

基于分簇与自适应加权的 WSN 数据融合算法

陶志勇, 王 雪

(辽宁工程技术大学 电子与信息工程学院, 辽宁 葫芦岛 125105)

摘 要: 针对无线传感器网络自身能量有限需要节能的问题, 提出了网络的最优分簇规模算法; 考虑到簇首的负担较重, 提出要选择剩余能量较高的节点作为簇首。这样可以有效延长 WSN 的生命周期, 提高 WSN 的使用效率。最后结合按测量方差的自适应加权数据融合算法对收集的数据进行融合, 减少数据传输, 进一步达到节能的作用。

关键词: 无线传感器网络; 分簇算法; 自适应加权; 数据融合

中图分类号: TP393

文献标识码: A

文章编号: 1000-7180(2013)11-0136-03

A Data Aggregation Algorithm for WSN Based on Clustering and Weighing-estimation

TAO Zhi-yong, WANG Xue

(School of Electrics and Information Engineering, Liaoning Technical University, Huludao 125105, China)

Abstract: Regarding to the situation that the wireless sensor networks have limited energy and require energy-saving, an optimally scaled clustering algorithm for network have been proposed in this paper. Moreover taking the heavier burden of cluster heads into account, the node that has higher residual energy is chosen as the cluster head, which can effectively extend the lifecycle and improve the efficiency of the use of WSN. Finally, with the way of coordinating the adaptive weighing-estimation data aggregation algorithm measurement variance of to integrate the collected data, reduce the data transmission, further to achieve the effect of energy saving.

Key words: wireless sensor network; clustering algorithm; weighing-estimation; data aggregation

1 引言

无线传感器网络(Wireless Sensor Network, WSN)中的数据融合技术, 是指在各个节点将检测数据通过多跳路由方式发送给收集数据的汇聚节点(或称基站)的路径过程中, 需要对数据进行处理以减少冗余信息^[1]。实验表明, 传输 1bit 信息 100m 需要消耗的能量大约相当于执行 3 000 条计算机指令消耗的能量, 也就是说, 传感器节点处理数据所需要消耗的能量要远远小于节点间无线通信传输所消耗的能量。因此, 虽然应用数据融合技术增加了中间节点的计算量, 但却减少了数据传输而节省了能量。将多个传感器的数据进行融合也可以增加测量结果的可信度, 提高系统的可靠度, 因此, 数据融合技术在

工业和军事等方面得到了广泛应用^[2]。

基于分簇算法构造和维护的无线传感器网络结构, 可以减少路由算法和洪泛广播的开销, 方便管理的节点和控制信道的接入, 也可以有效提高网络使用寿命和资源的使用效率。

本文提出了将最优分簇规模算法与自适应加权数据融合技术相结合的方法, 使无线传感器网络从最初的分簇开始到最终实现对数据的融合, 都实时考虑网络的能耗问题, 最终达到节约能量的目的。

2 WSN 中最优分簇规模计算

由于传感器网络逻辑结构各异, 其路由协议也被分为平面路由协议与分簇路由协议两种。前者里各传感器节点的地位相等, 网络内部也不会存在特

收稿日期: 2013-01-19; 修回日期: 2013-03-06

基金项目: 辽宁省教育厅科学基金项目(L2010167)

殊点,所有节点通过互相协作方式共同完成任务.而在分簇路由协议里,整个传感网络经常会被划分成一个簇首(Cluster Head)与数个簇成员节点(Cluster Member)组成的簇,其中,簇首要负责协调本簇内其他簇成员节点的工作,融合并向基站转发收集到的数据^[3].

假设在一个设定的监测范围内,部署着总数为 N 的同种类型的传感器节点,并设共会被分成 K 个簇.在已成型的簇内,各成员节点首先会通过消耗其自身的少部分能量来发送包括其剩余能量、编号以及地理位置等自身信息到其簇首,然后才开始转发采集到的数据信息.因为大多数的网络都应用统一的数据分组方式来完成数据的转发,那么各节点的编号以及位置等方面的消息就基本不变,自身信息的传输也只需要一次,进而能量的消耗也就有限.因而可以不必计较,文中提到的能量消耗只考虑用来传输数据所消耗的能量.根据一阶的能量消耗模型^[4],节点发送数据消耗的能量与距离成指数关系,即

$$E_{TX}(l, d) = \begin{cases} lE_{elec} + l\epsilon_{fs}d^2, & d < d_0 \\ lE_{elec} + l\epsilon_{mp}d^4, & d \geq d_0 \end{cases} \quad (1)$$

其中, $d_0 = (\epsilon_{fs}/\epsilon_{mp})^{1/2}$ 为发送节点和接收节点距离的阈值, $l, \epsilon_{fs}, \epsilon_{mp}, E_{elec}$ 分别为每个簇成员一次向簇首发送数据的比特数、空间无线数据传播常数以及收发消息的能量消耗常数.

接收节点的能量消耗:

$$E_{RX}(l) = lE_{elec} \quad (2)$$

簇首对簇内 n 个簇成员发来的数据进行数据融合的能量消耗为 $E(l, n) = nlE_{DA}$, 其中 E_{DA} 为簇首对每单位数据进行融合需要消耗的能量.

上述参数通常的典型值为 $E_{elec} = 50\text{nJ/bit}$, $\epsilon_{fs} = 10\text{pJ/bit/m}^2$, $\epsilon_{mp} = 0.0013\text{pJ/bit/m}^4$, $E_{DA} = 5\text{nJ/bit/signal}$.

假设每个簇的大小基本相同,簇内成员节点与簇首节点距离 d_{toCH} 较近,小于 d_0 ,簇首节点与基站距离 d_{toBS} 较远,大于 d_0 ;簇内每个成员节点采集到的数据包长度为 l ,簇首将这些数据经融合处理后,转发给基站的仍是一个长度为 l 的数据包.则簇内每个簇成员节点的能耗为:

$$E_{member} = lE_{elec} + l\epsilon_{fs}d_{toCH}^2 \quad (3)$$

簇首消耗的能量计算如下:①簇首接收簇内各成员节点需要消耗能量

$$E_{CHRX} = lE_{elec} \left(\frac{N}{K} - 1 \right) \quad (4)$$

②簇首对消息进行数据融合消耗能量:

$$E_{CHD} = lE_{DA} \left(\frac{N}{K} - 1 \right) \quad (5)$$

③簇首在对消息简单融合后,传给汇聚节点需要消耗的能量

$$E_{CHTX} = lE_{elec} + l\epsilon_{mp}d_{toBS}^4 \quad (6)$$

故簇头节点的能耗为:

$$E_{CH} = E_{CHRX} + E_{CHD} + E_{CHTX} \quad (7)$$

综上:整个网络完成一次数据采集的能耗为

$$E_{sum} = KE_{CH} + (N - K)E_{member} \quad (8)$$

上式中,为使 E_{sum} 最小,对变量 K 求偏导数即

$$\frac{\partial E_{sum}}{\partial K} = 0, \text{ 求出 } K \text{ 值即为最优分簇规模 } K_{opt}.$$

3 簇的形成与簇首选取

根据无线传感器网络本身的拓扑结构性性质,可使网络在完成一次分簇后就固定不变,让簇首轮流变换来降低多次成簇的能耗.由式(7)可知,簇首在整个网络中耗能最多.所以,在选择簇首时,要根据每个簇内各个节点的剩余能量的高低来决定出最终的簇首,以保障网络的生命周期^[5].

簇首选举之前,要首先运行一次 LEACH- F 算法或者 LEACH 算法,来得到一个具有最佳分簇规模的成簇网络.簇首选举时,可以令候选簇首记为 S_i ,其竞争半径记为 R ,大小由节点 S_i 到汇聚节点的距离决定.在竞争过程中,一旦 S_i 宣布竞争胜利,则其半径为 R 范围内的所有其他候选簇首均需退出竞争,网络成功建立.

簇首具体竞争步骤如下:①预设一个成员节点能够成为候选簇首的概率为阈值 T ,根据 T 来确定参与竞选的候选簇首,不能够参与竞选的成员节点直到竞选结束前需进入休眠状态.②参与竞选的候选簇首需广播其自身消息,包括剩余能量 E_i ,节点 ID 以及竞争半径 R .③每个候选簇首需要建立一个临时的簇首集合 S_{CH} ,用来收集其余候选簇首的广播消息,并与收集到的信息进行对比.剩余能量最大的候选簇首首先决定成为簇首,若发现剩余能量相等的情况,则选取节点 ID 距离汇聚节点较近的节点成为簇首.④簇首竞争一旦成功,便以广播的方式通知获胜消息,此时,要求其他候选簇首立即退出竞选,普通的成员节点也要从睡眠状态下唤醒,加入到相应的簇中去.

以上的簇首选举方式,与 LEACH 算法相比,避免了始终利用单一算法决定各节点角色而不考虑

节点剩余能量,造成网络内部负载不均衡的问题。

4 自适应加权数据融合算法

分簇结束并成功选举出簇首以后,簇首将对簇内每个节点收集的数据进行融合来减少数据量的传输。因为每个传感器的精度和可信度不可能完全一样,故可为每个传感器得到的测量值自适应地寻找与其对应的权数,使融合结果达到最优,即所谓的自适应加权数据融合算法^[6]。自适应加权数据融合算法的模型如图 1 所示。

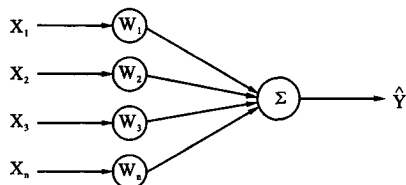


图 1 自适应加权融合算法模型

假设所需估计的真值为 Y , 各个成员节点采集到的数据的测量值分别为 $X_1, X_2 \cdots X_n$, 它们彼此之间互相独立并且是实际真值 X 的无偏估计量, 各传感器采集数据的加权因子分别为 $W_1, W_2 \cdots W_n$ 。设 n 个传感器采集数据的方差分别为 $\sigma_1^2, \sigma_2^2 \cdots \sigma_n^2$ 。

由图 1 可得数据融合结果为:

$$\hat{Y} = \sum_{i=1}^n W_i X_i \quad (9)$$

其中 W_i 为第 i 个传感器的加权因子, 且有:

$$\sum_{i=1}^n W_i = 1 \quad (10)$$

此时总的均方误差为:

$$\sigma^2 = E\left[\sum_{i=1}^n (X - X_i)^2 W_i^2\right] = \sum_{i=1}^n W_i^2 \sigma_i^2 \quad (11)$$

根据多元函数求极值的相关理论, 可以求出总均方误差最小时, 加权因子为:

$$W_i = \frac{1}{\sigma_i^2 \sum_{i=1}^n \frac{1}{\sigma_i^2}} \quad (12)$$

此时最小均方误差 σ^2 有最小值且为:

$$\sigma_{\min}^2 = \frac{1}{\sum_{i=1}^n \frac{1}{\sigma_i^2}} \quad (13)$$

由以上结果可知各传感器的权系数仅由其测量方差决定。只要使各个传感器的权数符合式(12), 就可以使总均方误差极小, 且此时有:

$$\hat{Y} = \frac{\sum_{i=1}^n \frac{X_i}{\sigma_i^2}}{\sum_{i=1}^n \frac{1}{\sigma_i^2}} \quad (14)$$

5 仿真实验

仿真实验模型由 100 个节点组成, 它们随机分布在 $100\text{m} \times 100\text{m}$ 的区域内, 每个节点的初始电池能量为 2J。考虑到系统中, 如果剩余节点总数低于整个网络节点总数的 30%, 那么网络的性能将会受到明显的影响, 就会失去研究的意义, 故本文考虑的网络生命周期为剩余节点总数大于 30% 的情况。从图 2 的仿真结果可以看出, 本文的分簇和簇首选举方法, 较单纯使用 LEACH 分簇算法相比, 明显地延长了系统的生命周期, 体现出了良好的能耗均衡性。

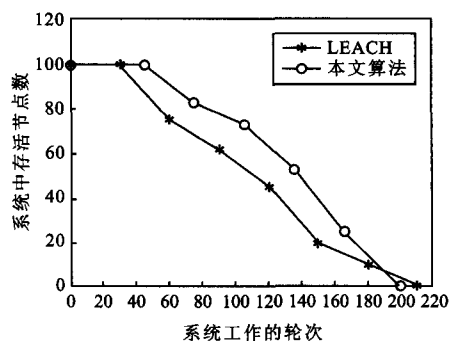


图 2 系统生命的比较

6 结束语

本文从无线传感器网络最初的网络分簇开始, 将网络分成最优规模形式, 利用选择剩余能量最多的节点作为簇首进行数据转发, 有效地延长了网络的生命周期。最后利用自适应加权数据融合算法将测量数据有效融合, 进一步减少了网络数据的传输, 达到了充分节能的目的。

参考文献:

- [1] 毕艳忠, 孙利民. 传感器网络中的数据融合[J]. 计算机科学, 2004, 31(7): 101-103, 117.
- [2] 许毅. 无线传感器网络原理及方法[M]. 北京: 清华大学出版社, 2012.
- [3] 彭爱平, 郭晓松, 蔡伟, 等. 基于估计机制的分簇传感器网络数据融合算法[J]. 传感技术学报, 2011, 24(1): 128-133.
- [4] Heitzelman w/ Chandrakasan A, Balakrishnan H. An applicationspecific protocol architecture for wireless microsensor networks[J]. IEEE Transactions on Wireless Communications, 2002, 1(4): 660-670.

(下转第 142 页)

Placement 阶段限制使用部分驱动能力弱的标准单元,并分析 congestion 问题,采用局部多次迭代的优化策略,消除 congestion. 布线阶段使用时钟连线优先的方法,并且设置串扰约束参数 Xtalk prevention threshold 为 0.3. 通过以上的预防措施,可以在布线后得到收敛的时序结果. 实验对比结果如表 2 所示.

表 2 布线后的时序结果对比

时序分析结果	Setup Case			Hold Case		
	WNS	TNS	NO. Viol	WNS	TNS	NO. Viol
串扰预防	-0.37	-7.36	92	-0.02	-0.16	8
普通方法	-1.01	-464.44	1032	-0.05	-19.40	121

通过采用二次综合、区域约束和串扰预防等措施可以保证复杂设计在 CMOS 深亚微米级制造工艺下实现渐进式的时序收敛. 在 65nm 工艺下,采用该方法完成了 530 万门的多核 DSP 芯片设计,保证了设计顺利交付 MPW 流片.

4 结束语

在芯片的设计中,后端物理设计作为重要环节已经形成了相对成熟的流程. 然而随着工艺的发展和设计规模的增大,后端设计变得越发复杂. 如何在有限的时间内完成所有设计流程,且保证时序收敛满足设计性能要求成为巨大的挑战. 所以针对不同的设计要求,需要采用灵活的设计方法来完善流程,

提高设计效率. 文中通过多核 DSP 芯片的后端设计,在层次化设计的基础上引入渐进式的收敛方法,通过采用二次综合、区域约束、串扰预防等措施实现阶段性的时序收敛,最终达到芯片设计收敛的目的.

参考文献:

[1] 谈晓婷,付宇卓,谢凯年. SoC 静态时序分析中时序约束策略的研究及实例[J]. 微电子学与计算机,2006,23(4):66-69.

[2] Shashank Prasad, Dongzi Liu, Oleg Levitsky, et al. Post assembly timing closure for multi million gate chips[C]//Proceedings of the 23rd IEEE International Conference on VLSI Design (VLSID'10), Bangalore, India; IEEE,2010.

[3] Jan M Rabaey. 数字集成电路设计透视[M]. 北京:清华大学出版社,2004.

[4] Synopsys Inc. IC compiler implementation user guide Version D-2010.03-SP2[R]. Synopsys Inc, 2010.

[5] Synopsys Inc. DC compiler user guide Version D-2010.03-SP2[R]. Synopsys Inc,2012.

作者简介:

郑天华 男,(1987-),硕士研究生. 研究方向为集成电路物理设计.

梁利平 男,(1969-),研究员. 研究方向为高性能多核 DSP 及 SOC 设计.

(上接第 138 页)

[5] 孙凌逸,黄先祥,蔡伟,等. 基于神经网络的无线传感器网络数据融合算法[J]. 传感技术学报,2011,24(1):122-127.

[6] 廖惜春,丘敏,麦汉荣. 基于参数估计的数据融合算法研究[J]. 传感器与微系统,2006,25(10):70-73.

作者简介:

陶志勇 男,(1978-),博士研究生,副教授. 研究方向为多媒体通信.

王雪 女,(1988-),硕士研究生. 研究方向为无线传感器网络.