

无线传感器网络分区能耗均衡的非均匀分簇算法

王毅, 张德运, 梁涛涛

(西安交通大学电子与信息工程学院, 710049, 西安)

摘要: 针对在无线传感器网络中分簇算法采用多跳通信方式时, 靠近汇聚节点的簇头节点由于转发大量数据而导致自身能量消耗过快且节点易失效, 从而造成网络分割, 形成“热区”的问题, 提出了一种新颖的基于分区能耗均衡的多跳非均匀分簇算法(CEB-UC). 其核心思想是: 将传感器网络合理分区, 使得在靠近汇聚节点分区内的簇数量较多, 各簇内传感器节点数较少; 在远离汇聚节点分区内的簇数量较少, 各簇内的传感器节点数较多, 从而保证承担数据中继转发任务的簇头节点能减少自身的簇内通信开销, 节约的能量可供簇间数据转发使用; 任意分区的簇头节点在选择下一跳中继节点时可综合考虑候选节点的位置及剩余能量. 实验以 HEED、LEACH、PEGASIS、EAR-ACM 等协议为参照, 结果表明, 当节点使用不同的数据融合策略时, CEB-UC 算法能有效平衡网络节点能耗, 延长网络部署半径, 降低 17%~41% 的簇头节点能耗, 提高 24%~53% 的网络寿命.

关键词: 无线传感器网络; 分簇; 能耗均衡; 网络寿命

中图分类号: TP393 **文献标识码:** A **文章编号:** 0253-987X(2008)04-0389-06

Cell Energy Balanced Uneven Clustering Hierarchy Scheme for Wireless Sensor Networks

WANG Yi, ZHANG Deyun, LIANG Taotao

(School of Electronics and Information Engineering, Xi'an Jiaotong University, Xi'an 710049, China)

Abstract: Employing clustering techniques in routing protocols can improve the scalability of wireless sensor networks. To solve the “hotspot” problem that is caused by the cluster heads close to the sink, which are burdened with heavy relay traffic via multi-hop communication and tend to die earlier and that leads to the network being partitioned, a novel cell energy balanced uneven clustering hierarchy scheme (CEB-UC) is presented. In CEB-UC, the entire network has been divided into some cells reasonably such that the number of clusters in each cell is different, and that the number of nodes in each cluster is different. The clusters closer to the sink node have fewer nodes than those farther away from the sink node. Thus these cluster heads can preserve some energy for inter-cluster data relaying. Compared with some classical clustering protocols, such as HEED, LEACH, PEGASIS, EAR-ACM, simulation results show that CEB-UC outperforms significantly in optimizing the cluster heads' energy consumption, balancing the nodes' energy consumption and enlarging the diameter of the entire network. When different data aggregation strategies are employed by the sensor nodes, the cluster heads' energy consumption is reduced by 17%-41%, and the network lifetime is prolonged by 24%-53%.

Keywords: wireless sensor networks; clustering; energy balance; network lifetime

在无线传感器网络拓扑协议中, 使用分簇技术 可改善无线传感器网络的扩展性^[1-4]. 已有的分簇算

法^[2-7]多采用单跳通信方式(簇头节点直接与汇聚节点通信,如 LEACH^[2]、LEACH-C^[3]、HYENAS^[4]等),但存在成簇开销大、网络部署面积受限、扩展性较差等诸多不足^[8].研究表明,传感器节点与汇聚节点之间通信若采取多跳的方式更有利于节能,但存有致命的缺陷:靠近汇聚节点的簇头节点需要转发大量的数据,从而造成负担过重,并过早地耗尽自身能量,最终导致网络分割、网络寿命降低.这就是所谓的“热区”问题.

综上,本文设计了一种分簇协议:分区能耗均衡的非均匀分簇(CEB-UC)算法.

1 系统模型

CEB-UC 算法按轮运行,每轮又分簇形成 T_{cluster} 和数据的收集、发送 T_{data} 2 个阶段,并保证 $T_{\text{data}} \gg T_{\text{cluster}}$.

1.1 网络模型

设有 N 个传感器节点随机均匀分布在一个二维平面区域,且该网络具有如下性质:①传感器节点部署后不再移动;②传感器节点同构,并在网络中的地位、作用相同;③节点具有感知半径 R_s 和数据传输半径 R_t ,且 $R_t > R_s$;④网络中只存在一个汇聚节点,且固定在节点部署区域之外的某一个定点上;⑤簇内通信采用单跳模式,簇头节点与汇聚节点通信采用多跳模式;⑥无线发射功率可控,即节点可以根据需要调整发射功率;⑦节点未安装 GPS 设备,也不知道其具体物理位置,但具有分区标识位.

由于算法需要节点的分区信息,所以本文根据网络部署模式提出了 2 种方法:①若网络由使用者部署,可设置节点预知分区信息;②若网络随机部署,则使用基于接受信号强度指示(RSSI)来估算分区.汇聚节点用最大功率来发送一系列的数据包,节点接收数据包,并计算平均 RSSI 强度、计算信号的传播损耗、估算节点与汇聚节点的相对距离,然后根据应用模型估算其所在分区^[9].

1.2 无线通信模型

本文与文献[3]使用了类似的无线通信模型.若节点 a 在向距离 d 以外的节点 b 发送字节为 l 的数据包时,其数据发送能耗为

$$E_T(l, d) = E_t(l) + E_c(l) + E_a(l, d) = \begin{cases} lE + l e_n d^2 & d \leq d_0 \\ lE + l e_f d^4 & d > d_0 \end{cases} \quad (1)$$

而节点 b 接收节点 a 发送的信息,其无线接收装置所需能耗为

$$E_R(l) = E_t(l) + E_c(l) = lE \quad (2)$$

1.3 数据融合模型

假设存在数据融合率 α ,若 $\alpha=1$,算法采用数据完美融合策略;若 $\alpha=0$,算法不采用任何的数据融合策略.在下述讨论中,用数据融合率 α 来代替特定的数据融合策略.

2 CEB-UC 算法描述

CEB-UC 算法将无线传感器网络部署区域(面积为 S)分为 K 个分区,记为 $P_1, P_2, \dots, P_K$,每个分区中具有相同(或大致相同)数目的传感器节点,节点数为 n ,各分区的簇头数为 $O_1, O_2, \dots, O_K$.规定: P_1 区内的簇头节点与汇聚节点可直接通信,其余分区的簇头节点通过多跳通信方式与汇聚节点通信. $P_1, \dots, P_{K-1}$ 分区内的簇头节点不仅要管理簇内节点,融合各节点数据并传输至下一跳节点,还要承担其他分区的数据中继任务,而在 P_K 分区内的节点则不需负担数据中继的责任.因此,定义 $P_1, P_2, \dots, P_{K-1}$ 内的簇头节点为完全责任簇头节点, P_K 内的簇头节点为部分责任簇头节点.对属于 p 分区任意传感器节点 m ,簇头竞争选举策略如图 1 所示.

3 CEB-UC 算法分析

3.1 对各分区簇头节点数 O_i 的分析

在 T_{data} 阶段,一个部分责任簇头节点的能耗为

$$E_{\text{CH}(K)} = l \left[E_R \left(\frac{n}{O_K} - 1 \right) + E_{\text{DA}} \frac{n}{O_K} + E_T \frac{\alpha n}{O_K} \right] = l \left[E \left(\frac{n}{O_K} - 1 \right) + E_{\text{DA}} \frac{n}{O_K} + (E + e_f D^4) \frac{\alpha n}{O_K} \right] \quad (3)$$

式中: D 是簇头节点与下一跳节点之间的距离; α 是数据融合率; E_{DA} 是簇头融合一个数据包的能耗.

P_K 分区内的任意簇内节点只需将数据发送至簇头节点即可,其能耗为

$$E_{\text{non-CH}(K)} = lE_T = l(E + e_n d_K^2) \quad (4)$$

式中: d_K 是节点与所在簇的簇头节点的距离.那么,根据文献[3], d^2 的数学期望值

$$E(d^2) = \iint (x^2 + y^2) \rho(x, y) dx dy = \iint r^2 \rho(r, \theta) dr d\theta \quad (5)$$

设分区的面积为 S_p ,将每个簇覆盖的区域看成是一个以簇头为圆心的圆盘,由此可得圆盘的半径 $R = (S / \pi K O_K)^{1/2}$,解式(2)得

```

1. Form_HELLO_Msg(ID, E_residual_m, Cell_Info_ID)
2. Broadcast_Msg(HELLO_ADV_MSG)
3. Receive_Msg(HELLO_ADV_MSG →
   Neighbour_Info_Table)
4. Begin
5. For each node i1 in Neighbour_Info_Table
6.   E_ave = E_ave + Neighbour_Info_Table[i1].E_residual_i1
7. End For
8. E_ave = E_ave / Neighbour_Info_Table.Length
9. If (E_residual_m < E_ave) End_Procedure
10. Else ΔE = (E_residual_m - E_ave) / (E_initial_m - E_ave)
11. Begin
12.   For each node i2 in Neighbour_Info_Table
13.     If (Neighbour_Info_Table[i2].
        Cell_Info_ID = p-1)
14.       np-1++;
15.     Else If (Neighbour_Info_Table[i2].
        Cell_Info_ID = p+1)
16.       np+1++;
17.     Else Continue
18.   End For
19. Δn = |np+1 - np-1|
20. Wm = w1ΔE + w2(np+1 + np-1) - w3Δn
21. Broadcast_Node_Weight(Wm, ID)
22. Receive_Node_Weight(W → Node_Weight_Table)
23. Begin
24.   For each node weight i3 in Node_Weight_Table
25.     If (Wm < Node_Weight_Table[i3])
26.       Is_Cluster_Head = False
27.       Broadcast_Msg(HEAD_FAL_MSG)
28.   End For
29. If (Is_Cluster_Head) Mark node m as Cluster_Head
30.   Broadcast_Msg(HEAD_SUS_MSG)

```

图1 属于 p 分区的任意传感器节点 m 的簇头竞争选举策略

$$E(d_k^2) = \int_0^{2\pi} \int_0^R \rho r^3 dr d\theta = \frac{\rho S^2}{2\pi K^2 O_K^2} \quad (6)$$

式中: ρ 代表节点的密度,代入式(6)可得

$$E(d_k^2) = \frac{\rho S^2}{2\pi K^2 O_K^2} = \frac{K n S^2}{2\pi S K^2 O_K^2} = \frac{n S}{2\pi K O_K^2} \quad (7)$$

综上所述,在 P_K 分区内的任意簇经一轮 T_{data} 的能耗为

$$E_{\text{cluster}} = E_{\text{CH}(K)} + \left(\frac{n}{O_K} - 1\right) E_{\text{non-CH}(K)} \approx E_{\text{CH}(K)} + \frac{n}{O_K} E_{\text{non-CH}(K)} \quad (8)$$

若算法使用了数据融合策略,则 P_K 分区的总能耗为

$$E_K = O_K E_{\text{cluster}} \approx n l \left[(2 + \alpha) E + E_{\text{DA}} + \alpha e_f D_K^4 + \frac{e_n n S}{2\pi K O_K^2} \right] \quad (9)$$

其中, D_K 与此时网络分区的簇头节点数 O_K 及簇头的分布 θ 有关^[10],可认为 D_K 是关于 O_K 及 θ 的函数,记为 $D_K = g(O_K, \theta)$.

当式(9)中 E_K 关于参数 O_K 的一阶导数为 0 时, P_K 分区的总能耗 E_K 可取得最小值,此时的 O_K 为最优值,即

$$\widetilde{O}_K = \frac{1}{g(O_K, \theta)} \left(\frac{n S e_n}{4\alpha\pi K e_f g'(O_K, \theta)} \right)^{1/3} \quad (10)$$

若簇头节点均为完全责任的,它们需为其他分区的数据中继消耗额外的能量.不失一般性,考虑 $P_{K-i} (i \in [1, \dots, K-1])$ 分区的情况,可得该分区的总能耗为

$$E_{K-i} = 2nlE + 2 \left(\sum_{p=K-i+1}^K O_p \right) lE + nlE_{\text{DA}} + l \left(\sum_{q=K-i}^K O_q \right) e_f D^4 + nle_n d_{K-i}^2 \quad (11)$$

由于 CEB-UC 算法重视各分区的能耗均衡问题,于是可得到关系: $E_K = E_{K-1} = E_{K-2} = \dots = E_1$.

设 $E_K = E_{K-i}$,可得

$$O_{K-i}^2 \left[2 \left(\sum_{p=K-i+1}^K O_p \right) E + \left(\sum_{p=K-i+1}^K O_p - \widetilde{O}_K \right) e_f D^4 - \frac{n^2 S e_n}{2\pi K \widetilde{O}_K^2} \right] + O_{K-i}^3 e_f D^4 + \frac{n^2 S e_n}{2\pi K} = 0 \quad (12)$$

为了简化计算,在不影响分析结果的情况下,用网络各跳节点距离的平均值

$$\bar{D} = \frac{1}{K} \sum_{n=1}^K D_K \approx \frac{1}{K} (\Phi) \quad (13)$$

代替式(12)中的 D ,并定义如下 3 个符号,即

$$A = e_f \bar{D}^4 \quad (14)$$

$$B = 2 \left(\sum_{p=K-i+1}^K O_p \right) E + \left(\sum_{p=K-i+1}^K O_p - \widetilde{O}_K \right) e_f \bar{D}^4 - \frac{n^2 S e_n}{2\pi K \widetilde{O}_K^2} \quad (15)$$

$$C = \frac{n^2 S e_n}{2\pi K} \quad (16)$$

将式(14)~式(16)代入式(12),可得

$$AO_{K-i}^3 + BO_{K-i}^2 + C = 0 \quad (17)$$

解之可得

$$O_{K-i} = \frac{1}{6A} (4((27A^2C + 2B^3)^2 - 4B^6)^{1/2} - 108A^2C + 8B^3)^{1/3} - \frac{B}{3A} - \frac{2B^2}{9A((27A^2C + 2B^3)^2 - 4B^6)^{1/2} - 108A^2C + 8B^3} \quad (18)$$

3.2 实际应用的限制及 K 的选择

影响系统选择 K 的因素是:网络时延 t_{TD} 和节点的接受度 β_{AD} .

通常,网络时延分为系统时延(SD)和数据传输时延(DTD).在实际应用中,由于存在大量的未知因素,所以很难精确计算出网络时延.一般而言,系统时延与节点的物理设备及操作系统密切相关,主要由传感器模块、微处理芯片和存储器、操作系统3个模块的时延组成.数据传输时延与无线通信模块的收发数据时延和链路传输时延相关,其中传感器节点使用的无线通信设备及媒体访问控制(MAC)协议是主要因素.由于所有的传感器节点同构,所以可认为传感器节点的系统时延是一致的.由此可知,网络时延可定义为以路径跳数 c_h 、相邻节点距离 d_{ij} 为自变量的函数,即

$$t_{TD}(c_h, d_{ij}) = t_{SD}c_h + t_{DTD}(c_h, d_{ij}) \quad (19)$$

同理可得,网络 TD 时延的最大值为

$$\max t_{TD} = t_{SD}K + t_{DTD}(K, \Phi) \quad (20)$$

式中: Φ 表示网络直径.

节点的接受度表示簇头节点在一个 MAC 周期同时接纳及管理的簇内节点的数量.本文选用时分多址(TDMA)协议,则簇头节点负责为簇内所有的传感器节点分配时槽.因此,配置合理的 TDMA 协议可提高节点的接受度.在 CEB-UC 算法中,距离汇聚节点最远的 P_K 分区,其内簇的数量最小,单簇的覆盖面积最大,簇内节点最多,故该分区簇头节点的接受度总是最大的.

在实际应用中,通常对系统的实时性要求比较高,因此网络时延必须控制在合理的范围之内.同理,MAC 协议在限制 P_K 分区的簇头节点接受度的同时,也影响了 K 的选取.本文将在实验分析中给出了 K 的最优值.

4 实验与性能分析

4.1 O_i 的选取实验及分析

实验 1 采用 NS2 离散事件模拟器来构建模拟环境,网络环境配置参数如表 1 所示.

将以上参数代入式(10)、式(18),可推导出 $O_i \in [4, 19]$,其中 $O_6 \in [5, 8]$.为了验证 3.1 节的理论,取 100 次平均数据作为实验结果,并分析如下.

每个分区内簇头节点数量与每个簇的单轮平均能耗之间的关系如图 2 所示.通过分析可知,算法应首先确定 P_6 分区的簇头节点的数量(见图 2a),在 P_6 分区内,当簇头节点数量 $O_6=6$ 时,簇的单轮平

表 1 网络环境参数设置

参数	参数值
网络覆盖区域 S/m^2	100×600
K	6
分区节点数 n	100
节点无线通信距离 D/m	$50 \sim 200$
$e_n/pJ \cdot b^{-1} \cdot m^{-2}$	10
$e_f/pJ \cdot b^{-1} \cdot m^{-4}$	0.001 3
$E/pJ \cdot b^{-1}$	50
L/b	256

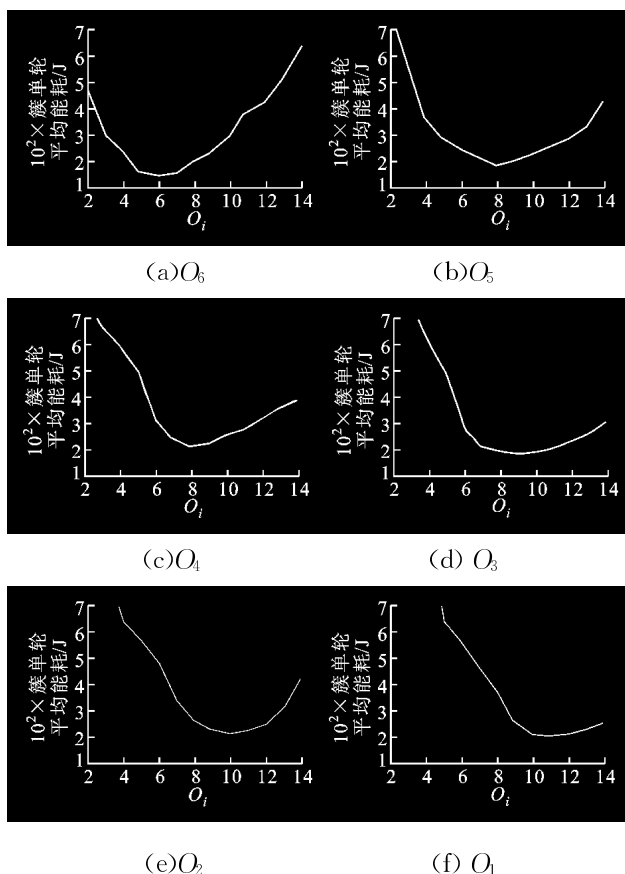


图 2 分区内簇头节点数量与簇的单轮平均能耗的关系

均能耗最小,同样可得各个分区的簇头节点数量依次为 $O_5=8$ 、 $O_4=8$ 、 $O_3=9$ 、 $O_2=10$ 及 $O_1=11$,这些数值均在理论推导的值域范围之内.

4.2 K 的选取及接受度性能实验及分析

实验 2 参数设置见表 1,其中传感器节点数 $N=1\ 200$, K 在 $0 \sim 12$ 区间内变化,实验结果见图 3.

当 $K=0$ 时,算法退化为 LEACH 类算法,此时有部分节点因无法与汇聚节点通信而空转,另有部分节点可与汇聚节点通信,但其传感器节点距汇聚

节点过远,通信能耗过大,所以第一个死亡簇头节点会较早出现.随着 K 增大,第一个死亡簇头节点出现的时间也将逐渐提前.

实验3 考察了簇头节点接受度的最大值选取,参数设置见表1.选择 P_6 分区的簇头节点为考察对象,MAC协议选择 TDMA 协议,实验结果如图4所示.当 β_{AD} 为20时,网络寿命最长;当 β_{AD} 小于20时,分区簇的数量增多,该分区总的维护开销较大;当 β_{AD} 大于20时,簇的时间延迟、信号干扰的概率都将增大,同时簇头与簇内节点之间相互通信的能耗增大,簇头节点能耗增大,分区要求发起簇重组的频率加快,从而加大了分区总能量的开销.

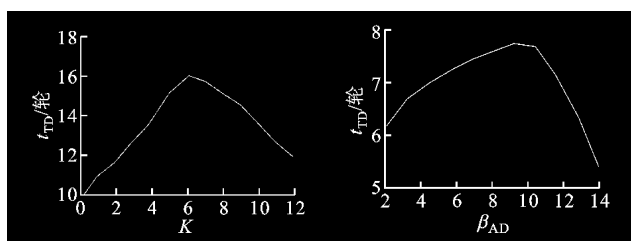


图3 分区数量与网络寿命的关系

图4 节点接受度与网络寿命的关系

4.3 多个分簇协议的实验及结果分析

实验4 将 CEB-UC 与 LEACH-C、EAR-ACM 算法比较,由于 LEACH-C 算法要求网络中的所有节点均可与簇头节点直接通信,故在使用 LEACH-C 算法时,将节点的通信范围设置为 $50\text{ m} \leq D \leq 300\text{ m}$,其余设置见表1.实验4以网络寿命作为性能指标,并定义其为网络运行分簇算法的轮数.在不

表2 传感器节点存活率随网络运行时间的变化

算法	节点存活率/%																	
	轮数为1 900						轮数为2 500						轮数为3 000					
	P_1	P_2	P_3	P_4	P_5	P_6	P_1	P_2	P_3	P_4	P_5	P_6	P_1	P_2	P_3	P_4	P_5	P_6
LEACH-C	91	88	82	0	0	0	59	54	48	0	0	0	15	11	9	0	0	0
HEED	95	93	90	0	0	0	69	63	57	0	0	0	20	17	11	0	0	0
EARACM	100	100	100	100	99	97	82	78	71	65	61	55	33	28	22	17	14	9
PEGASIS	100	100	100	99	97	94	79	76	67	61	57	49	27	22	15	11	9	0
CEB-UC	100	100	100	100	100	99	89	88	86	85	84	83	67	64	65	63	59	59

5 结束语

本文提出了一种基于分簇技术的传感器网络拓扑算法,其具有以下优点:①优化了网络中的能耗,

同数据融合策略的影响下,各个分簇算法的运行时间与网络中存活节点数的关系如图5所示.从中看出,CEB-UC 算法在多种数据融合下,比 LEACH-C、EAR-ACM 算法的网络寿命更长、能耗更低.

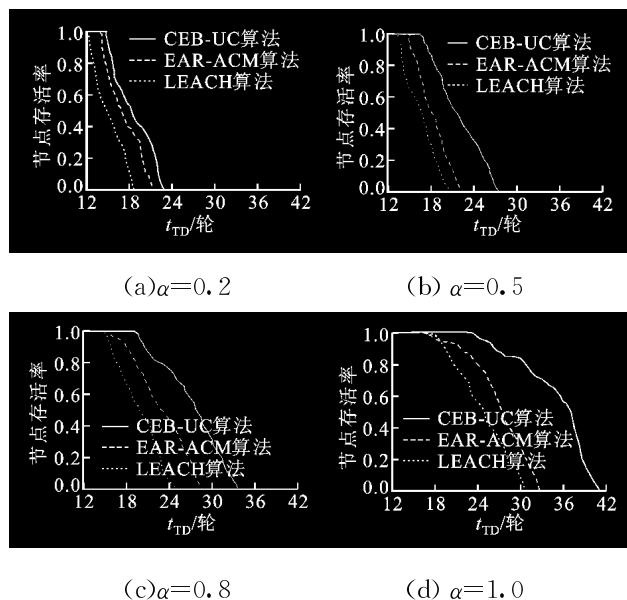


图5 不同 α 时网络运行时间与节点存率的关系

实验5 考察了网络中传感器节点存活率随网络运行时间的变化.在使用 HEED、LEACH-C 算法时,将节点的通信范围设置为 $50\text{ m} \leq D \leq 300\text{ m}$,其余设置见表1,实验结果如表2所示.在节点的存活时间这一指标上,CEB-UC 算法均优于其他4种经典算法.由此可以看出,CEB-UC 算法不仅显著延长了网络的存活时间,且能很好地均衡节点的能耗.

解决了“热区”问题;②分簇算法稳定,形成簇的数目波动较小;③与使用单跳数据传输模式的协议相比,显著改善了网络的扩展性,与其他多跳数据传输模式的协议相比,有效延长了网络的存活时间.

参考文献:

- [1] AKYILDIZ I F, SU W. Wireless sensor networks; a survey [J]. *Computer Networks*, 2002, 38(4): 393-422.
- [2] HEINZELMAN W, CHANDRAKASAN A, BALAKRISHNAN H. An application specific protocol architecture for wireless microsensor networks [J]. *IEEE Transactions on Wireless Communications*, 2002, 1(4): 660-669.
- [3] HEINZELMAN W. Application-specific protocol architectures for wireless networks [D]. Boston, USA: Massachusetts Institute of Technology, 2000.
- [4] TILLAPART P, THUMTHAWATWORN T, PAKDEEPINIT P et al. Method for cluster heads selection in wireless sensor networks [C]// *Proceedings of the 2004 IEEE Aerospace Conference*. Piscataway, USA: IEEE, 2004: 3615-3623.
- [5] YOUNIS O, FAHMY S. Distributed clustering in ad-hoc sensor networks; a hybrid, energy-efficient approach [J]. *IEEE Transactions on Mobile Computing*, 2004, 3(4): 660-669.
- [6] LINDSEY S, RAGHAVENDRA C S. PEGASIS: power-efficient gathering in sensor information systems [C]// *Proceedings of the IEEE Aerospace Conference*. Montana, USA: IEEE Aerospace and Electronic Systems Society, 2002: 1125-1130.
- [7] YI S H, PARK G Y, HEO J Y, et al. Energy aware routing based on adaptive clustering mechanism for wireless sensor networks [C]// *Proceedings of 19th Annual IFIP WG11.3 Working Conference on Data and Application Security*. Berlin, Germany: Springer-Verlag, 2005: 1115-1124.
- [8] MHATRE V, ROSENBERG C. Design guidelines for wireless sensor networks: communication, clustering and aggregation [J]. *Ad Hoc Networks*, 2004, 2(1): 45-63.
- [9] BAHİ B, PADMANABHAN V N. RADAR: an in-building RF-based user location and tracking system [C]// *Proceedings of INFOCOM 2000*. Piscataway, USA: IEEE, 2000: 775-784.

(编辑 苗凌)