

低能耗的分布式数据融合改进算法

邓亚平, 牛 康

(重庆邮电大学 计算机科学与技术学院, 重庆 400065)

(niukang285@yahoo.com.cn)

摘 要:通过对串行数据融合算法的研究,发现在实际情况中,要使所有的传感器节点保存全局路由信息非常困难。为降低额外的能量消耗,提出一种新的分布式数据融合算法。该算法中,多个传感器节点同时启动数据融合过程,以多路方式向簇头传递信息,所有节点都是自行决定下一跳的传输方向,避免了保存全局路由信息。从仿真结果看,该算法可以减少时延,有效降低能量消耗。

关键词:无线传感器网络; 数据融合; 分布式算法; 能量消耗

中图分类号: TP393 **文献标志码:** A

Improved distributed algorithm of reducing energy consumption in data aggregation

DENG Ya-ping, NIU Kang

(College of Computer Science and Technology, Chongqing University of Posts and Telecommunications, Chongqing 400065, China)

Abstract: Through studying the algorithm of serial data aggregation, it is found that it's difficult in a real scenario to save the global routing knowledge for all sensors. In order to reduce the additional energy consumption, the authors proposed a new distributed algorithm to aggregate data. Multiple sensor nodes initiate the process of data aggregation simultaneously, transmitting information to cluster head through multi-channel approach. All sensors choose the next transmission direction by themselves, avoiding saving the global knowledge. Simulation results show that this algorithm can reduce the transmission delay and energy consumption effectively.

Key words: Wireless Sensor Network (WSN); data aggregation; distributed algorithm; energy consumption

0 引言

微电子技术、计算机技术和无线通信等技术的进步,推动了低功耗多功能传感器的快速发展,使其在微小体积内能够集成信息采集、数据处理和无线通信等多种功能。多个传感器节点的集合构成了无线传感器网络,传感器节点之间以 Ad Hoc 的方式相互连接^[1]。传感器节点之间相互协作,采集并处理监测对象的相关数据,通过多跳路由将用户需要的信息发送到基站。无线传感器网络将虚拟的信息世界与客观的物理世界联系在一起,改变了人类与自然界的交互方式。鉴于单个传感器节点的监测范围和可靠性是有限的,在部署网络时,需要使传感器节点达到一定的密度以增强整个网络的鲁棒性和监测信息的准确性,有时甚至需要使多个节点的监测范围互相交叠。这种监测区域的相互重叠导致邻近节点报告的信息存在一定程度的冗余。比如对于监测温度的传感器网络,每个位置的温度可能会有多个传感器节点进行监测,这些节点所报告的温度数据会非常接近或完全相同。在这种冗余程度很高的情况下,把这些节点报告的数据全部发送给汇聚点与仅发送一份数据相比,除了使网络消耗更多的能量外,汇聚节点并未获得更多的信息^[2-3]。

数据融合是无线传感器网络中重要的研究领域之一。使用数据融合技术能够有效地解决无线传感器网络中能量的限制。通过合并多个数据源产生的数据,去除冗余信息,数据融合能够有效地减少网络中的数据传输量,从而节省传感器节

点的能量,延长无线传感器网络的生命期。目前已有大量的研究工作从不同角度来力求延长传感器网络的寿命。

1 相关工作

文献[4]提出了 LEACH 算法。它是一种分布式自组织的协议,其核心思想是减少与基站直接通信的节点数量来达到节能的目的。LEACH 协议按轮运行,每轮以自组织的方式随机选出部分传感器节点作为簇头,簇中的成员节点把收集的数据传送给簇头,由簇头进行融合并将融合后的数据发送给基站。当有些簇非常大时,簇头的能量消耗就相当快。

文献[5]提出了 PEGASIS 协议。该协议的具体方法是,把系统中所有传感器节点用贪婪算法构成一个边长之和接近最小的链,任意节点从位于链上的一个相邻节点接收数据并把接收到的数据与本节点的数据融合再转发给另一相邻节点,最后由指定节点把最后融合的数据发送到基站。与 LEACH 协议相比,PEGASIS 协议与基站直接通信的节点更少,从而显著减少了每一轮的能量损耗。然而,PEGASIS 算法要求所有节点知道网络全局信息,而保存网络的全局信息对于单个节点非常困难。另外,如果链上的某一节点死亡则使得从链端到该节点的所有数据丢失,因此 PEGASIS 的容错性不佳。

文献[6]提出基于空间填充曲线的融合方法,利用空间填充曲线通过传感器网络建立了一个路由路径来进行串行数据融合。这个算法保证了每个传感器节点从一个邻居节点接

收稿日期:2009-01-07;修回日期:2009-03-02。

作者简介:邓亚平(1948-),男,重庆人,教授,主要研究方向:计算机网络与通信、信息安全;牛康(1985-),男,陕西宝鸡人,硕士研究生,主要研究方向:计算机网络与通信。

收或者向一个邻居节点仅发送一次数据。另外,路由的生成和数据融合是在同一个过程中完成的,因此就减少了能量的消耗。然而,由于只有一个节点来启动算法,它要求某种形式的全局同步,这也造成了额外的能量消耗。

文献[7]对 LEACH 算法做了改进,提出了两种不同的协议:能量 LEACH 协议和多跳 LEACH 协议,分别提高了能量利用率和传输效率。文献[8-9]中节点的寿命比 LEACH 和 PEGASIS 中更长。

以上算法表明要实现低能耗和最大化传感器节点的寿命,需要满足以下条件:

1) 相对平均的能量消耗:所有节点成为融合节点的次数应该大体相等。

2) 最小的运行次数:节点最多只进行接收、融合、发送信息三个过程。

本文提出一种分布式的算法来收集各个传感器节点的数据。这个算法的实现过程中,节点不需要知道系统的全局信息,这样可以降低系统中节点总的能量消耗。算法的运行允许多个传感器节点启动数据融合过程,数据融合就以并行的方式进行,因而相比文献[6]减少了延迟。算法动态地选择不同的节点作为簇头,这就保证了能量的消耗是动态地分布在系统中所有节点上的。

2 算法模型

2.1 算法的描述

在我们的方案中,每个传感器节点最多有四个邻居节点,也就是可以有四个方向(up, down, left, right)进行传输。用 up_i , $down_i$, $left_i$, $right_i$ 局部变量来表示传感器 i 在不同方向的邻居节点。这些变量包括了各个方向离 i 最近的节点的身份信息,变量的选择可以在传感器的部署阶段进行。节点之间可以相互通信,也可以与基站进行通信,任何一个节点都有可能成为融合节点或者簇头,要向基站传输数据信息。传感器节点是均匀分布在给定区域内的。假设每个节点都配备指令天线用来区分不同的方向,指令天线的使用可以使传感器节点能够选择自己想要的传输方向。

数据融合过程由基站来启动,基站发送一个信标信号,表示新一轮的数据收集过程开始。然后,依据本地知识,每个节点自己决定是否开始消息传送,也就是说可以有多个传感器节点同时来启动算法,并行地遍历所有节点。基站不要求掌握全局信息,因为各个节点都是自己做出决定。

当新一轮数据融合开始后,需要选择多个初始节点启动算法。因传感器网络的结构在不停的变化之中,本轮中离基站近的节点在下一轮可能是离基站较远的,所以每轮中离基站最远的节点是不同的。每一轮中的初始节点可以选择离基站最远的多个节点,这样就保证了每次选择的初始节点是不同的节点,而且最后成为簇头的节点是离基站很近的,减少了额外的能量消耗。

每个节点分别选择自己的邻居节点(up_i , $down_i$, $left_i$, $right_i$)进行数据传输,它们选择下一跳邻居节点需遵循以下原则:

- 1) 如是初始节点,则先上下,后左右;
- 2) 如果既有上方节点又有下方节点,则随机选择。左右也相同;
- 3) 如果不是初始节点,则此节点选择来源节点方向的反

方向进行传输;

4) 直到遇到一个已经发送过自己信息的节点为止;

5) 每一条路径到达的最后一个节点成为簇头,由簇头向基站转发本簇中所有节点总的融合信息;

6) 每一个节点在启动算法时确定要转发的方向。

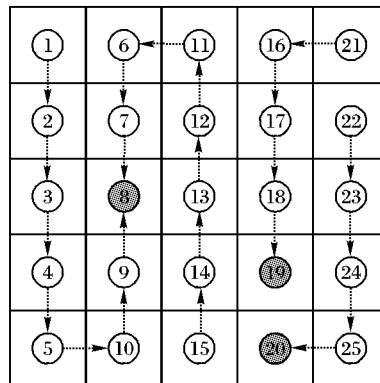


图1 分布式融合方案

在图1中,节点1、15、21、22作为初始节点来启动算法。节点1有两个邻居节点,根据规则,节点1选择下方节点2转发数据,由于节点2是从上方节点1处接收的数据,它将1转发的数据和自己收集的数据进行融合后,转发给下方节点3,以此类推到达节点10,它有三个邻居节点,因节点5和15均已发送过信息,因此,节点10将融合后的数据转发给节点9,9再转发到节点8,而另一条路径转发到了节点7。如果节点7接收到信息比节点8要早,那么节点7可以继续将数据转发给节点8,因为8还没有转发过数据。这样就保证了每一个节点向它的邻居节点最多发送一次信息,但却不能避免多次收到信息。最后节点8、19、20成为簇头,由它们向基站发送信息。

该算法的主要特征是:1) 传感器节点只根据本地知识进行操作,并且系统中没有时钟同步;2) 多个传感器节点启动算法;3) 在给定的区域内节点是均匀分布的;4) 节点之间可以相互通信,也可以和基站进行通信。

2.2 能量消耗模型

利用文献[4]中描述的能量消耗模型, d_i 表示一个簇头到基站之间的距离,发送消息的长度是 l 比特,将 l 比特消息传输 d_i 距离所消耗的能量为 E_T :

$$E_T = lE_e + l\varepsilon_l d_i^4 \quad (1)$$

此外,接收 l 比特消息消耗的能量为 E_R :

$$E_R = lE_e + lE_{BF} \quad (2)$$

式(2)包括波束形成方法的开销 E_{BF} ,这种方法可以减少能量的消耗。以上公式中能量消耗值如下:1) 短距离传输的能量消耗: $\varepsilon_l = 10 \text{ pJ/bit/m}^2$; 2) 长距离传输的能量消耗: $\varepsilon_l = 0.0013 \text{ pJ/bit/m}^4$; 3) 传输或接受信号的电路所消耗的能量: $E_e = 50 \text{ nJ/bit}$; 4) 波束形成消耗的能量: $E_{BF} = 5 \text{ nJ/bit}$ 。

2.3 算法的初始化

初始化要求每个节点选取一组参数,这些参数在算法运行过程中会用到,不需要建立路由表。每个节点将按照以下步骤进行初始化:

- 1) 为方向建立参数值;
- 2) 建立传输消息到基站的能量消耗;
- 3) 为初始发送速率赋值;

4) 选取开始一轮的随机等待时间;

5) 初始化所有局部变量。

当等待时间结束,新一轮数据融合过程还未开始时,节点开始启动算法。每个节点选择下一个目的节点传输信息,并在本地保存此目的节点。

3 仿真及其分析

假设节点均匀地分布在一个 XY 坐标内,共有 4×10 个节点,坐标原点在左下角的节点处,此节点标为第 0 个节点,沿着 X 坐标依次标为第 10 个节点,第 20 个节点...,一直到第 30 个节点;然后 Y 轴上的第 2 个坐标标为第 1 个节点,沿着 X 轴方向依次为第 11 个节点,第 21 个节点...,一直到第 31 个节点; Y 轴上第 3 个节点标为第 2 个节点,以次类推,一直到第 39 个节点。仿真环境采用的是 NS2。

仿真实验中, 4×10 个节点均带有 1 个单位的电量,数据包发送速率为 100。

图 2 给出了节点在数据融合过程中能量消耗的情况。图中的三种柱状分别是节点在不同时刻的剩余能量,能量值分别在 $0.7 \sim 1$, $0.3 \sim 0.7$, $0 \sim 0.3$ 。从图中可以看到,在 30 s 之前节点能量消耗不大,30 s 之后节点进行数据融合及数据传输过程,能量消耗增加,100 s 左右大多数节点 1 个单位的能量值基本耗尽。

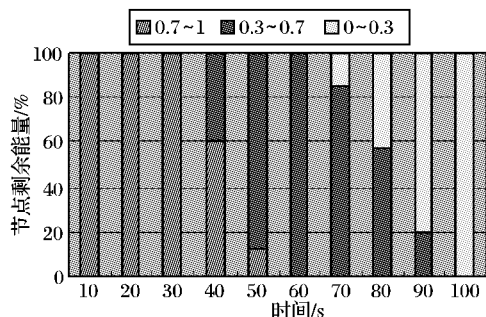


图2 数据融合过程中的能量消耗

在数据融合过程中,所有节点成为融合节点或簇头的几率差不多,能量消耗比较平均。将轮数和初始能量定的较小得情况下才能够看到比较确切的不同节点的能量消耗值。表 1 给出了不同时刻带有不同电量值的节点占总节点数的百分比。

表 1 不同电量节点占总节点百分比

电量值	时间/s										%
	10	20	30	40	50	60	70	80	90	100	
0.7 ~ 1	100	100	100	60	12	0	0	0	0	0	
0.3 ~ 0.7	0	0	0	40	88	100	85	57	20	0	
0 ~ 0.3	0	0	0	0	0	0	15	43	80	100	

图 3 给出的是本文提出的算法和基于空间填充曲线以及 LEACH 算法中节点平均能耗的比较。从图中可以看到, LEACH 算法中节点的平均能量消耗较快,在大约 57 s 时节点能量已耗尽,节点死亡。基于空间填充曲线的融合算法明显要好于 LEACH 算法。本文提出的改进算法比 LEACH 和基于空间填充曲线的融合算法效果要更好一些,节点的寿命最长。图中,在大约 30 s 之前,改进算法与空间填充曲线算法节点的能耗基本相同。在 35 s 之后,可以看到改进算法效果更好,整体的节点能耗更小,节点的生存期更长。

4 结语

数据融合技术可以减少数据传输量,节省节点的能量,延长节点的生命周期。本文在原有数据融合方法的基础上提出一种新的分布式数据融合算法,在算法中节点不需要知道全局信息,所有节点自行作出决定,并行地向基站传输数据,从而降低了额外的能量消耗,减少了整个网络的延迟。由于能量消耗是均匀地分布在网内所有节点上的,因此就延长了网络的生命周期。在数据融合过程中,如何提高它的容错性是进一步研究的工作。

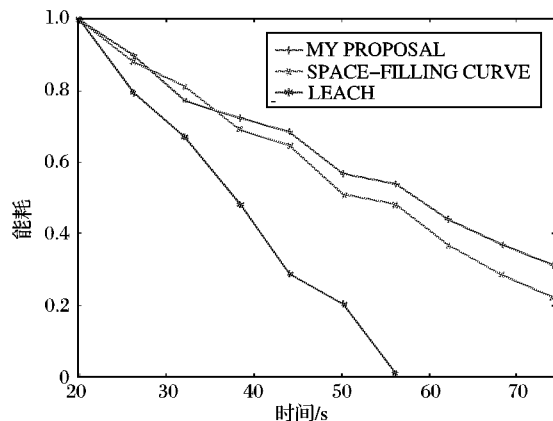


图3 不同算法的节点能耗比较

参考文献:

- [1] 孙利民, 李建中, 陈渝, 等. 无线传感器网络[M]. 北京: 清华大学出版社, 2005.
- [2] AKYILDIZ L F, SU WEILIAN, SANKARASUBRAMANIAM Y. A survey on sensor networks[J]. IEEE Communications Magazine, 2002, 40(8): 102 - 114.
- [3] ELENA F, MICHELE R, JORG W, *et al.* In-network aggregation techniques for wireless sensor networks: A survey[J]. IEEE Wireless Communication, 2007, 14 (2): 70 - 87.
- [4] WENDI H, ANANTHA C, HARI B. Energy-efficient communication protocol for wireless microsensor networks[C]// Proceedings of the 33rd Hawaii International Conference on System Sciences. Washington, DC: IEEE Computer Society, 2000: 8020 - 8030.
- [5] STEPHANIE L, CAULIGI S. PEGASIS: Power-efficient gathering in sensor information systems[EB/OL]. [2008 - 09 - 20]. <http://ceng.usc.edu/~raghu/pegasisrev.pdf>.
- [6] SWAPNIL P, SAMIR R. Serial data fusion using space - filling curves in wireless sensor networks[EB/OL]. [2008 - 09 - 20]. <http://www.cs.sunysb.edu/~samir/Pubs/serial.pdf>.
- [7] FAN XIANGNING, SONG YULIN. Improvement on LEACH protocol of wireless sensor network[C]// Proceedings of the 2007 International Conference on Sensor Technologies and Applications. New York: IEEE, 2007: 260 - 264.
- [8] GOYENECHE M, VILLADANGOS J, ASTRAIN J J, *et al.* A distributed data gathering algorithm for wireless sensor networks with uniform architecture[C]// Proceedings of the 15th Euro micro International Conference on Parallel, Distributed and Network-Based Processing. Washington, DC: IEEE Computer Society, 2007: 373 - 380.
- [9] TAN H O, KORPEOGLU. Power efficient data gathering and aggregation in wireless sensor networks[J]. ACM SIGMOD Record, 2003, 32(4): 66 - 71.