

# 无线传感器网络分布式成簇算法优化<sup>①</sup>

陶志勇, 方 宁

(辽宁工程技术大学 电子与信息工程学院, 葫芦岛 125105)

**摘 要:** 能量异构问题在通信过程中会成为无线传感器网络不可避免的现象, 为解决能量负载不均衡问题, 对一种适用于能量异构环境的无线传感器网络分布式成簇算法进行改进。算法结合能量水平相对估计因子使高能量节点更易成为簇头; 簇间通信时根据通信代价选择单/多跳方式, 同时设置低能量的簇拒绝转发其他簇的数据以达到均衡能耗的目的。模拟实验结果表明, 新算法能更好的均衡节点能量负载, 其能量有效性得到了很大的提高。

**关键词:** 无线传感器网络; 能量异构; 成簇算法; 能量有效性; 能量均衡

## Improvement of Distributed Clustering Algorithm in Wireless Sensor Networks

TAO Zhi-Yong, FANG Ning

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

**Abstract:** Energy-heterogeneous is unavoidable for wireless sensor networks when communicating. An improved novel clustering algorithm for wireless sensor networks is proposed to adapt for energy-heterogeneous environments, and to solve the problem of imbalance in energy load for whole network. In the algorithm, the relative estimation factor of nodes' current energy level which can make higher energy nodes become cluster heads easier. When communicating between different clusters it will chose single hop or multi hop according to the communication costs. Lower energy clusters will not retransmission other clusters' data to balance the energy load. The result of simulation shows that the algorithm has been improved greatly and the energy load for whole network is more balanced.

**Key words:** wireless sensor network; energy-heterogeneous; clustering algorithm; energy-efficient energy balance

### 1 引言

无线传感器网络(Wireless Sensor Network, 简称 WSN)所用的路由协议具有节点较为固定、能量高度受限、网数据冗余大、应用相关、以数据为中心等特点。从网络通信的角度看, 传感器网络的基本特点是能量低、规模大、移动性弱、拓扑易变化、使用数据融合技术和通信的不对称性。这使它的路由协议设计面临着节能和提高扩展性两个关键性问题。

分簇是无线传感器网络路由算法研究的主要方向之一, 具有较好的扩展性和能量有效性。簇头节点负责管理簇内节点、协调成员节点之间的任务、收集簇内信息并进行数据融合、信道分配及簇间传递转发; 成员节点则只收集数据然后传输给簇头节点。

LEACH(Low-energy adaptive clustering hierarchy)<sup>[1]</sup>是 WSN 中最早的分层路由算法, 其成簇思想成为许多分簇协议的基础, 如 TEEN(Threshold sensitive energy efficient sensor network protocol)<sup>[2]</sup>, HEED(Hybrid energy-efficient distributed clustering)<sup>[3]</sup>等。

LEACHE 协议由一个相等的概率随机循环选择簇头。所有操作按轮(Round)循环进行, 每一轮分为簇的建立阶段和数据传输阶段。所有节点轮流担任簇首以此平均地分配负载。每一个节点随机产生[0,1]之间的数, 若小于阈值  $T(n)$  则成为簇首。

$$T(n) = \begin{cases} \frac{P}{1 - P[r \bmod 1/P]}, & n \in G \\ 0, & \text{其他情况} \end{cases} \quad (1)$$

<sup>①</sup> 基金项目: 辽宁省教育厅科学技术研究自然科学基金(10287)

收稿时间: 2011-11-11; 收到修改稿时间: 2011-12-26

其中  $P$  是簇头数目的百分比,  $r$  是当前轮数,  $G$  是未当选簇头的节点集合。其他节点依据接收的广播信号强度决定加入哪个簇, 并回复给该簇簇头。数据传输时, 普通节点依照 TDMA(时分复用)时隙向簇头发送数据。簇头进行数据融合并发送给基站。

随着通信的进行网络逐渐呈现能量异构特点, 这使得 LEACH 在实际应用中难免差强人意。鉴于实际应用中的能量异构现象, 一个好的分簇算法应能适应能量异构环境, 最大限度地延长网络生命周期。

## 2 EHDC成簇算法

为解决网络负载不平衡的问题, 文献[4]提出 EHDC 算法, 算法在成簇阶段考虑了节点的剩余能量, 利用前一轮的本地信息预算节点的下一轮能量水平使能量高的节点更有机会出任簇头。这样使得当前能量较高的节点能够拥有更大的概率出任簇头, 以此协调整个网络负载, 从而均衡网络的整体能耗。

然后引入基于能量的相对估计因子——节点能量评估因子, 以此调整节点出任簇头的概率及其轮转周期, 力求能量高的节点更易成为簇头节点从而实现均衡网络能耗的目的。定义节点能量估计因子如下: 设节点  $i$  在第  $r$  轮的能量评估因子是  $\delta(i, r)$ , 则

$$\delta(i, n) = \begin{cases} \frac{E_{current}(i, r)}{E_{net-average}}, & (r=1) \\ \frac{E_{current}(i, r)}{E_{cluster-average}(j, r-1)}, & (r>1) \end{cases} \quad (2)$$

其中,  $E_{current}(i, r)$  是节点  $i$  在第  $r$  轮的初始能量;  $E_{net-average}$  是网络平均能量;  $E_{cluster-average}(j, r-1)$  是节点  $i$  所在第  $j$  簇在本轮(第  $r-1$  轮)结束后的簇中节点的平均能量。式(2)表明, 当前能量越大的节点, 其能量评估因子越大。由于簇头轮转过程中( $r>1$ )仅利用前一轮结束后的平均能量就可以得到下一轮节点的能量评估因子, 使得 EHDC 算法在簇头轮转过程中仍然具有分布式算法的优点。

$P_{opt} = \frac{k_{opt}}{N}$  表示优化簇头比例,  $k_{opt}$  是网络最优簇头数,

$$\text{由文献[5]可知: } k_{opt} = \frac{\sqrt{N}}{\sqrt{2\pi}} \sqrt{\frac{\xi_{fs}}{\xi_{mp}}} \frac{M}{d_{toBS}^2}$$

令网络得轮转周期为  $T_r = 1/P_{opt}$ , 可使每一个节点在每个轮转周期至少成为一次簇头。引入能量评估因子

可以调节节点的轮转周期, 使能量更高的节点获得更有机会出任簇头:  $P(i, r) = P_{opt} \delta(i, r)$ ;

$$T_r(i) \frac{1}{P_{opt} \delta(i, r)} = \frac{T_r}{\delta(i, r)}$$

此时的簇首节点数为:

$$E(\#CH) = \sum_{i=1}^N P(i, r) = \sum_{i=1}^N P_{opt} \delta(i, r) = P_{opt} \sum_{i=1}^N \delta(i, r) = P_{opt} N = k_{opt}$$

引入能量评估因子  $\delta(i, r)$  后, 最优簇头数  $k_{opt}$  仍能得到保证。这样, 簇头选择的阈值可变为:

$$T(i) = \begin{cases} \frac{P_{opt} \delta(i, r)}{1 - P_{opt} \delta(i, r)} \left[ r \bmod \frac{1}{P_{opt} \delta(i, r)} \right], & i \in G \\ 0, & \text{其他情况} \end{cases} \quad (3)$$

其中  $r$  为当前的轮数,  $G$  为节点最近  $\frac{1}{P_{opt} \delta(i, r)}$  轮

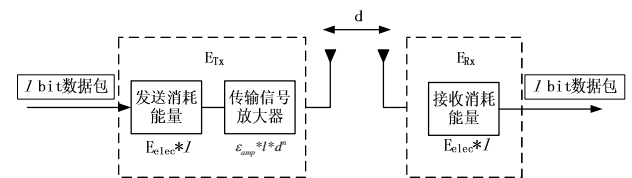
中没有成为簇头的节点集合。式(3)表明, 节点的能量评估因子越大, 其出任簇头的机会越多, 轮转周期越短, 从而达到使网络中所有节点几乎同时耗尽能量的目的。

由此, 能量评估因子大的节点出任簇头的机会就越多, 轮转周期就越短, 以此来使网络中所有节点几乎同时耗尽能量的目的。

## 3 基于能量评估因子的算法优化

随着网络消耗逐渐增大, 造成网络内部数据转发量较大的簇耗费更多的能量, 最终导致网络出现空洞。为此, 我们对 EHDC 算法进行改进并命名为 EHDC-E。

为均衡各簇间的能量消耗, 本文采用多跳和单跳相结合的通信方式。采用文献[5]中的信道模型, 发送 1 比特信息到距离为  $d$  的接收方时, 则发送能量开销为  $E_{Tx}(l, d)$ 。传输模型如下:



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

设  $d_0 = \sqrt{\frac{\epsilon_{fs}}{\epsilon_{mp}}}$ , 其中  $\epsilon_{fs}$ 、 $\epsilon_{mp}$  分别是自由空间信道和多径衰落信道模型下功率放大器的能量消耗常数。

基站首先向全网广播消息，簇头节点接收基站的广播消息后计算与基站的近似距离，簇间传输时，簇头根据这个距离决定采用单跳/多跳形式。若  $d > d_0$ ，则采用多跳形式；若  $d \leq d_0$  则采用单跳形式。

若采用多跳路由，引入参数  $c$  作为选择标准。设当前簇头靠近基站一端的邻居簇头剩余能量为  $E_{current}$ ，与当前簇头的距离为  $d$ ，簇头与基站的距离为  $d_{direct}$ 。若该邻居簇头不与基站直接通信，则令  $c = E_{current}/d$ ，优先选取  $E_{current}$  最大的；若该邻居簇头与基站直接通信，则令  $c = E_{current}/d_{direct}$ ，优先选取  $d_{direct}$  最小的。每次都选取  $c$  最大的簇头作为下一跳节点，这样能量越大、位置越有利的邻居簇头越有可能当选为下一跳节点<sup>[6]</sup>。簇间路由的确定也决定了各簇间能量消耗的不等，因此也证明了各簇独立进行簇头转换的必要性和正确性。

另一处改进思想是：簇的平均能量小于一个基于初始能量的阈值时，该簇的簇首节点将设置一个标志位。在接下来的通信中，被标示的簇将仅仅进行数据的采集和自身数据传输而拒绝数据转发工作，以此节省自身能量。簇间通信时，能量较低的簇杜绝了其他簇首节点的转发数据，选择次优簇首传输数据包从而实现网络能量负载的均衡。

成簇完成后，进入数据采集和转发阶段。各簇头节点向 CH-BS 层的其他簇头节点发送 CH-CH-HELLO 消息<sup>[7]</sup>，消息格式如表 1 所示。

表 1 簇头的 CH-CH-HELLO 消息格式

ID	nodeType	mesType	turn	curEnergy	Identify	disToBS
----	----------	---------	------	-----------	----------	---------

其中，ID 是节点标号；nodeType 是节点类型(簇头节点/非簇头节点)；mesType 为消息类型(簇内消息/簇间消息)；turn 为轮标志符以区别不同轮的消息；curEnergy 为簇头节点当前能量；Identify 为剩余能量标志，节点剩余能量大于基于初始能量的阈值，则为 0，否则置为 1；distToBS 为簇头到基站的距离。

CH-BS 混合多跳路由中各个簇头节点执行过程如下：

a)簇头广播自身 CH-CH-HELLO 消息并接收其他簇头的同类消息；记录各消息中簇头的当前剩余能量以及个簇头节点到 BS 的距离。

b)判断簇头节点与基站距离是否小于单跳距离  $d_0$ ，如果是，则将数据直接发送给基站，否则进入步

骤 c)。

c)选择各个簇头中参数  $c$  最小的，查看其 Identify 标示位是否为 1，若为 0 则将其作为下一跳节点，否则查看次优节点(即参数  $c$  为次优值的簇头)并进入步骤 c)。

d)直至将数据包传至基站为止。

e)簇头节点数据转发结束。

CH-BS 混合多跳路由选择算法的流程图如图 1 所示：

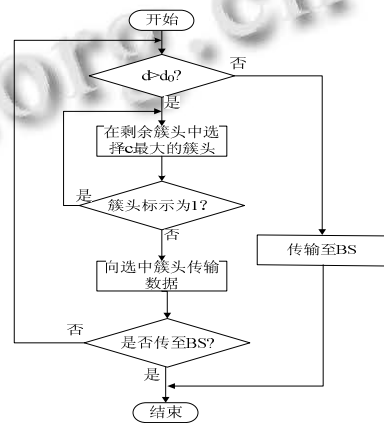


图 1 CH-BS 路由选择流程图

## 4 算法的仿真及结果分析

### 4.1 对节点能量评估因子的计算和分析<sup>[4]</sup>：

$r=1$  时， $\delta(i,1)=1$ ；

$r>1$  时，设节点  $i$  根据  $r-1$  轮收到的广播信息强度决定加入第  $j$  簇，向相应簇头发送请求信息时附带自身能量值  $E_{current}(i, r-1)$ ，簇首根据加入请求数得出本簇节

点总数  $C_j$  及  $\sum_{i=1}^{C_j} E_{current}(i, r-1)$ ，然后估算本簇在本轮能

耗，进而得出本簇平均能量  $E_{cluster-average}(j, r-1)$ 。向成员节点发送 TDMA 时隙表时附带平均能量，簇内节点在第  $r-1$  轮结束后，就可以根据自身的当前能量

$E_{current}(i, r)$ ，获得下一轮能量评估因子值  $\delta(i, r)$ 。计算如下：设区域  $M \times M$  内有  $N$  个传感器节点，第  $r-1$  轮有  $k$  个簇，第  $j(1 \leq j \leq k)$  簇内共有  $C_j$  个传感器节点，簇内节点  $i$  的本轮初始能量为  $E_{current}(i, r-1)$ ，则第  $r-1$  轮第  $j$  簇的起始总能量为：

$$E_{begin-total}(j, r-1) = \sum_{i=1}^{C_j} E_{current}(i, r-1)。$$

所以，由文献[5]可知在  $r-1$  轮的节点  $i$  的消耗能量为：

$$E_{consume-node}(i, j, r-1) = \begin{cases} lE_{elec}C_j + lE_{DA}C_j + l\epsilon_{mp}d_{toBS}^4, & (i \text{ 是簇头}) \\ lE_{elec} + l\epsilon_{fs} \frac{1}{2\pi} \frac{M^4}{k^2}, & (i \text{ 不是簇头}) \end{cases} \quad (5)$$

其中, 1 是节点发送或接收数据的 bit 数;  $E_{elec}$  是发送电路和接收电路消耗的能量;  $E_{DA}$  是簇头进行数据融合时消耗的能量;  $d_{toBS}$  是簇头到基站的距离。则  $r-1$  轮中第  $j$  簇耗能为:

$$E_{consume-total}(j, r-1) = \sum_{i=1}^{C_j} E_{consume-node}(i, j, r-1) = (E_{elec}C_j + lE_{DA}C_j + l\epsilon_{mp}d_{toBS}^4) + (C_j - 1) \times (lE_{elec} + l\epsilon_{fs} \frac{1}{2\pi} \frac{M^4}{k^2})$$

第  $j$  个簇的剩余能量为

$$E_{end-total}(j, r-1) = E_{begin-total}(j, r-1) - E_{consume-total}(j, r-1)$$

簇头估算第  $j$  簇平均能量为:

$$E_{cluster-average}(j, r-1) = \frac{E_{end-total}(j)}{C_j}$$

节点  $i$  当前能量为:

$$E_{current}(i, r) = E_{current}(i, r-1) - E_{cluster-node}(i, j, r-1)$$

则

$$E_{end-total}(j, r-1) = \sum_{i=1}^{C_j} E_{current}(i, r)$$

因此, 第  $r$  轮能量评估因子:

$$\delta(i, r) = \frac{E_{current}(i, r)}{E_{cluster-average}(j, r-1)}$$

由此可知, 当前能量越大的节点其能量评估因子越大:

$$\begin{aligned} \sum_{i=1}^{C_j} \delta(i, r) &= \sum_{i=1}^{C_j} \frac{E_{current}(i, r)}{E_{cluster-average}(j, r-1)} = \frac{\sum_{i=1}^{C_j} E_{current}(i, r)}{E_{cluster-average}(j, r-1)} \\ &= \frac{E_{end-total}(j, r-1)}{E_{cluster-average}(j, r-1)} = C_j \end{aligned}$$

就全网来说:

$$\sum_{i=1}^N \delta(i, r) = \sum_{j=1}^k \sum_{i=1}^{C_j} \frac{E_{current}(i, r)}{E_{cluster-average}(j, r-1)} = \sum_{i=1}^N C_j = N$$

### 4.2 仿真分析

为更好地对比协议的性能, 在实验中分别对几种算法进行对比并采用 MATLAB 软件给出仿真。由于实际应用中节点很难绝对平均分布, 在此我们选择将 100

个节点随机分布于检测地区中, 节点分布图如图-2, 仿真参数如表 2。

EHDC-E、EHDC、LEACH 的生存时间依次递减, 这是因为改进后的协议体现出良好的能耗均衡性。显然, 成簇稳定期越长, 网络能量有效性越好。

表 2 仿真参数表

参数	值
网络范围/m	100×100
基站位置/m	(50,175)
节点数	100
初始能量/J	2
$E_{elec} / (nJ \cdot bit^{-1})$	50
$\epsilon_{fs} / (pJ \cdot (bit \cdot m^2)^{-1})$	10
$\epsilon_{mp} / (pJ \cdot (bit \cdot m^4)^{-1})$	0.0013
$E_{DA} / (nJ \cdot (bit \cdot signal)^{-1})$	5
数据包大小/bit	500

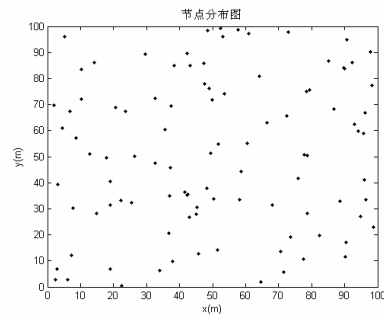


图 2 节点分布图

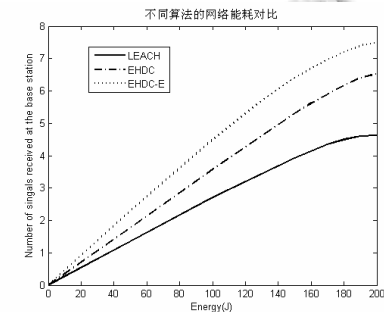


图 3 基站接收数据

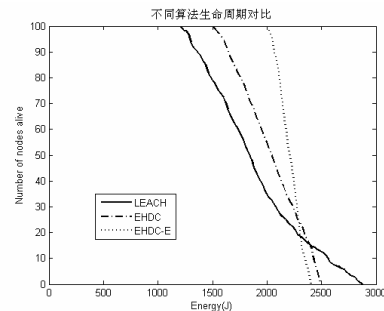


图 4 算法的生命周期比较

算法的能量消耗均衡性网络能量消耗均衡性是衡量分簇算法性能的重要指标,具体表现在每轮的耗能及网络死亡期方面。显然,网络每轮消耗的能量越小越均匀,网络死亡期越短,网络能量消耗均衡性越好<sup>[4]</sup>。

协议中采用适用于能量异构情况的簇头选举机制同时采用了合理的簇间通信方式机制,使得平均能量较低的簇不再转发和传送数据包而是仅仅收集数据,这样就有效地均衡网络的整体消耗。改进后的协议始终选择能量高的节点作为簇头和簇间通信转发节点,将簇头剩余能量和簇头转换相关联,从而能很好地均衡网络节点能耗。

## 5 结语

本算法主要从簇头选择、簇间传输路径两个方面对分簇路由算法进行分析与改进。簇头选择中,剩余能量较高的节点竞争为候选簇头,簇间通信过程中,运用单跳/多跳的数据传输模式,基于通信代价和簇内剩余平均能量选择簇间数据转发模式,有效均衡网络负载。通过分析和仿真证明,算法在网络生存周期、负载均衡等多方面都显示了很好的性能。下一步计划研究路由协议的容错能力和具体的数据融合算法。

## 参考文献

- 1 Heinzelman W, Chandrakasan A, Balakrishnan H. An application specific protocol architecture for wireless microsensor networks. *IEEE Trans. on Wireless Communications*, 2002,1(4):660-670.
- 2 Manjeshwar A, Grawal D P. TEEN: a protocol for enhanced efficiency in wireless sensor networks. *Proc. of the 15th Parallel and Distributed Processing Symp.* San Francisco: IEEE Computer Society, 2001,2009-2015.
- 3 Younis O, Fahmy S. Heed: a hybrid, energy-efficient distributed clustering approach for ad-hoc sensor networks. *IEEE Trans. on Mobile Computing*, 2004,3(4):660-669.
- 4 刘新华,李方敏,旷海兰,等.基于能量异构的无线传感器网络分布式成簇算法. *小型微型计算机系统*,2010,1:26-31.
- 5 w'Chandrakasan HA, Balakrishnan H. An application specific protocol architecture for wireless microsensor networks. *IEEE Trans. on Wireless Communications*, 2002,1(4):660-670.
- 6 张强,卢潇,崔晓臣.基于能量高效的无线传感器网络 LEACH 协议改进. *计算机工程与设计*,2011,32(2):427-429.
- 7 周钰川,施荣华,周媛媛.WSN 中基于非均匀簇的混合多跳路由协议. *计算机应用研究*,2011,28(2):642-644.

(上接第 239 页)

- 2 胡磊.基于模板的可视化软件理解研究与实现.武汉:武汉大学,2004.
- 3 Harbison III SP, Steele Jr GL. C 语言参考手册.北京:机械工业出版社,2003.40-45.
- 4 CRegKey.[2011-8-31].[http://msdn.microsoft.com/zh-cn/library/xka57\\_xy4 \(v=VS.80\).aspx](http://msdn.microsoft.com/zh-cn/library/xka57_xy4 (v=VS.80).aspx).
- 5 Mayers.STL 高效编程.北京:机械工业出版社,2006.64-80.
- 6 Booch G, Rumbaugh J, Jacobson I.UML 用户指南.北京:机械工业出版社,2006.164-170.
- 7 Petzold C.方敏等译. windows 程序设计.北京:清华大学出版社,2010.240-245.
- 8 Hind M, Pioli A. Which Pointer Analysis Should I Use. *ACM SIG-SOFT Software Engineering Notes*. 2000, 25(5):113-111.
- 9 李莹,张琴燕.程序理解. *计算机应用研究*,2001,18(6):40-41.