

Clustering Algorithm Based on Uneven Distributed Double Cluster Heads for Wireless Sensor Networks*

XIONG Ke, FAN Xiaoping*, LIU Shaoliang, LIAO Zhifang, ZHANG Chunhe
(School of Information Science and Engineering, Central South University, Changsha 410075, China)

Abstract: Aimed at the problem of unbalanced residual energy of each node caused by inter-cluster communications in wireless sensor network routing protocols, a novel algorithm named clustering algorithm based on uneven distributed double cluster heads (BUDC) is presented. BUDC selects main cluster head at random as LEACH scheme, and decides the probability of the relay cluster head generating, according to the distance between the areas node belongs to and base station. To reduce main cluster head's load of relaying message, the relay cluster head is determined by the probability and its residual energy. Simulation shows that BUDC is more effective than LEACH in balancing energy consumption and improving system lifetime.

Key words: wireless sensor network; routing protocols; uneven distributed; double cluster heads
EEACC: 7230; 6150P

一种基于非均匀分布双簇头的 无线传感器网络分簇算法*

熊科, 樊晓平*, 刘少强, 廖志芳, 张纯和

(中南大学信息科学与工程学院, 长沙 410075)

摘要: 针对无线传感器网络分簇路由协议中因簇间路由而产生的节点间剩余能量不平衡问题, 提出一种非均匀分布双簇头的分簇方法。此方法基于 LEACH 算法产生主簇头, 然后根据簇头所属区域离基站远近及其剩余能量确定各簇内数据转发簇头的产生概率, 再按此概率产生非均匀分布的数据转发簇头, 以减轻信息转发负担。仿真表明新算法和 LEACH 等算法相比能更有效地平衡网络中的能量消耗, 延长整个网络的生存周期。

关键词: 无线传感器网络; 路由协议; 非均匀分布; 双簇头

中图分类号: TP393

文献标识码: A

文章编号: 1004-1699(2008)07-1207-05

传感器网络路由协议中^[1], 分簇路由协议通过轮换簇头的方式均衡网络能量消耗。簇头和基站的通信一般采用单跳或多跳方式。单跳方式导致离基站较远的节点能量消耗过快, 多跳方式使离基站较近的节点能量消耗过快, 因而都会提前死亡, 造成整个网络的覆盖面积缩小, 甚至出现不连通的区域, 影响数据采集的完整性, 破坏路由链路^[2]。

LEACH^[3]协议是最早出现的分簇路由协议, 其基本思想是通过等概率周期性地选择簇头, 将整个网络的能量负载平均分配到每个传感器节点, 从

而达到减少网络能量耗费、延长网络生命周期的目的。LEACH 每一轮循环的基本过程是, 在簇的建立阶段, 每个节点选取一个介于 0 和 1 之间的随机数 a , 如果 a 小于阈值 T , 该节点成为簇头并广播成为簇头的消息。其他节点根据接收到广播信号的强弱来决定加入哪个簇并发送加入请求。在数据传输阶段, 簇内的所有节点按照 TDMA 时隙向簇头发送数据。簇头将数据融合之后, 把结果发送给基站, 在持续工作一段时间之后, 进行下一轮簇头选择和新的簇的建立。阈值 T 的计算公式为:

基金项目: 国家自然科学基金资助项目 (60776834)

收稿日期: 2007-12-19 修改日期: 2008-03-25

$$T(n) = \begin{cases} \frac{p}{1 - p \times \left[r \bmod \frac{1}{p} \right]}, n \in G \\ 0, \text{其他} \end{cases} \quad (1)$$

其中: $p = 0.05$, 为预先设定的簇头节点数与节点总数的比率值; r 为当前轮数, $r \bmod (1/p)$ 为这一轮周期中当选过簇头的节点个数; G 为在 $1/p$ 中未当选过簇头的节点集合。

LEACH 簇头和基站采用单跳通信导致每一轮簇头节点剩余能量不均衡, 因此簇头采用多跳和基站通信, 如 HEED^[4] 和 TEEN^[5] 协议。HEED 以节点的剩余能量为依据, 节点成为簇头的概率为:

$$CH_{\text{prob}} = C_{\text{prob}} \times \frac{E_{\text{residual}}}{E_{\text{max}}} \quad (2)$$

其中: $C_{\text{prob}} = 0.05$ 为初始概率; E_{residual} 为节点剩余能量; E_{max} 为节点初始能量, 这使得剩余能量多的节点更有机会成为簇头。

HEED 虽然能够平衡能量消耗, 延长网络生存周期, 但其簇的形成过于复杂, 要损耗一部分能量。TEEN 运用绝对阈值和相对阈值来控制簇头节点数据发送的次数, 由于有两个前提条件, 当节点两个条件都不满足时, 造成节点假死亡, 且 TEEN 并未考虑簇头之间的能量平衡; 因此, 为了平衡传感器网络簇间路由的能量消耗, 需要研究能更有效减少选择簇头时的能耗和均衡簇头能耗的分簇算法。

1 问题描述

1.1 网络模型

对于传感器节点随机分布于正方形区域的传感器网络, 作如下假设:

传感器网络部署之后, 节点都不会随着时间的推移而移动;

每个节点都是同构的, 具有相同的初始能量;

基站位于离传感器网络较远的区域, 网络中部分节点不能直接和基站通信;

传感器节点的发射功率可以进行离散的调节, 簇内通信时为最小值;

传感器节点数量可大于 1 000。

1.2 簇间能量消耗不均衡问题

文献[1]把 LEACH 算法和 MTE 以及 DIRECT 路由协议进行了比较。MTE 采用多跳通信方式, 靠近基站的节点由于要转发数据而消耗更多的能量, 因此这一区域的传感器节点的能量会最先耗尽。在 DIRECT 中, 根据能量公式可知, 离基站越远收发数据所消耗的能量越多。而 LEACH 通过簇头节点的随机轮换, 可以使整个网络的能耗平衡。但 LEACH 只

适应 1 000 个节点以下的网络^[6]。当网络节点数多于 1 000, 如果簇头间通信采用 MTE 和 DIRECT 协议, 簇头间能量消耗将不均衡; 如果在簇头构成的骨干网中, 继续使用 LEACH 算法进行分层, 又将增加节点的通信和计算能耗, 也可能产生通信冲突。对此可采用自适应双簇头的方法, 在一个簇内根据参数生成一个除主簇头以外的副簇头, 用来转发簇间信息, 分担主簇头的通信负担。

文献[7]已证明, 在大型传感器网络中(节点数大于 1 000), 簇内采用多跳通信的 M-LEACH 协议, 每轮节点平均消耗的能量比直接用 LEACH 所消耗的均能小。根据 M-LEACH 协议, 簇间通信中各簇头节点直接和基站通信, 并没有平衡簇间的能量消耗。

由于簇间路由使离基站近的簇头所要转发的数据多于其他区域的簇头, 消耗的能量多, 导致整个网络的能量不均衡, 这就产生了所谓“热区”(hot hops)问题。Soro S 等提出通过非均匀分簇来解决此问题^[8]。借鉴非均匀的思路, 本文提出一种基于非均匀分布双簇头的分簇算法 BUDC (clustering algorithm based on uneven distributed double cluster heads)。该算法在簇头选举的过程中, 簇内根据已选出的簇头和基站的距离来产生用于转发簇间数据的簇头(Relay Cluster Head, 简称 RCH), 使得整个网络中簇头的分布密度随簇头离基站的距离的减少而增大。LEACH 算法没有考虑到簇头间能量消耗的平衡, 而 BUDC 算法在簇间采用多跳通信, 产生非均匀的 RCH 来减少主簇头的通信负担, 均衡每一轮中簇头的能量消耗, 延长传感器网络生命周期。

2 BUDC 算法

BUDC 是使簇头按一定机率在簇内选择剩余能量最多的节点转发数据; 并且这种机率随着节点与基站的距离的减小而增大。

2.1 主簇头选举

LEACH 协议中所有节点成为簇头的概率都为 p , 如果在网络中不同区域通过改变 p 值来产生非均匀分布簇头, 则所有节点不能同时在相同轮次过后都充当过一次簇头。对此, 采用 LEACH 算法产生簇头, 再根据节点和基站距离的远近, 产生非均匀分布的 RCH。簇头间通信采用多跳。簇头采集簇内信息并通过 RCH 传至基站。

2.2 RCH 簇头产生机制

把传感器网络分成如图 1 所示的区域, n 为区域序号, 代表节点离基站的远近程度, n 越大离基站

越远。多跳通信方式导致离基站越近的区域内的簇头所需承担的转发任务越多,能量消耗越多。如果离基站越近的区域可承担转发任务的簇头越多,也即簇头分布密度越大,那么各簇头转发数据所消耗的能量就越小。为了使离基站越近的区域中簇头的分布密度越大,各区域内产生 RCH 的概率设定为 $P = 1/n^3$,再考虑到节点的剩余能量,可得产生 RCH 的阈值:

$$T_R = \frac{E_{residual}}{E_{max}} \times \frac{1}{n^3} \quad (3)$$

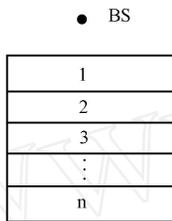


图 1 网络区域划分示意图

簇头将本轮簇头选举阶段产生的随机数 a 与 T_R 进行比较,若 a 小于 T_R ,簇头从簇成员中挑选剩余能量最多的节点作为 RCH,否则,本簇内无 RCH 产生。

2.3 簇的形成

簇成员节点收到簇头广播信息后,根据接收信号的强弱选择加入簇,并通过基于信道竞争的 CS-MA 协议来向簇头发送请求信息。在这一期间,所有的簇头节点都必须处于激活状态。簇头确认节点的入簇请求后,为各成员节点分配一个 TDMA 时序表,完成一轮新的分簇。

2.4 下一轮簇头的选举

簇形成以后,簇间采用多跳通信。但簇头并不知道相邻簇内是否有双簇头。因此簇头在开始传数据前先要向基站查询,以建立到基站的次优路径。在每一轮结束前,各簇的簇头和 RCH 都自动转为普通节点。在下一轮选举时,当选过簇头的节点不参与簇头和 RCH 选举。

2.5 BUDC 算法描述

BUDC 算法流程如图 2 所示。在簇头和副簇头形成阶段,节点间的通信消息类型定义为四种,分别为 ADmessage (area, NodeID, Rejoinmessage (area, NodeID, $E_{re}(i)$), AD_CHRmessage (NodeID, NodeID, CH) 和 CF_CHRmessage (NodeID, CH)。Area 为节点所在区域相应区号;NodeID 为每个节点唯一网络标号; $E_{re}(i)$ 为节点的剩余能量;CH 为布尔变量,为真时代表节点上一轮没有充当过副簇头。在 AD_CHRmessage 消息中,第一个 NodeID

为簇头标识,第二个 NodeID 为簇成员节点中剩余能量最多的节点标识。

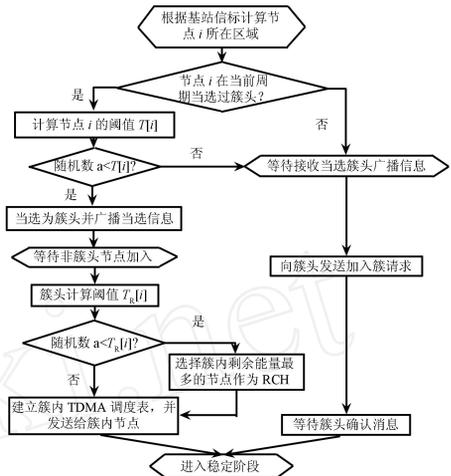


图 2 BUDC 算法流程

算法伪代码如下:

```

计算出每个节点所属层 (area)
初始化 BeAclusterhead = false, RCH = false
产生 0-1 间的随机数 a
If a < T(i) then
    节点当选簇头, 变量 BeAclusterhead = true, 并广播
    当选消息 ADmessage (area, NodeID)
Endif
If a > t(i) then
    接收从当选簇头当选消息, 并根据接收信号的强弱
    来决定加入簇并发送加入请求 Rejoinmessage (area, NodeID, E_re(i))
Endif
If 节点是簇头 BeAclusterhead = true
    ⑪ 计算出 T_R(i)
    ⑫ If a < T_R(i) // 表示该簇可以产生副簇头
    ⑬ 选择成员能量表中能量最多的节点作为副簇头, 再
    发送请求 AD_CHRmessage (NodeID, NodeID, CH)
    ⑭ Endif
    ⑮ Endif
    ⑯ 成员节点接收到广播, 和自身 ID 做比较, 如果 ID 和
    本身 ID 相同, 记住簇头节点 ID 号, 并发送确认消息 CF_
    CHRmessage (NodeID, CH)

```

3 仿真及分析

在簇头选择产生转发数据簇头时,无需对节点的剩余能量进行测量,可根据文献[9]提出的传感器节点接收和发送数据的能量模型进行简单计算得到通信能量消耗,即若发送 1 bit 数据到距离为 d 的目的地址所消耗的能量为:

$$E_{Tx}(l, d) = \begin{cases} le_1 + le_2 d^2, & d < d_0 \\ le_1 + le_3 d^4, & d > d_0 \end{cases} \quad (4)$$

传感器节点接收数据所消耗的能量为:

$$E_{RX}(l, d) = le_1 \quad (5)$$

其中, e_1 为射频装置和接收电路每发送单位数据所消耗的能量, J/bit ; e_2 为发射功率放大器将一比特数据发送单位平方米所耗费的能量值, $J/(\text{bit} \cdot \text{m}^2)$; d_0 为单跳通信的最大距离。

每个节点根据其到基站的距离, 可以确认出所属区域。根据文献[9], 任何节点在转发一个数据包所消耗的能量为:

$$E_n = E_{TX}(l, d) + E_{RX}(l, d) \quad (6)$$

由于节点和其邻居节点的距离小于 d_0 , 所以式(6)可进一步表示为:

$$E_R = 2le_1 + le_2 d^2 \quad (7)$$

在第 L 层的簇头节点平均所需要转发的数据包 L_n 为

$$L_n = m \times \frac{A^2 - nL^2}{A^2 \times (n^2 - (n-1)^2) \times d \times L^2} \quad (8)$$

仿真平台采用 MATLAB, 分别在三种不同的场景对 LEACH、M-LEACH、HEED 和 BUDC 进行实验, 其参数如表 1 所示。在边长为 A 的正方形区域, 随机分布的 n_0 个坐标点代表传感器节点, 基站位于 (x, y) 处, round 表示分簇的轮数, 取 round = 1 500。为便于比较, 假定在不同的场景每个节点所具有的初始能量不同, 场景 1、2、3 里节点的初始能量分别为 0.5 J、2 J、4 J。每个节点接收和发送电路需要消耗能量为 $e_1 = 50 \text{ nJ}/\text{bit}$, $e_2 = 10 \text{ pJ}/(\text{bit} \cdot \text{m}^2)$, $e_3 = 0.0013 \text{ pJ}/(\text{bit} \cdot \text{m}^4)$, 每个数据包大小固定为 2 000 bit, 每个节点数据融合所消耗的能量为 $E_{DA} = 5 \text{ nJ}/(\text{bit} \cdot \text{signal}^{-1})$, $d_0 = 70 \text{ m}$ 。节点死亡发生在剩余能量为零时。

表 1 三种场景参数

	场景 1	场景 2	场景 3
节点数 (n_0)	100	600	1000
基站坐标 (x, y)	100, 275	400, 875	600, 1275
场景边长 A/m	200	800	1200
$E_{\text{initial}}/\text{J}$	0.5	2	4

针对表 1 所示三种场景参数仿真的结果分别如图 3、图 4 和图 5 所示。若以网络节点死亡数超过 80% 的时间来衡量整个网络的生存周期, 由于节点随机分布在 $200 \text{ m} \times 200 \text{ m}$ 域中时, LEACH 和 HEED 算法中簇内通信采用单跳方式, 而 M-LEACH 算法中, 簇内采用多跳通信, 所以其簇内的能耗多于采用 LEACH 和 HEED 算法时簇内的能量消耗。从图 3 看出, 采用 LEACH 比采用 M-LEACH 的网络能耗多 13%。HEED 分簇数目固定且分布均匀。从图 3、图

4 和图 5 中可看出 HEED 从第一个节点死亡到网络存活节点数剩余 20% 时, 所经历的时间相对长, 没有很好地均衡整个网络的能量消耗, 而在应用 BUDC 算法的网络中第一个节点死亡的时间延后了 200 轮。无线传感器网络使用 BUDC 算法与使用 LEACH 算法相比, 第一个节点死亡出现的时间要晚, 这是因为两种算法中, 簇头分布是随机的, 簇头和基站的通信距离有可能大于单跳传输的最大距离。与场景 1 相比, 场景 2 和场景 3 的节点数量和通信距离急剧增长, 根据能量模型, 当节点和基站的距离大于 600 时, 节点所消耗的能量大于 0.5 J, 这造成 LEACH 和 M-LEACH 的第一个死亡节点都出现在前 200 轮。场景 2、3 中, BUDC 有效地平衡了簇间路由的能量消耗, 从图 4、图 5 可看出采用 BUDC 的网络的生存周期明显长于其他协议, 第一个节点死亡时间在 200 轮和 450 轮之间。

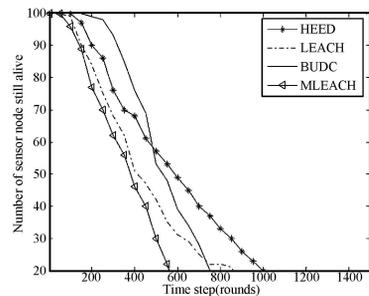


图 3 场景 1 中 LEACH、M-LEACH、HEED 和 BUDC 算法下的存活节点数

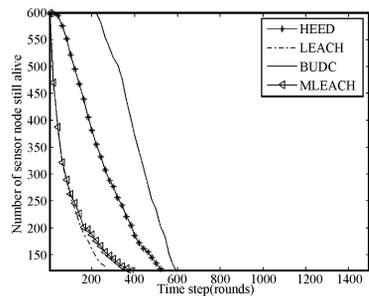


图 4 场景 2 中 LEACH、M-LEACH、HEED 和 BUDC 算法下的存活节点数

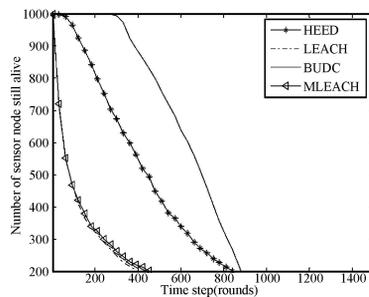


图 5 场景 3 中 LEACH、M-LEACH、HEED 和 BUDC 算法下的存活节点数

在 BUDC 算法中,阈值 T_R 的参数 n 的幂次,也即转发簇头产生概率的选取对整个网络的生存周期有直接影响。依次取 n^2 、 n^3 、 n^4 和 n^5 ,采用场景 3 进行横向对比,从图 6 可看出,在网络生存周期内,当 T_R 的参数 n 的幂取 3 时,节点死亡的比例基本保持不变,这使得在第一个节点死亡后,网络的存活节点数是均匀递减的,有利于维持网络的稳定性。

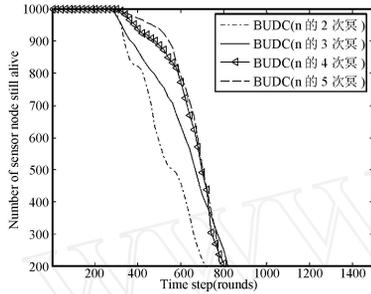


图 6 场景 3 中 BUDC 算法参数取 n^2 、 n^3 、 n^4 、 n^5 时的网络存活节点数

4 结论

针对中型传感器网络中,簇间路由所产生的簇头能量消耗不均衡,本文提出了簇间多跳通信方式的 BUDC 算法。和 LEACH 等算法相比,该算法能产生非均匀分布的 RCH,既能减轻簇头间的通信负担,又能减少主簇头的能量消耗。仿真结果表明,与平均分簇的 LEACH、M-LEACH 和 HEED 相比,BUDC 算法能更有效地平衡整个网络的能量消耗,延长网络生命周期。

参考文献:

- [1] Akyildiz I F, Su W, Sankarasubramanian Y, et al. Wireless Sensor Networks: A Survey. Computer Networks[J], IEEE Communications Magazine, 2002, 38: 393-422.
- [2] 邢云冰,史浩山. 基于备用节点的无线传感器网络 LEACH 协议的改进[J]. 传感技术学报,2007,20(7):1592-1596.
- [3] Heinzelman W, Chandrakasan A, Balakrishnan H, An Application-Specific Protocol Architecture for Wireless Microsensor Networks[C]// IEEE Transactions on Wireless Communications, October 2002, 1(4).
- [4] Younis O, Fahmy S, Distributed Clustering in Ad hoc Sensor Networks: A Hybrid, Energy Efficient Approach[C]// Proceedings of IEEE INFOCOM, March 2004.
- [5] A Manjeshwar, D P Agrawal. TEEN: a Routing Protocol for Enhanced Efficiency in Wireless Sensor Networks[C]// Parallel and Distributed Processing Symposium Proceedings 15th International, 2001-04: 23-27.
- [6] Younis O, Fahmy S. An Experimental study of Routing and Data Aggregation in Sensor Networks [C]// Proceedings of the International Workshop on Localized Communication and Topology Protocols for Ad hoc Networks (LOCAN), Washington, DC, 2005.
- [7] Mhatre V, Rosenberg C, "Homogeneous Vs Heterogeneous Clustered Networks: A Comparative Study"[C]// Proceedings of IEEE ICC 2004, June 2004.
- [8] Soro S, Heinzelman W. Prolonging the Lifetime of Wireless Sensor Networks Via Unequal Clustering [C]// Proceedings of the 5th International Workshop on Algorithms for Wireless, Mobile, Ad Hoc and Sensor Networks, Denver, CO, 2005.
- [9] Heinzelman W, Chandrakasan A, Balakrishnan H. A Energy Efficient Communication Protocol for Wireless Microsensor Networks[C]// Proceedings of the 33rd Annual Hawaii International Conference on System Sciences, Maui, HI 2000: 1210.



熊 科(1982-),男,硕士研究生.主要研究方向为无线传感器网络,xkelora@163.com



刘少强(1964-),男,博士,副教授.主要研究方向为智能检测与控制,无线传感器网络等,liussqq@126.com



樊晓平(1961-),男,博士,教授,博士生导师.主要研究方向为机器人控制,无线传感器网络,智能交通系统等,xp-fan@ieee.org