

无线传感器网络簇内数据汇聚方法

郭江鸿^{1,2}, 陈德礼¹, 刘志宏²

(1 广东嘉应学院 计算机系, 广东 梅州 514015; 2 西安电子科技大学 计算机学院, 陕西 西安 710071)

摘 要: 数据汇聚是降低传感器网络传输能耗的重要手段. 分簇传感网数据汇聚通常包括簇内汇聚及汇聚结果传输, 现有方案大多关注簇构造或簇内汇聚结果传递到基站的效率. 文章提出一种分簇传感网簇内数据汇聚方法, 通过位置信息选择代理点, 仅与代理点有不同测量值的节点进行数据发送, 减少了簇内数据汇聚中的数据传输量. 分析与试验表明, 与相关方案比较, 文章方案簇内汇聚的数据传输量较小, 降低了通信开销.

关键词: 传感器网络; 数据汇聚; 代理节点

中图分类号: TP301

文献标识码: A

文章编号: 1000-7180(2013)09-0013-04

Inner-cluster Data Aggregation Scheme for WSN

GUO Jiang-hong¹, CHEN De-li¹, LIU Zhi-hong²

(Department of Computer Science and Technology, Jiaying University, Meizhou 514015, China;

(Department of Computer Science and Technology, Xidian University, Xi'an 710071, China)

Abstract: Data aggregation is an important method to reduce energy consumption in wireless sensor networks. Data aggregation for cluster-based wireless sensor networks usually consists of in-cluster aggregation and fusion data transmission, most of existing schemes are mainly focus on the cluster construction or the efficiency of fusion data transmission. Authors proposed an inner-cluster data aggregation scheme for wireless sensor networks, agent nodes are selected with assist of location information and only the nodes have different readings with agent nodes transmit their data to cluster head, thus reduce the inner-cluster transmissions. Analysis and experiments show that our scheme reduces inner-cluster transmissions effectively and lowers the communication overhead.

Key words: sensor network; data aggregation; agent nodes

1 引言

无线传感器网络 (Wireless Sensor Network, WSN) 一般通过随机抛撒或确定性投放等方式进行部署, 在战场监控、灾难拯救、目标跟踪等方面有广泛应用. 无线传感网的一个重要功能是收集部署区域的物理数据, 数据节点通过彼此协作完成测量任务并将结果送往远方基站进行进一步处理. 传感器节点一般依靠电池供电且往往部署在较为恶劣的环境中, 大规模更换节点电池较为困难, 因此如何减少节点能耗成为设计传感器网络各种协议的主要考虑因素之一.

由于无线通信的特点, 一个节点发送数据将导致所有其通信范围内所有节点接受数据, 通信开销是传感器网络能耗的主要开销, 数据汇聚技术是通过减少冗余数据传输降低通信开销的重要手段.

目前提出来诸多传感器网络数据汇聚方案. 定向扩散方案^[1]是典型的基于平面型网络的协议; LEACH 方案^[2]与 HEED 方案^[3]对成簇及簇内汇聚点的选择方面做了研究; NECCHI 等提出 EERINA 协议^[4]将 Gossip 算法应用于分簇网络的数据汇聚; MOSTAFIZUR 等对 EERINA 协议进行了改进^[5]; ZHENG 等^[6]使用 Slepian-Wolf 编码在分簇网络中进行分布式数据汇聚; 杨军等研究了分簇

网络的无缝覆盖并采用中心化自回归模型对簇内汇聚结果进行预测^[7];杨国宁等提出了基于最优融合集合的数据融合方法^[8].

由于无线通信的广播特性,一个节点发送的数据将被所有其通信范围内的邻居节点接收,减少数据发送是降低通信开销最有效的手段,但上述基于分簇网络的数据汇聚方案主要关注于簇的构造或簇内汇聚结果到基站的发送效率,簇内汇聚中所有簇内节点都进行数据发送,通信开销相对较大.

为减少簇内汇聚所需的通信量,文章提出一种无线传感器网络簇内数据汇聚方法(Inner-cluster data aggregation for wireless sensor networks, ICDA),首先等分簇为 4 个区域,节点通过位置信息确定所在区域及本节点是否为备选代理点;簇头从备选代理点中为每个区域选取代理点;在数据汇聚过程中,代理点先进行发送,仅与代理点有不同测量值的节点进行后续发送.

2 网络模型及相关假设

文中方案主要针对分簇传感网,并做如下假设:

(1)簇内节点在簇头 1-hop 范围内,如图 1 所示.

(2)节点在部署后通过硬件(如 GPS)或定位算法^[9]确定位置.

(3)簇的覆盖范围相对较小,除个别节点受噪音干扰导致测量值与真实值偏差较大,大部分节点测量值与真实值偏差较小.

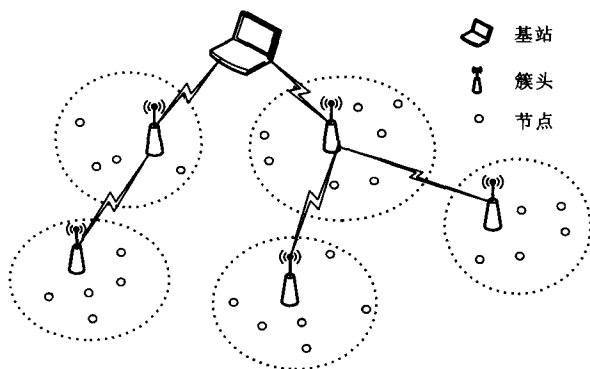


图1 分簇传感网结构

3 ICDA 方案简介

在网络部署后,经过初始化阶段的网络分簇及节点定位,每个节点知道簇头 ID 及自身位置信息. ICDA 方案主要由子区域划分、备选代理点判定、代理点判定、数据汇聚等部分组成.

3.1 区域划分

簇头 A 根据其位置信息 (X_A, Y_A) , 将簇 4 个区域 S_1, S_2, S_3, S_4 , 其覆盖范围分别为 $(0, \pi/2], (\pi/2, \pi], (\pi, 3\pi/2], (3\pi/2, 2\pi]$. 区域 S_1 如图 2(a) 中浅色阴影扇形区域所示. 簇头在簇内广播 (X_A, Y_A) , 簇内每个节点 N_i 根据位置信息 (X_{N_i}, Y_{N_i}) 判断所属区域:

$$\theta = \arctan\left(\frac{Y_{N_i} - Y_A}{X_{N_i} - X_A}\right) \quad (1)$$

设 $\theta \in (0, \pi/2]$, 则:

$$\begin{cases} \text{若 } X_{N_i} > X_A, \text{ 则 } N_i \in S_1 \\ \text{若 } X_{N_i} < X_A, \text{ 则 } N_i \in S_3 \end{cases} \quad (2)$$

同理可判断 $\theta \in (\pi/2, \pi]$ 时节点位于区域 S_2 或 S_4 .

簇内节点发送区域归属信息 $\langle \text{id}, \text{Area} \rangle$ 到簇头, Area 为节点所在区域;通过区域归属信息,簇头得到每个区域的节点数.

3.2 备选代理点判定

设区域 1 的范围为 S_1 , 以 S_1 分界点 A、B 为圆心, 以通信半径为半径做圆 SA, SB , 定义区域 S_1 的代理区域 $R_{S_1} = S_1 \cap SA \cap SB$, 如图 2(a) 中深色阴影部分所示.

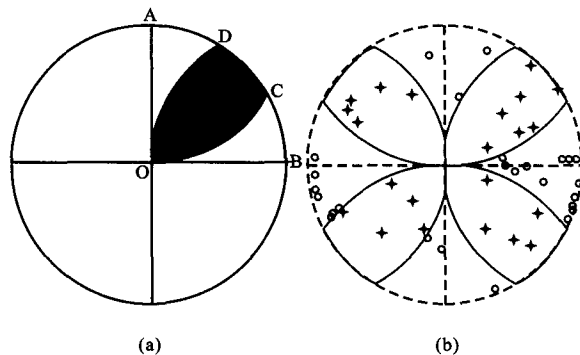


图2 区域与代理区域

定理 1 ICDA 中区域 S 的代理区域 R_S 中任一点 V 到区域 S 中任一点 U 的距离小于等于节点通信半径.

定理 1 可通过平面几何容易得证, 因此文章省略对此证明.

图 2(b) 为 50 个节点在簇内随机分布, 图中“o”为普通数据节点, “*”为备选代理点.

3.3 代理点判定

簇头从各区域备选代理点中随机选取一个节点作为该区域代理点, 因为代理点的通信范围覆盖整个簇, 簇内任一节点至少有一个代理点数据作为参照对象. 同时, 簇头保存代理点信息如表 1 所示. 表

中 N_1, N_2, N_3, N_4 为区域 $S1, S2, S3, S4$ 中的代理点, $Num_1, Num_2, Num_3, Num_4$ 分别表示各区域节点数. 簇头在簇内广播代理点信息 $\langle N1, S1, N2, S2, N3, S3, N4, S4 \rangle$.

表1 代理点表

ID	Area	Num
N_1	S1	Num_1
N_2	S2	Num_2
N_3	S3	Num_3
N_4	S4	Num_4

3.4 数据汇聚

收到数据请求后,各节点进行测量.各区域代理点先发送其测量值,区域内其它节点比较自己的测量值与代理点测量值,比较结果不相同则行数据发送,否则保持沉默;若某区域内无代理点,则该区域内所有节点都发送数据.

为进行数据汇聚,簇头需建立一张汇聚表(汇聚表初始化如表2),并将接受到的各代理点测量值填入表中,此后收到的数据将是与代理点测量值不同的数据.表中 Num 为计数值,表示测量值相同且位于相同区域的节点数.

表2 初始化汇聚表

ID	Area	Data	Num
N_1	S1	D_N_1	Num_1
N_2	S2	D_N_2	Num_2
N_3	S3	D_N_3	Num_3
N_4	S4	D_N_4	Num_4

设簇头收到节点 N_w 的测量值 D_N_w , N_w 位于区域 $S_i, i \in [1, 4]$, 簇头按下列步骤进行处理:

- (1) 首先在汇聚表中查找是否存在某节点 N_k , N_k 与 N_w 位于同一区域且其测量值 $D_N_k = D_N_w$;
- (2) 若存在符合条件的 N_k , 则 N_k 的计数值+1, N_k 所在区域代理点的计数值-1, 丢弃 N_w ;
- (3) 若找不到符合条件的 N_k , 则将 $\langle N_k, S_i, D_N_k, 1 \rangle$ 插入表中.

通过上述步骤,簇头可得到簇内各区域的不同测量值及相应计数,在此基础上可进一步进行求均值、求和等运算,完成簇内汇聚.

文中主要关注簇内汇聚中的数据传输量问题,因此对簇头汇聚结果传输到基站的过程不做过多阐述,可选用现有方案完成后续过程.

4 分析与实验

4.1 分析

采用文献[8]的方法以正态函数输出随机值的

方式模拟温度传感器的测量值,数据精度为 0.1. 设节点测量值服从 (μ, σ) 分布, μ 为簇内期望真值, σ 为标准差. 因簇覆盖范围不大,在此范围内大部分节点测量值与真值偏差较小,因此设 $\sigma = 0.5$; 同时对节点测量值取整,即数据精度为 1.

图3为数据精度分别为 0.1 及 1 时,随机生产 30 个服从 $(30, 0.5)$ 正太分布的数据在重复 100 次实验后的均值. 从图3中可看出,两种数据精度下数据均值基本都在 $[\mu - 0.2, \mu + 0.2]$ 范围内,同时,传感器本身的测量误差难以避免,因此适当减小数据精度不会对汇聚结果造成大的影响.

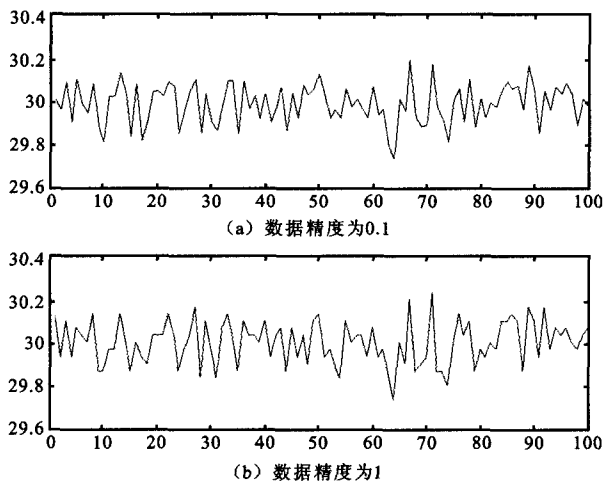


图3 不同数据精度下簇内数据均值

设 n 个测量值集合为 $\{Data_i\}, i \in [1, n]$, 节点测量值为 $Data_i$ 的概率为 p_i , 则任一节点 N_j 与所在区域代理点取相同测量值的概率 P 为:

$$P = \sum_{i=1}^n p_i^2 \quad (3)$$

若簇内节点数为 N , 则理论上进行数据发送的簇内节点数 N_{trans} 为:

$$N_{trans} = 4 + (N - 4) \times (1 - P) \quad (4)$$

4.2 模拟实验

文中采用图1所示的分簇网络(共5个簇), 设簇内节点数为 40, 节点通信半径为 40m, 各簇的测量期望值为 30, 标准差 $\sigma = 0.5$. 文中以数据汇聚中进行数据的节点数衡量各方案进行数据汇聚的簇内通信开销, 用 Matlab 进行仿真并与文献[6-7]做比较.

图4为一次实验中为5个簇生成的数据精度为 0.1 的随机数据图. 图5中 ICDA 表示采用图4数据的一次簇内汇聚中进行数据发送的簇内节点数, $ICDA *$ 表示重复 50 次生成随机数据时簇内汇聚中进行数据发送的簇内节点数均值.

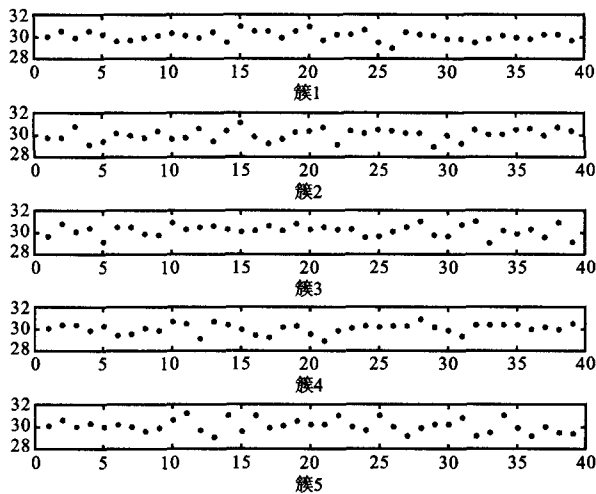


图4 一次实验随机数据图

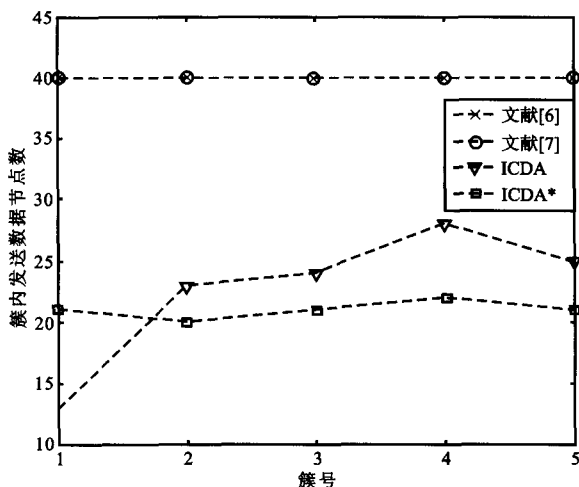


图5 簇内发送数据节点数

从图5可看出文中方案的簇内发送数据节点数低于文献[6]及文献[7],一次实验中簇内发送节点数为文献[6]及文献[7]的32.5%,57.5%,60%,70%,62.5%;50次实验簇内发送节点数的均值分别为文献[6]及文献[7]的52.5%,50%,52.5%,55%,52.5%。主要原因在于文中各节点先与所在区域代理点的测量值比较,比较结果不同才进行发送,同时,文中在不对汇聚结果有较大影响的情况下适当降低了数据精度,提高了簇内数据冗余度,增加了节点与代理点取相同测量值的概率。

在数据做取整处理后,节点测量值为28、29、30、31、32的概率分别为0.13%、15.73%、68.72%、15.73%、0.13%,根据式(3)计算簇内发送数据节点数的理论值为21。图5中进行50次实验后5个簇的簇内发送数据节点数均值为:21、20、21、22、21,与理论值基本吻合,说明文中方案具有一定的可行性。

5 结束语

文中提出了一个无线传感器网络簇内数据汇聚方案,通过节点位置信息进行簇内区域划分,为每个区域选择代理点,仅与所在区域代理点读数不同的节点进行数据发送,同时,在对汇聚结果影响较小的前提下适当降低了数据精度。理论分析与模拟实验均证明了文中方案可有效减少分簇传感网中簇内汇聚的发送消息节点数,降低了数据汇聚的簇内通信开销;同时,实验值与理论值的对比也证明了文中方案的可行性。

参考文献:

- [1] Intanagonwiwat, Govindan C, Estrin R, et al. Directed diffusion for wireless sensor networking [J]. IEEE/ACM transactions on networking, 2003, 11(1): 2-16.
- [2] Heinzelman W R, Chandrakasan A P. An application-specific protocol architecture for wireless micro-sensor network [J]. IEEE Trans Wireless Communication, 2002, 4(1): 660-670.
- [3] Younis O, Fahmy S. HEED: a hybrid, energy-efficient, distributed clustering approach for Ad hoc sensor networks [J]. IEEE Trans. Mobile Computing, 2004, 3(4): 366-379.
- [4] Necchi L, Bonivento A, Lavagno L. Eerina: an energy efficient and reliable in-network aggregation for clustered wireless sensor networks [C] // Proc of Wireless Communications and Networking Conference. China: Hangkong, 2007: 3364-3369.
- [5] Mostafizur M, Mozumdar R, Nan Guofang. An efficient data aggregation algorithm for cluster-based sensor network [J]. Journal of Networks, 2009, 4(7): 598-606.
- [6] Zheng Jun, Wang Pu, Li Cheng. Distributed data aggregation using slepian-wolf coding in cluster-based wireless sensor networks [J]. IEEE Trans on Vehicular Technology, 2010, 59(5): 2564-2574.
- [7] 杨军, 张德运. 基于分簇的无线传感器网络数据汇聚传送协议 [J]. 软件学报, 2010, 5(21): 1127-1137.
- [8] 杨国宁, 冯秀芳. 一种基于最优融合集的多传感器数据融合算法 [J]. 软件学报, 2012, 23(zk1): 134-140.

作者简介:

郭江鸿 男, (1975—), 博士研究生, 讲师. 研究方向为无线移动安全、网络与信息安全、无线传感器网络。