

高效无线传感器网络分簇数据收集协议

常 波¹, 徐保国²

(1 淮阴工学院 计算机工程学院, 江苏 淮安 223003; 2 江南大学 信控学院, 江苏 无锡 214122)

摘 要: 提出了一种基于簇结构的数据收集协议 ECDGP(Energy-efficient Cluster-based Data Gathering Protocol)。ECDGP 使用了一种基于置信度的分簇算法进行簇头竞选, 并根据网路覆盖要求选择活动节点, 通过控制分簇中活动节点数目, ECDGP 减少了能耗和延长了网络生命期。仿真结果表明在能耗和网络寿命上, ECDGP 大幅度优于传统的分簇路由算法。

关键词: 数据收集; 高效; 分簇算法; 置信度

中图分类号: TP393

文献标识码: A

文章编号: 1000-7180(2011)05-0075-04

Energy-efficient Cluster-based Data Gathering Protocol for Wireless Sensor Networks

CHANG Bo¹, XU Bao-guo²

(1 Faculty of Computer Engineering, Huaiyin Institute of Technology, Huai'an 223003, China;

2 School of Communication and Control Engineering, Jiangnan University, Wuxi 214122, China)

Abstract: In this paper, an energy-efficient cluster-based data gathering protocol (ECDGP) is proposed. ECDGP uses a new clustering algorithm based on belief degree for clusterhead election. Furthermore, ECDGP selects the active nodes according to the demand of network coverage. By controlling the number of active nodes in a cluster, ECDGP reduces energy consumption and prolongs the network lifetime. Simulation results have shown that ECDGP significantly outperforms the traditional clustering routing algorithms in terms of energy consumption and network lifetime.

Key words: data gathering; energy-efficient; clustering algorithm; belief degree

1 引言

无线传感器网络(Wireless Sensor Networks, WSN)是由大量微型传感器节点,通过无线通信方式形成的一个自组织网络系统。传感器节点通常采用电池供电,而能量主要消耗在通信过程中,故在 WSN 的应用中节约能源是个关键问题,必须设计能量高效的数据收集算法以尽量延长网络寿命。

数据收集协议依据网络底层的拓扑结构分为平面型数据收集和层次型数据收集(特别是簇结构)。与平面型相比,分簇数据收集具有拓扑管理方便、扩展性强、能量利用率高、数据收集效率高、负载平衡

性好、数据融合简单等优点^[1-2]。

由于分簇数据收集可以有效减少数据延迟和增强网络的可扩展性,目前已经有很多学者对其进行了广泛的研究。低能耗自适应成簇算法 LEACH (Low-Energy Adaptive Clustering Hierarchy)^[3]协议是一种典型的基于簇的数据收集协议,该协议采用动态分簇算法平衡网络能量消耗、延长网络生命周期。为了提高 LEACH 的性能, Lindsey 等人在 LEACH 协议基础上提出了 PEGASIS(Power-Efficient Gathering in Sensor Information Systems)协议^[4],该协议通过算法将网络中所有节点构成一条链,并在链上选择一个节点作为头节点,链两端数据

收稿日期: 2010-06-18; 修回日期: 2010-09-01

基金项目: 国家“八六三”计划项目(2007AA10Z241)

沿链传输到头结点,头节点与基站直接通信.由于节点采用小功率与最近邻居节点通信以及通信量的减少,PEGASIS比LEACH协议更能延长网络的生命周期.然而长链引起的传输延时限制了它在大规模网络中的应用.在文献[5]中,提出了使用K-means技术的平衡并行聚集协议,该协议能在节点之间平坦地分配能量损耗,因此延长了系统的生命期.但算法要求每个节点都能获得自身的位置信息,故算法的可测量性不好.在文献[6]中,提出了对门限敏感的高能效传感器网络协议TEEN(Threshold sensitive Energy Efficient sensor Network protocol),该协议使用软门限和硬门限来减少传送数目,故能节约能耗.但是,如果CHs不在彼此的传输范围内,消息将会丢失.

文中提出的高能效分簇数据收集协议ECDGP(Energy-efficient Cluster-based Data Gathering Protocol)属于一种事件驱动协议,即事件区中的节点能被聚合成簇.在簇头CH(clusterhead)选举时段,根据相邻节点的置信度(由置信度函数计算),各节点竞争成为CH.此外为了使网络中的能量负载平衡,ECDGP使用多径方案从CH到汇聚节点传输数据.

2 ECDGP协议的设计

文中假设:

(1) N 个传感器节点随机均匀部署在一个二维方形区域内,所有传感器节点都被事先编排唯一的ID号,位于事件区中的节点组成一个簇.

(2) 所有节点不能获知其位置信息,节点密集部署.

(3) 各节点能够根据已接收的信号强度计算近似通信距离.

(4) 各节点都有功率调节装置,即无线发射功率可控,能够根据接收者的距离远近来调节发射功率的大小.

ECDGP协议工作时段可分为路由初始化、簇生成和数据传输三个时段.在簇生成时段,位于事件区的节点将被组成一个簇.在数据传输阶段,CH收集来自活动节点的数据并通过多径方法把它们传递到汇聚节点.

2.1 路由初始化

ECDGP协议的开始时段,一个带有汇聚节点信息的节点通过向邻居节点发送一个广播消息(ADV消息)开始路由的初始化,每个收到ADV消

息的节点再向邻居节点广播ADV消息.在网络初始化结束阶段,每个节点构造一个邻居节点的信息表,见表1.

表1 邻居节点信息表

ID	H_i	E_i	CH	d_{ij}	T_{ag}	q	S
1	H_1	E_1	not		0	q_1	0
2	H_2	E_2	not		0	q_2	0
...	...	...	...	...	...	...	...
n	H_n	E_n	not		0	q_n	0

表中,ID是邻居节点的标识索引; H_i 是从节点*i*到汇聚节点的跳数; E_i 是节点*i*的剩余能量; d_{ij} 是节点*i*和节点*j*之间的距离; T_{ag} 标识邻居节点是否位于事件区, $T_{ag} = 1$ 表示邻居节点位于事件区,否则 $T_{ag} = 0$; q_i 表示节点*i*成为CH的概率,S表示节点的状态, $S = 0$ 表示节点是不活跃的,否则 $S = 1$ 表示节点是活跃的.

2.2 簇生成时段

(1) 簇的生成

位于事件区的节点监测到一个事件,即开始了簇生成时段.每个能感知事件的节点将广播一个Hello消息,如果邻居节点收到这个消息且它们也位于这个事件区,就修改表1中的值(如 d_{ij}).Hello消息的格式包含三个字段:MessageType、ID、 E_i ,其中MessageType字段是Hello消息的标识,ID字段表示节点标识符, E_i 字段给出节点*i*的剩余能量.

为了估算节点*i*和*j*之间的能耗,使用文献[3]中的无线模型.在节点*i*和*j*之间(相距 d m)传输和接收一个 l 比特数据包所需的能耗可表示为

$$E_T(i, j) = l \times (E_{\text{elec}} + E_{\text{amp}} \times d^\gamma) \quad (1)$$

$$E_R = l \times E_{\text{elec}} \quad (2)$$

式中, E_{elec} , E_{amp} 和 E_R 分别是发送器、发送放大器和接收器的每比特能耗, $\gamma \in \{2, 4\}$ 是路径损耗指数.

定义1 置信度:表示一个节点成为CH的概率,其值由节点间的相对距离确定,即距离越近,来自邻居节点的置信度越高,置信度函数 $B(i)$ 由 $b_{ij} \in [0, 1)$ 给出, b_{ij} 与节点间的距离关系可表示为

$$b_{ij} = \begin{cases} \frac{2\text{arc cot}(d_{ij})}{\pi}, & \text{if } d_{ij} \leq R \\ 0, & \text{if } d_{ij} > R \end{cases} \quad (3)$$

式中, R 是节点的通信半径.节点*i*的整个置信度是其所有邻居节点置信度之和,即

$$B(i) = \sum_{(d_{ij} \leq R)} b_{ij} \quad (4)$$

在一个簇中,CH 要做更多的工作,因此剩余能量多的节点将更有机会成为 CH. 根据参考文献[7],当 CH 位于事件中心附近时,能耗最小,同时考虑到式(4),定义节点成为 CH 的目标函数为

$$q_i = (E_i/E_{\text{Avg}})^{k_1} \times B^{k_2}(i) \times (1/d_{i,C})^{k_3} \quad (5)$$

服从 $E_i \geq E_{\min}$ (5)

式中, E_i 是节点 i 的剩余能量, E_{Avg} 是节点 i 的邻居节点的平均能量, $d_{i,C}$ 是节点 i 与事件中心间的距离, k_1, k_2, k_3 分别是 $E_i, B(i)$ 和 $d_{i,C}$ 的权重系数, $E_{\min}$ 是成为 CH 的能量门限,由下式计算:

$$E_{\min} = a \times [N_A \times (E_{\text{elec}} + E_D) \times l + l \times E_{\text{amp}} \times d^2] \quad (6)$$

式中, a 是系数, N_A 是簇中的活跃节点数, E_D 是数据融合中每比特消耗的能量,式(6)中其他参数的含义与式(1)和式(2)相同.

图1表示簇生成算法的伪代码. K 表示存储 CH 广播数的临时变量; T_1 是等待时间,其值与网络规模有关; Threshold 与文献[5]中的值相同; RSS_i 是节点 i 感知的已接收到的来自事件区信号的强度; T_i 是节点竞争成为 CH 而延迟的时间,由下式给出:

$$T_i = p \times T/q_i \quad (7)$$

式中, p 是控制延时的时间常数, T 是初始化等待时间.

```

Begin
1:  schedule a node wait time with  $T_1$  sec.delay
2:  while( $T_1 \neq 0$ )
3:    if  $\text{Threshold} \leq \text{RSS}_i$  then
4:      if  $d(v_i, v_j) \leq R$  and  $\text{Threshold} \leq \text{RSS}_j$ 
5:        Use(3),(4) to calculate  $B(i)$ ;
6:      end-if
7:    end-if
8:  end-while
9:  If  $\text{Threshold} \leq \text{RSS}_i$  then
10:    node  $i$  use(5),(7) to calculate  $q_i$  and  $T_i$ ;
11:    if  $T_i \neq 0$  then
12:      wait;
13:    else
14:      collect the sender of any other incoming CH
15:      advertisement in  $K$ ;
16:      if  $K=0$  then
17:        send CH advertisement message;
18:      else
19:        if  $q_i$  is the biggest then
20:          send a join-request to the node  $j$ ;
21:        end-if
22:      end-if
23:    end-if
24:  end-if
End

```

图1 簇生成算法伪码

(2) 簇内活动节点的选择

WSN 中覆盖范围是一个基本要求,这已在文献[8-9]中进行了大量研究.覆盖范围是能被一个传感器节点检测到的区域.如果区域 S 中的位置由至

少 K 个节点覆盖,则区域 S 覆盖度为 K .多数情况下,检测区的完全覆盖是不必要的^[9].因此,采用部分覆盖并把簇覆盖的概念引入事件区.对给定的事件区域 S ,部署事件区中的节点集合. ECDGP 通过强制最大化休眠节点数目来减少能耗,保留的活动节点数目要能确保 S 可满足期望的覆盖度.但找到一个最小支配集是一个 NP 完全问题, ECDGP 使用下面的方法来解决这个问题.

设 N 是事件区中的所有节点集合, M 是覆盖集(当网络未初始化时 M 为空).首先,采用文献[10]中的算法得到活动节点数 m ,然后使用式(8)来选择覆盖集中的节点.

$$P_i = N_i^{k_4} \times E_i^{k_5} \quad (8)$$

式中, N_i 是节点 i 的邻居节点数, E_i 是节点 i 的剩余能量, k_4 和 k_5 是分别是用来控制 N_i 和 E_i 的权重系数.选择活动节点的算法可描述为:当 $N > m$ 时,循环计算 P_i 的值,若 P_i 最大,则设置节点 i 在 M 中,而把节点 i 和 N_i 从集合 N 中删除.通过循环计算比较,最后选择出 m 个活动节点.

2.3 数据收集和传输时段

数据收集和传输模型如图2所示.活动节点只向 CH 传输感知数据,由 CH 通过多径方式和多跳方式把收集的数据传输到汇聚节点.在数据传输期间,每个中间活动节点都会对已接收的数据进行延时,并且是使用多径方式进行延时.多径路由算法使用多条路径传输数据,这就带来负载平衡和容错问题.在 WSN 中,这种数据传输方式使总能量分布在各节点间,这对增加网络生命期来说很重要^[11].因此,使用文献[11]中提出的多径方案在 CH 和汇聚节点间传输数据.

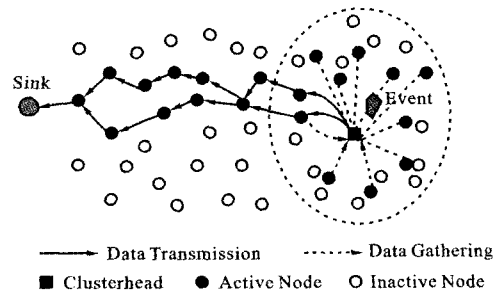


图2 数据收集和传输模型

2.4 CH 管理

由于 CH 需要消耗更多的能量来传输数据和进行数据聚合,所以对 CH 来说负载平衡机制很重要.为了将能量负载均匀地分配到各节点上, ECDGP 协议按轮运行,并在每一轮中对簇头进行轮换.在每

一轮中,当簇生成以后,簇头将聚合其收到的各成员节点的采集信息,并将聚合信息直接传输到汇聚节点.当 CH 的能耗低于 E_{Ava} (E_{Ava} 与簇大小有关)时,根据式(5)、式(6)和表 1,重新选择一个新的 CH.由于减少了与汇聚节点直接通信的节点数量以及通信量,ECDGP 协议可以有效地延长网络生命周期.

3 协议性能分析

对文中提出的 ECDGP 数据收集机制和现有的两类典型数据收集方法 LEACH 和 PEGASIS 进行比较.仿真场景设置为:网络节点均匀分布在 $200\text{ m} \times 200\text{ m}$ 的范围.分布在事件区中的节点(半径为 20 m)是信源.

仿真讨论的性能指标如下.

能耗:包括网络一个周期数据收集的总能耗以及单个周期节点最大能耗.

网络生命期:定义为网络从开始运行到第一个节点由于能量耗尽而死亡的持续时间.

仿真用到的其他参数如表 2 所示.

表 2 仿真参数

Parameters	Value
Number of nodes	500~1 000
Number of sink	1
Sensing radius	40m
Initial energy	2 Joules
Packet size	512 Bytes
E_{elec}	50 nJ
E_{amp}	100 pJ
E_D	5 nJ
Location of sink	(0,200)
Size of control packet	20 Bytes

文中对 LEACH 算法、PEGASIS 和 ECDGP 三种算法下的网络能耗和网络生命期进行了仿真,结果如图 3 和图 4 所示.

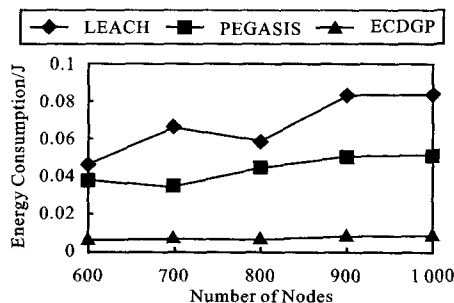


图 3 能耗的比较

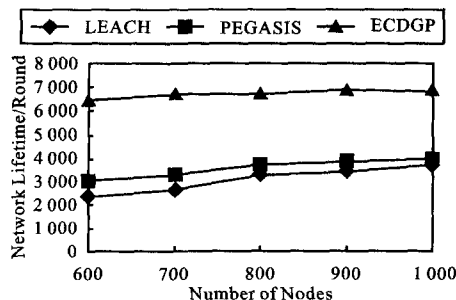


图 4 网络生命期的比较

由图 3 可见,与其它两种算法相比,ECDGP 的能耗最低,原因是在 ECDGP 协议中的分簇与事件有关,而且 ECDGP 能控制活动节点的数量,这对降低能耗非常有利.

由图 4 可见,三种算法实现的网络生命期都随网络规模的增加成线性增加.原因是参与数据传输的延迟节点的数量与网络中的总节点数成比例,即更多的节点参与到数据传输的路由上,能够平衡网络能耗从而延长网络生命期.而通过 ECDGP 算法得到的网络生命期几乎是 LEACH 或 C-PEGASIS 算法的两倍.

4 结束语

为了提高 WSN 中数据收集的能量效率,ECDGP 使用置信度和簇内覆盖,它集中了分级路由和多径路由的优点. ECDGP 还使用簇内聚集来减少通信量,使用多径方案来均匀分布网络中节点间的能耗.仿真结果表明,和传统算法相比,采用 ECDGP 方式提供了更低的能耗,更长的生命期.

参考文献:

- [1] 李明,曹建农,陈贵海,等. EADEEG: 能量感知的无线传感器网络数据收集协议[J]. 软件学报, 2007, 18(5): 1092~1109.
- [2] 刘韬,谢储晖. 无线传感器网络能量均衡路由算法[J]. 微电子学与计算机, 2009, 26(10): 5~7.
- [3] Heinzelman W R, Chandrakasan A, Balakrishnan H. Energy-efficient communication protocol for wireless microsensor networks[C]// Proc of the Hawaii International Conference on System Sciences, Washington, DC, USA: IEEE Computer Society, 2000: 3005~3014.
- [4] Lindsey S, Raghavendire CS. PEGASIS: Power-efficient gathering in sensor information systems[C]// Proceedings of the IEEE Aerospace Conference. [S. L.] IEEE Computer Society, 2002(3): 1125~1130.

(下转第 81 页)

户端程序 11 的返回结果以及其直接控制运行程序的返回结果,通过监控方 4 上的控制命令发送接口 8 返回结果给中心分析和控制 7。

这样,就实现了从信息的订购、采集、分析处理、故障发现到故障排除的自动化。当然,这只是一种实现方式,根据具体的应用环境还会有其它的实现方式。如不采用 FTP 协议传递文件等。

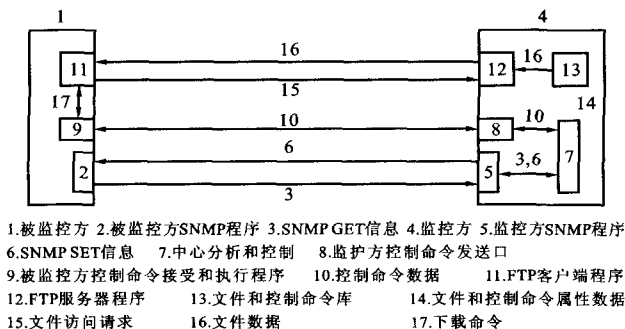


图3 具有自恢复能力的机群监控系统实现

4 结束语

文中提出的具有自恢复能力的机群监控系统结构是基于动态、主动监视和控制的设计思路,不同于传统的静态配置和管理的设计思想,可以极大提高机群监控系统的主动控制、精确控制和升级维护能力以及整个机群系统的可靠性。采用该结构的机群监控系统已在实际的 Linux 机群存储系统中得到应用,所开发的系统提供了良好的接口,允许用户动态

添加新的信息处理程序和故障处理程序,具有良好的可扩展性。

参考文献:

- [1] 武林平,孟丹,詹剑锋,等. 机群系统可靠性研究的新思路—主动故障检测机制研究[EB/OL]. (2007)[2010-08-13]. <http://www.docin.com/p-16397101.html>.
- [2] Treaster M. A survey of fault-tolerance and fault-recovery techniques in parallel systems. cs DC/0501002 [R]. CORR; National Center for Supercomputing Applications of University of Illinois, 2005.
- [3] 张洋,陈文波,李廉. 基于 SCMSWEB 的网络资源监控研究与实现[J]. 微电子学与计算机, 2008, 25(10): 166—168.
- [4] 李明江. SNMP 简单网络管理协议[M]. 北京:电子工业出版社, 2007.
- [5] 李娟,毛宏燕,马光思. SNMP v3 数据采集器 DCP 的研究与实现[J]. 微电子学与计算机, 2008, 25(12): 137—140.
- [6] Sun Jimeng, Hoke Evan, Strunk John D, et al. Intelligent system monitoring on large clusters[C]// Proceedings of the 3rd Workshop on Data Management for Sensor Networks; in Conjunction with VLDB 2006. NY, USA: ACM, 2006: 47—52.

作者简介:

范中磊 男, (1970—), 博士, 教授. 研究方向为分布式计算与存储。

(上接第 78 页)

- [5] Tan L, Gong Y, Chen G. A balance parallel clustering protocol for wireless sensor networks using K-means techniques[C]// Proceedings of the 2008 Second International Conference on Sensor Technologies and Application. Washington, DC, USA: IEEE, 2008: 300—305.
- [6] Manjeshwar A, Agrawal D P. Teen: A routing protocol for enhanced efficiency in wireless sensor networks [C]// Proc 1st International Workshop on Parallel and Distributed Computing Issues in Wireless Networks and Mobile computing. San Francisco, California, USA: IEEE, 2001: 189—196.
- [7] Zhao X, Makki S K, Pissinou N. Determining the routing path for in-networking aggregation[C]// Proc of the IEEE ICHIT. Cheju Island: IEEE, 2006: 318—327.
- [8] Yang S, Dai F, Cardei M, etc. On Multiple Point Coverage in Wireless Sensor[C]. IEEE Conf on Mobile Adhoc and Sensor Systems. Washington, DC, USA: IEEE, 2005: 757—764.

- [9] Liu Y, Liang W. Approximate coverage in wireless sensor[C]// IEEE Conf on Local Computer Networks 30th Anniversary. Washington, DC, USA: IEEE Computer Society, 2005: 68—75.
- [10] 刘明, 龚海刚, 毛莺池, 等. 高效节能的传感器网络数据收集和聚合协议[J]. 软件学报, 2005, 16(12): 2106—2116.
- [11] Yang J, Xu M, Xu B. A multipath routing algorithm based on parametric probability for wireless sensor networks[C]// International Conf on Intelligent Computation Technology and Automation. Washington, DC, USA: IEEE Computer Society, 2009: 417—420.

作者简介:

常波 女, (1970—), 讲师, 硕士. 研究方向为无线通信、自组织网络。

徐保国 男, (1950—), 教授, 博导. 研究方向为无线传感器网络及其控制、工业过程控制与优化、智能仪器及现场总线技术。