_i} \int_0^{\phi_i} (x^2 + y^2)^{1/2} dx dy = \frac{1}{12} \pi n^{\frac{i-1}{2}} s_i^{1/2} (n-1) \quad (11)$$

其中 $b_i = \frac{1}{2} s_i^{1/2}$ 。当 $i=1$ 时, $d_1^{(l)} = \frac{1}{12} \pi s^{1/2} (N_1^{(l)} - 1)$ 。

最终总能耗为:

$$\begin{aligned} E &= \sum_{i=1}^T \sum_{l=1}^{|C_i|} E_i^{(l)} = \sum_{l=1}^{|C_1|} c_s + c d_1^{(l)} + \sum_{i=2}^T \sum_{l=1}^{|C_i|} c_s + c d_i^{(l)} M_i^{(l)} \\ &= n^{T-1} c_s + c \frac{1}{12} \pi s^{1/2} \sum_{l=1}^{|C_1|} (N_1^{(l)} - 1) + \sum_{i=2}^T n^{T-i} c_s + \\ &\quad \sum_{i=2}^T c \frac{1}{12} \pi n^{\frac{i-1}{2}} s^{1/2} (n-1) K \log N_{i-1}^{(l)} \\ &= n^{T-1} c_s + c \frac{1}{12} \pi s^{1/2} \sum_{l=1}^{|C_1|} (N_1^{(l)} - 1) + n^{T-1} c_s S_2 + \\ &\quad c \frac{1}{12} \pi s^{1/2} (n-1) K \sum_{i=2}^T \sum_{l=1}^{|C_i|} n^{\frac{i-1}{2}} \log N_{i-1}^{(l)} \end{aligned} \quad (12)$$

设 $S'_1 = \sum_{i=1}^{T-1} i n^{-i/2}$, 得到 S'_1 的闭型为:

$$S'_1 = \frac{n^{-1/2} (1 - n^{-(T-1)/2})}{(1 - n^{-1/2})^2} - \frac{(T-1) n^{-T/2}}{1 - n^{-1/2}} \quad (13)$$

设 $S'_2 = \sum_{i=1}^{T-1} n^{-i/2}$, 其闭型为:

$$S'_2 = \frac{n^{-1/2} (1 - n^{-(T-1)/2})}{1 - n^{-1/2}} \quad (14)$$

因此,总能耗 E 的下限为:

$$\begin{aligned} \Omega(E) &= n^{T-1} c_s (1 + S_2) + c \frac{1}{12} \pi s^{1/2} \\ &\quad \left[(N - n^{T-1}) + K(n-1) n^{T-1} S'_1 \log n \right] \end{aligned} \quad (15)$$

E 的上限为:

$$\begin{aligned} O(E) &= n^{T-1} c_s (1 + S_2) + c \frac{1}{12} \pi s^{1/2} \\ &\quad \left[(N - n^{T-1}) + K(n-1) n^{T-1} (S'_1 + S'_2) \log n \right] \end{aligned} \quad (16)$$

利用类似的推导方法,获得了参考文献[9]中 NCS 算法采用相同数据融合架构时的传输能耗:

$$\begin{aligned} E_{\text{NCS}} &= n^{T-1} \mathbf{c}_s (1 + S_2) + \mathbf{c} \frac{1}{12} \pi s^{1/2} \\ \log N &\left[(N - n^{T-1}) + K(n-1)n^{T-1}S_2' \right] \end{aligned} \quad (17)$$

参考文献[10,11]中 HCS 算法的能耗为:

$$E_{\text{HCS}} = n^{T-1} \mathbf{c}_s (1 + S_2) + \mathbf{c} \frac{1}{12} \pi s^{1/2} \left[(N - n^{T-1}) + K(n-1)n^{T-1} \log NS'_2 \right] \quad (18)$$

为了忽略所有常值参数的影响,假设 c_s 、 c 、 K 、 s 均为1。图5给出了簇规模为4,网络规模为300、400、500、600、700、800时的最终传输能耗 E 。本文 HDACS 算法与其他算法相比,能效最高。在利用 Java 开发的 SIDnet-SWANS 仿真平台上设置一个二维不规则网络,以证明本文分层模型的可行性和可靠性。

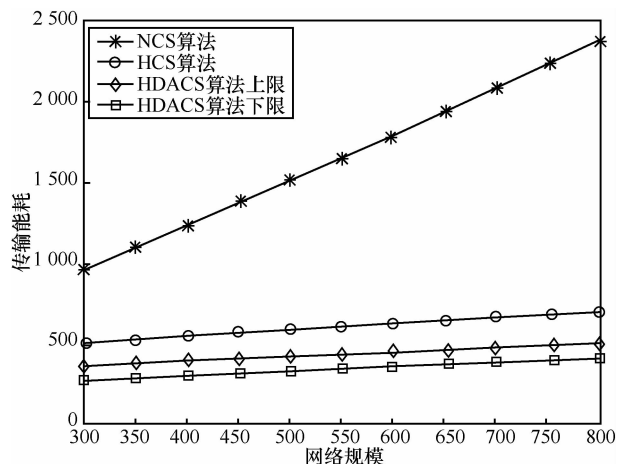


图 5 网络规模增加时的能耗变化

6 仿真实验

在栖息地监测监控等多种实际应用中,温度、声音、压力测量值的数据字段往往是平滑的,忽略稀疏度 K 的变化在每一层的影响。为了获得具有相同稀疏度 K 的稀疏信号表示,往往选择带有均衡噪声的平滑数据字段。在采集随机测量值前对每个数据收集簇进行离散余弦变换(DCT)。选择 DCT 的主要原因是:与离散傅立叶变换(DFT)相比,离散余弦变换生成信号表示消失矩的速度更快,且生成的系数为实数;它不像小波变化那样,要求测量基数为 2 的幂。通过将低于阈值的这些数值设为 0,然后对 DCT 系数进行截断,使信号进一步稀疏化。阈值设为第一个主要数值的 $\alpha\%$ 。在实际仿真中, α 选为 0.01、0.005。

本文研究的对象是密集型大规模网络拓扑,采用多尺度路由协议进行数据传输,与分层数据融合机制非常匹配,可以保证任意两个节点间均有最短路径。在本文仿真时,CS 编码采用随机高斯矩阵,数据重构算法采用参考文献[12]中的 CoSaMP 算法。与其他重构算法相比^[13,14],CoSaMP 算法既可以保证运算速度,又可以提供严格的误差边界。实验结果取 50 次仿真运行的平均值。实验中的其他仿真参数见表 2。

表 2 实验环境中的基本参数

参数	取值
节点 i 发送单位数据的能耗 $E_{i\text{-TCR-RX}}$	50×10^{-6} J/bit
节点 i 接收单位数据的能耗 $E_{i\text{-TCR-EC}}$	50×10^{-6} J/bit
节点 i 成功发送一位数据通过 1 m 距离的能耗 $E_{i\text{-TCR-PA}}$	100×10^{-9} J/bit/m ²
节点 i 上微控制单元的能耗 $E_{i\text{-MCU}}$	50×10^{-6} J/bit
节点 i 上中央处理器的计算能耗 $E_{i\text{-SB}}$	50×10^{-6} J/bit
每个数据分组的长度	1 024 bit
每个控制分组的长度	64 bit

参考文献[15]的 SIDnet-SWANS 是针对各种应用的感知器网络仿真环境,提供各种基于 Java 的虚拟工具。利用这一平台测试本文算法的性能。图 6 给出了 400 个节点条件下,在 SIDnet-SWANS 平台上部署新设计的压缩感知数据融合架构时用户界面的一个快照。基于 SIDnet-SWANS 平台评估算法性能,并验证理论分析相关结论的正确性。

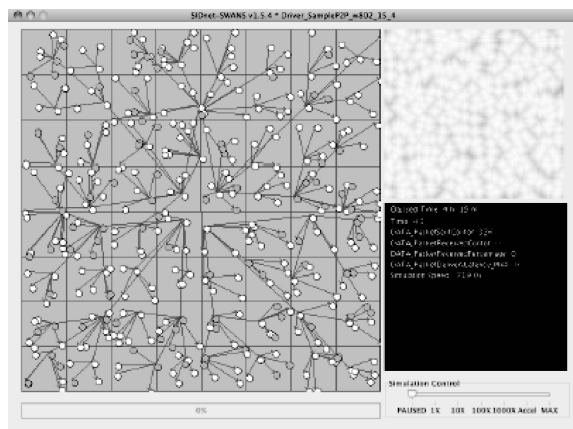


图 6 SIDnet-SWANS 仿真平台的用户接口

在 5 种网络规模下测试算法性能, 节点数量分别为 300、400、500、600、700, 数据字段平滑且带有均匀分布的加性白噪声。在各种网络规模下, 设置 $n=4, T=3$ 。第一层的叶节点数量 N_1 可调, 视二维随机部署传感器网络的特征而定, 因此 $N=N_1+n^{T-1}$ 。在恢复阶段, 借鉴参考文献[13]中基



于模型的 CoSaMP 算法的思路, DCT 可使系数的支持位置可视, 同时针对基于 DCT 的全体信号设计一种新的 CoSaMP 算法, 实现数据的准确恢复, 将信噪比 (SNR) 定义为每个传感器的信号功率比与融合中心恢复误差的对数。得到的仿真结果如图 7 所示。从图 7 可以看出, 感知器规模变化并不会影响 SNR 性能。

图 8 给出了 400 个节点时网络传输能耗的比较情况。Ratio1 定义为本文 HDACS 算法和 NCS 数据融合算法的传输能耗之比, Ratio2 定义为本文 HDACS 算法和 HCS 数据融合算法的传输能耗之比。从图 8 可以看出, Ratio1 低于 0.5, 意味着与 NCS 相比, 节约了 50% 的传输能量; Ratio2 几乎等于或小于 1, 这是因为在第 1 层大部分节点只传输数据。本文 HDACS 算法和 HCS 算法采用相同的策略, 对这些叶节点只传递原始数据, 这就解释了为什么 Ratio2 值等于 1。对于第 2 层及更高层作为数据采集簇的节点, Ratio2 小于或等于 0.633, 与预期相符, 最高层节点可以节约 70% 的能耗。此外, 迄今获得的结论在某种程度上取决于 MAC 层每次传输的帧大小。如果数据量增大, 则会把数据分割成多个帧再传输, 这必将导致功率消耗。因为本文 HDACS、NCS 和 HCS 算法的比较主要是 $\log N_i$ 和 $\log N$ 的比较, 假设一个帧的大小为 m , 则数据量为 N_i 和 N 时的帧数量分别为 $\left\lceil \frac{\log N_i}{m} \right\rceil$ 和 $\left\lceil \frac{\log N}{m} \right\rceil$ 。如果 $N_i \approx n^i$, 则两个帧的数量分别为 $\left\lceil \frac{i \log n}{m} \right\rceil$ 和 $\left\lceil \frac{T \log n}{m} \right\rceil$ 。当 $i=2, n=4, m=4$ 且 $T=4$ 时, 帧数分别为 1 和 2, 这便解释了为何使用 HDACS 算法可以节省 50% 的传输能量。

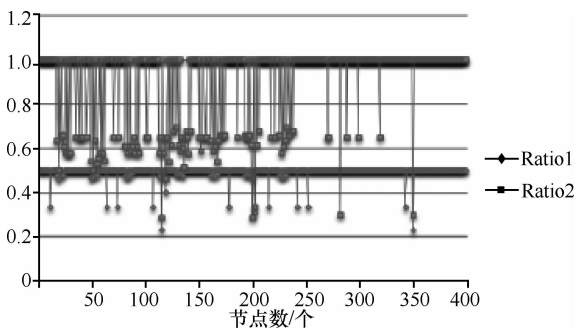


图 8 网络有 400 个节点时的传输能耗率

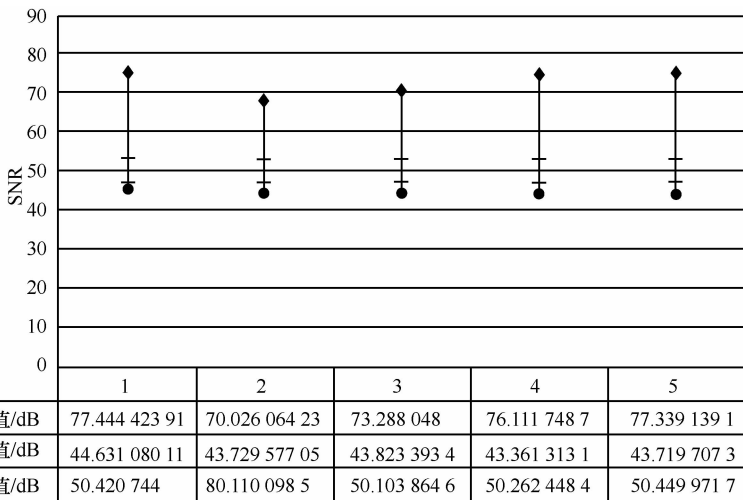


图 7 网络有 300、400、500、600、700 个节点时的数据恢复性能

7 结束语

本文针对大规模密集型传感器网络, 提出一种基于压缩感知的高能效分层数据融合架构。然后从数据传输总次数、数据压缩率和传输能耗 3 个方面对分层数据融合模型进行了理论分析。最后在 SIDnet-SWANS 仿真平台上部署了本文模型, 并在不同的二维随机部署感知器网络下展开了测试。测试结果表明, 本文模型可以保证从所有节点采集到的数据的准确性。与先前研究相比, 本文模型可以显著降低传输能耗。在下一步工作中, 将考虑稀疏度的可变性, 这就意味着数据字段的复杂性增加, 需要建立自适应模型来处理数据融合字段的动态性。此外, 需要研究其他压缩感知恢复算法, 以降低恢复操作的复杂性, 提高信号恢复的准确性。

参考文献

- 1 许建, 杨庚, 陈正宇等. 基于二次独立集的数据融合调度算法. 通信学报, 2014, 12(1):1306~1313
- 2 师东生. WSN 中一种基于混合 CS 的分簇数据收集方案. 计算机应用研究, 2014, 31(3): 892~897
- 3 杨庚, 李森, 陈正宇等. 传感器网络中面向隐私保护的高精度数据融合算法. 计算机学报, 2013, 36(1): 189~199
- 4 乐俊, 张维明, 肖卫东等. 无线传感器网络中一种基于非均匀划分的分簇数据融合算法. 计算机研究与发展, 2011, 48(Z2): 247~254
- 5 余阳, 吴银峰, 冯仁剑等. 应用于管道气体泄漏监测的 WSN 层级式数据融合方法. 高技术通讯, 2012, 22(1):1~7
- 6 林蔚, 尹娟. 无线传感器网络含有缺失数据的数据融合研究. 计算机应用研究, 2012, 28(4):631~636

- 7 马守明, 王汝传, 叶宁. 基于信誉度集对分析的 WSN 完全数据融合. 计算机研究与发展, 2011, 48(8): 1652~1658
- 8 李小亚, 潘家辉. 无线传感器网络能量平衡的数据汇聚路由算法. 华南师范大学学报(自然科学版), 2012, 44(2): 53~57
- 9 Luo C, Wu F, Sun J, *et al.* Compressive data gathering for large-scale wireless sensor networks. Proceedings of the 15th Annual International Conference on Mobile Computing and Networking, Beijing, China, 2009: 145~156
- 10 Luo J, Xiang L, Rosenberg C. Does compressed sensing improve the throughput of wireless sensor networks. Proceedings of 2010 IEEE International Conference on Communications (ICC), Cape Town, South Africa, 2010: 1~6
- 11 Xiang L, Luo J, Vasilakos A. Compressed data aggregation for energy efficient wireless sensor networks. Proceedings of 8th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks (SECON), Salt Lake City, Utah, USA, 2011: 46~54
- 12 Gilbert A C, Strauss M J, Tropp J A, *et al.* One sketch for all: fast algorithms for compressed sensing. Proceedings of the Thirty-Ninth Annual ACM Symposium on Theory of Computing, San Diego, California, USA, 2007: 237~246
- 13 Park J, Jung B C, Kim J M, *et al.* Sparse signal recovery with parallel orthogonal matching pursuit and its performances. Journal of the Korea Institute of Information and Communication Engineering, 2013, 17(8): 1784~1789
- 14 Zhang W, Huang B, Zhou T. An improvement on StOMP for

sparse solution of linear underdetermined problems. Proceedings of 32nd Chinese Control Conference (CCC), Xi'an, China, IEEE, 2013: 1951~1956

- 15 Jeong Y S, Han Y H, Park J J, *et al.* MSNS: mobile sensor network simulator for area coverage and obstacle avoidance based on GML. EURASIP Journal on Wireless Communications and Networking, 2012(1): 1~15

[作者简介]



赵建军,男,保定学院实验师,主要研究方向为数据挖掘、网络安全。

王怀宇,男,保定学院讲师,主要研究方向为数据挖掘、嵌入式系统。

赵泽阳,男,保定学院硕士生,主要研究方向为数据收集。

陈生昌,男,博士,浙江大学教授、博士生导师,主要研究方向为压缩感知理论、数据收集技术、计算地球物理。

(收稿日期:2014-03-20)

(上接第91页)

牌,具有使用便捷、安全性高、运行效率高、成本低、功能强大和人性化等特点,非常适合移动应用的身份认证。

参考文献

- 1 郑强, 高能, 张令臣. 基于 SM3 算法的动态口令卡的设计与实现. 计算机应用与软件, 2013, 30(2): 14~17
- 2 张利华, 章丽萍, 张有光等. 基于口令的远程身份认证及密钥协商协议. 计算机应用, 2009, 29(4): 924~927
- 3 林元明. 基于手机令牌的身份认证系统的研究与实现. 厦门大学硕士学位论文, 2009
- 4 韦广剑. 基于 Android 令牌的动态密码认证系统的研究与实现. 武汉理工大学硕士学位论文, 2012
- 5 郭建昌, 郭茂文, 黎艳. 一种电信级手机令牌动态口令系统的架构设计及其实现. 电信科学, 2012, 28(10): 94~98
- 6 叶晰, 叶依如, 岑琴. 基于手机令牌的动态口令身份认证系统的实现. 计算机系统应用, 2010, 19(10): 32~36
- 7 刘超凡, 张永良, 肖刚等. 基于滚动指纹数字签名的电子商务安全认证. 计算机应用, 2012, 32(2): 475~479
- 8 赵新杰, 王韬, 郭世泽等. AES 访问驱动 Cache 计时攻击. 软

件学报, 2011, 22(3): 572~591

- 9 闫茂德, 纪志强, 贺显曜等. AES 与 ECC 混合加密算法的无线数据通信系统设计. 微电子学与计算机, 2007, 24(7): 135~138

[作者简介]



郑兆华,女,海南大学信息科学技术学院讲师,主要研究方向为信息安全与移动互联网应用。

彭金莲,女,海南大学信息科学技术学院教授,主要研究方向为农业信息技术与网络安全。

程杰仁,男,海南大学信息科学技术学院教授,主要研究方向为网络安全。

(收稿日期:2014-08-20)