

# 基于混合分簇的无线传感器网络路由协议<sup>\*</sup>

李捷 陈志国 李致远

(河南大学计算机与信息工程学院 开封 475004)

**摘要** 提出一种基于混合分簇的无线传感器网络路由协议(EHCRP)。该协议综合了集中式和分布式分簇的优点,使得簇头分布更加均匀并减少了簇重构的开销。在此分簇的基础上,使用改进的多路径路由树建立簇间路由,使其不但可以减少重传次数,提高数据传输的稳定性,还缓解了簇头的能耗。然后使用 CMMBCR 算法建立一条能量高效的簇内路由。仿真结果表明:该协议较之经典的分簇路由协议不仅有效地均衡了网络能量消耗、显著地延长了网络生命周期,而且具有较强的鲁棒性。

**关键词** 混合分簇,多路径路由树,能耗均衡,J-Sim 仿真

## Hybrid Cluster-based Routing Protocol in Wireless Sensor Networks

LI Jie CHEN Zhi-guo LI Zhi-yuan

(College of Computer & Information Engineering, Henan University, Kaifeng 475004, China)

**Abstract** With the application of large-scale Wireless Sensor Networks (WSNs) in recent years, how to further prolong the lifetime of sensors network is a challenge to the research of routing protocol. In this paper, an energy-efficient hybrid clustering routing protocol (EEHCRP) for wireless sensor networks was proposed. The algorithm integrated the advantages of concentrated clustering and distributed clustering. It could make the distribution of cluster head more evenly, and reduce the overhead of cluster reconstruction. On the basis of the clusters, the inter-cluster routing was established through multiple path routing tree and gateways mechanism. The method could reduce retransmission, make data transmit more steadily and alleviate the energy consumption of cluster head. Then an energy efficient intra-cluster routing was also established by CMMBCR algorithm. Finally, EEHCRP was compared to classical clustering routing protocol (LEACH, LEACH-C, HEED) by simulation. Simulation results show that the EEHCRP algorithm can not only effectively balance the energy consumption and further prolong the lifetime of the network, but also have lower overhead, better robustness.

**Keywords** Hybrid clustering, Multiple path routing tree, Energy consumption-balancing, J-Sim simulation

## 1 引言

无线传感器网络(Wireless Sensor Networks, WSNs)路由协议的研究一直是 WSNs 研究的核心问题之一。WSNs 与传统网络在协议设计方面有很大的差异,其中最主要的差别是 WSNs 首要设计目标是能量的高效使用。WSNs 主要撒布在野外且节点数量非常庞大,根据具体应用的要求,这些节点需要在不能补充能量的情况下工作几个月甚至几年。为了延长网络的寿命,采用基于分簇的路由协议是目前解决这一问题的主要方法<sup>[1-3]</sup>。而簇头的产生是簇形成的基础,分簇路由协议的第一步就是考虑怎样产生簇头。目前,按簇头的生成方式主要分为:集中式和分布式。集中式的簇头产生算法是指由基站负责挑选簇头<sup>[3]</sup>;分布式算法又包括两类:一类是由节点根据某个阈值自主决定是否当选簇头<sup>[1]</sup>;另一类是通过节点之间的信息交互动态产生簇头<sup>[2]</sup>。研究发现这两种成簇方法各有优劣,如果综合起来使用,可以达到互补的效果,进一步提高路由协议的效率。

文献[1]提出分布式的分簇路由协议 LEACH(low energy adaptive clustering hierarchy)。由于减少了与基站的直接

通讯的节点数及通讯量,与节点直接与基站通信相比,LEACH 协议降低了约 8 倍的能耗<sup>[1]</sup>。但是 LEACH 算法存在以下问题:(1)簇头分布不均匀,簇首节点负载较大导致能量消耗不均衡;(2)簇重构开销大。文献[2]提出一种完全分布式的分簇协议 HEED(hybrid energy-efficient distributed clustering),该协议根据剩余能量、簇内通讯代价两个参数选择簇头,簇头的产生在有限次迭代内完成。使得簇头在全网分布较均匀,获得较合理的拓扑结构。但是 HEED 协议在成簇阶段要广播大量消息,增大了能量开销。LEACH-C (LEACH-centralized)<sup>[3]</sup>和 LEACH-F (LEACH-fixed)<sup>[3]</sup>都是在 LEACH 基础上做了一些改变,根据全局信息挑选簇头,属于集中式的分簇协议。LEACH-C 是每个节点把自身地理位置和当前能量报告给基站,基站根据所有节点的报告计算平均能量。当前能量低于平均能量的节点不能成为候选簇头。从剩余候选节点中选出合适数量和最优地理位置的簇头集合是一个 NP 问题。基站根据所有成员节点到簇头的距离平方和最小的原则,采用模拟退火算法解决该 NP 问题。最后,基站把簇头集合和簇的结构广播出去。LEACH-F 与 LEACH-C 成簇方法相同,只是基站为每个簇生成一个簇头列表。指

<sup>\*</sup> 本文得到国家 863 计划项目(2007AA01Z478)和河南省自然科学基金(0611054700)资助。李捷 博士,副教授,研究方向为无线传感器网络;陈志国 教授,研究方向为计算机网络、软件工程等;李致远 硕士研究生,研究方向为无线传感器网络。

示簇内节点轮流当选簇头的顺序。一旦簇形成之后,簇的结构就不再改变,簇内节点根据簇头列表依次成为簇头。该协议的优点是无须每轮循环都构造簇,减少了构造簇的开销,但其不能动态处理节点的加入、失败和移动的问题使其不适合真实的网络应用。分布式分簇算法利用局部拓扑信息分簇导致簇头在全网内分布不均匀,能量消耗不均衡。此外,簇不断地重构,实际有些时候是不必要的,最终导致额外开销太大;集中式分簇算法利用全局拓扑信息分簇使簇头分布更加均匀,能量消耗均衡,减少了簇的重构,但该类算法的网络开销大、可扩展性差,不适合大规模的无线传感器网络应用。

因此,本文提出一种基于混合分簇的无线传感器网络路由协议 EEHCRP (an energy-efficient hybrid clustering routing protocol for wireless sensor networks),该协议综合两类分簇协议的优点,克服彼此的缺点,使得簇头分布在全网更加均匀,簇重构的代价更小,能量消耗更加均衡,延长网络的生命周期。传感器节点部署在野外,容易损坏,使得网络拓扑频繁变化。为了提高数据传输的效率,增强路由协议的鲁棒性,使其易从差错中恢复,在上述分簇的基础上利用多路径路由树建立簇间路由,与传统多路径算法不同的是,该算法不再使簇首节点之间直接通讯,而是利用多网关机制,从而缓解了簇头节点的能量消耗,获得了负载均衡的效果。在簇内,使用 CMBCR(ConditionalMMBCR)<sup>[4]</sup> 协议对重叠节点和簇首节点之间建立一条能量之和最小的路由。本文第 2 节介绍网络模型,第 3 节详细描述 EEHCRP 协议的设计,第 4 节是仿真实验,最后是结论。

## 2 网络模型

本文假设  $N$  个传感器节点随机布设在一个  $M \times M$  的正方形区域内,并假设该网络具有以下性质:

- (1)部署后节点不能移动。
- (2)节点的初始能量相同。
- (3)节点能量、计算能力受限,而基站的能量、计算能力不受限。
- (4)基站部署在正方形的左上角,坐标(0,0)。
- (5)节点已获取自己的地理位置信息并把这些信息发送到基站。

Clustering Partition(m)

```

begin
  Init  $V_0 = \{v_1, v_2, \dots, v_m\}_0$ 
   $S_i = \{s_j; d(s_j, v_i) \leq d(s_j, v_j), \forall j \neq i\}$ 
   $D_0 = \frac{1}{n} \sum_i \min_j d(s_i, v_j)$ 
  while  $\epsilon < \epsilon_m$ 
     $V_k = \{v_1, v_2, \dots, v_m\}_k$ 
     $S_i = \{s_j; d(s_j, v_i) \leq d(s_j, v_j), \forall j \neq i\}$ 
     $D_m = \frac{1}{n} \sum_i \min_j d(s_i, v_j)$ 
     $\epsilon = \frac{|D_{m-1} - D_m|}{D_m}$ 
     $k = k + 1$ 
  end
end

```

在  $1000 \times 1000$  区域内随机部署 960 个节点,取其 5%<sup>[4]</sup> 作为算法运行的参数。经过  $k$  次迭代后得到 48 个码本(圆心的位置)及这 48 个簇的划分,如图 1 所示为分簇效果图。加号表示簇首节点,圆点表示传感器节点。

### 3.1.2 簇建立及簇更新

BaseStation 首先运行胞腔划分算法得到图 1 的均匀分簇。基于该网络满足周期性,以下仅以首次分簇后的第一个  $T$  为例说明,算法描述如下:

$\Delta t$  之后,  $S_i$  发送消息 ( $S_{id} E_{current}$ ) 给 CH

到基站。

(6)节点无线发射功率可调,即节点可以根据接收者的距离自动调整其发射功率。

前三项属性是一般无线传感器网络的典型设置,第四项是考虑到节点的地理位置表达,在正方形区域内建立笛卡儿坐标系,坐标原点即为基站,第五项中节点的地理位置信息可通过 GPS、定位算法获取。第六项的假设是为了便于功率控制,高效利用能量。

表 1 符号描述

符号描述	BaseStation 基站	CH 簇头	CHid 簇头 id	$S_i$ 簇成员 节点	$S_{id}$ 节点 id	Clustering Partition 胞腔划分算法
符号描述	$n_c$ 交集区域 节点数	$V_0$ 初始码本	$\gamma$ 簇头能量状态	$E_{ch}$ 连通强度	$\Delta t$ 固定时间间隔	$T$ 周期
符号描述	$N_c$ 交集区域 节点总数	$E_{current}$ 节点当前 能量状态	$E_{ave}$ 整个簇平 均能量	$E_{chave}$ 全网平 均能量	$Q$ 全网的性能质量 分为 4 个等级 (excellent, good, fair, poor)	$N$ 全网的节点数

## 3 EEHCRP 协议设计

为了实现网络负载均衡、可扩展性好、鲁棒性强的目标,把传感器网络的运行时间划分为相同的周期,每个周期分为两个阶段:簇的建立和数据传递阶段。协议描述中所需的符号见表 1。

### 3.1 簇的形成阶段

#### 3.1.1 胞腔划分算法

为了得到分布更加均匀的簇,使用矢量量化中胞腔的相关理论<sup>[5]</sup>找到一组码本和胞腔的划分,使平均失真最小。在该算法中,将传感器节点看成样本矢量,圆心的集合作为码本集合,每一个胞腔中的节点就组成一簇,并用距离来衡量失真,则这样的划分可以使得平均距离最小。设样本矢量集合  $S = \{s_1, s_2, \dots, s_n\}$ ,  $n$  是样本个数,  $s_i$  是一个矢量,  $V = \{v_1, v_2, \dots, v_m\}$  是码本集合,  $m$  是码本的个数,  $v_i$  是与  $s_i$  同维的矢量,  $d(s_i, v_i)$  表示  $s_i$  与  $v_i$  之间的欧几里得距离。具体算法伪码如下:

- \* 基站从集合  $S$  中随机挑选  $m$  个点作为初始码本
- \* 划分胞腔的依据
- \* 平均失真计算
- \*  $\epsilon_m$  是停止门限,  $0 < \epsilon_m < 1$
- \* 第  $k$  次循环的第  $j$  个码本是由第  $k-1$  次循环得到的第  $j$  个胞腔内的所有点的概率中心求得
- \* 前后两次计算的平均失真的相对变化

CH 接受消息 ( $S_{id} E_{current}$ ), 以二维表形式存储节点信息

CH 根据  $E_{current}$  计算  $E_{ave}$ , 每个簇与相邻簇的 CH 之间互相发送本地节点 id 信息, 由 CH 计算与相邻簇之间 id 信息相同的节点数  $n_c$

if  $E_{ch} < E_{ave}$

按照 HEED<sup>[3]</sup> 选簇头的方法进行局部簇更新

else

CH 发送消息 ( $CH_{id} E_{ave} n_c$ ) 给 BaseStation

BaseStation 首先根据  $E_{ave}$  计算出  $E_{chave}$ , 根据  $n_c$  计算  $N_c$ ,

然后计算  $\gamma = N_c / N$

$$Q = \alpha \times E_{chave} + (1 - \alpha) \times \gamma, 0 < \alpha < 1, \alpha \text{ 根据具体的应用}$$

而定

if  $Q \leq \text{fair or poor}$

由基站再次运行胞割划分算法进行全局的簇更新  
else

按照上述方式进行局部簇更新

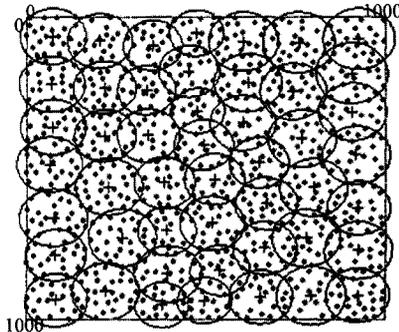


图1 算法运行后的分簇效果图

### 3.2 数据传递阶段

数据的传输是分簇间和簇内两级路由来实现的。节点在传输数据时先根据簇间路由，确定发送的数据要经过哪几个簇才能到达基站，当数据经过具体簇时，根据簇内路由，确定如何将信息传送到下一个簇。

#### 3.2.1 簇间通讯

在生成上述交叠簇的基础上，为簇首节点与基站之间建立一棵具有冗余度的路由树<sup>[6,7]</sup>，使得每个节点具有两个双亲，分别被称为默认双亲和备份双亲。它们均指向基站，并根据动态链路的质量动态切换。图2所示是具有备份链路的路由树模型。为了保证能量有效性，默认双亲采用最小期望传输能量的标准进行选择，备份双亲的选择上增加了信号强度标准。该切换机制是为了使得路由算法具有自适应性，增强路由算法的鲁棒性。图3所示为簇间多路径路由树建立的效果图，其中左上角节点为基站、数据源0号节点为目标区域、其它为簇首节点。对于传统无线传感器网络多路径算法<sup>[6,7]</sup>，是把目标区域采集到的数据根据建立起来的多路径路由树沿默认路径向基站转发，如果默认路径失效，就自动切换到备用路径上转发数据；本文的算法延用了上述传统算法的机制，而与其不同的是，数据不再使用簇首节点直接通讯，而是基于图1的交叠簇，通过簇与簇之间的重叠节点进行中继通讯以缓解簇首的能耗。

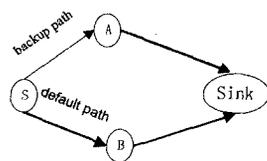


图2 多路径路由树模型

如图1所示，两个簇的交集部分有多个节点称之为重叠节点，重叠节点类似于网关的作用把两个簇连接起来<sup>[8]</sup>。重叠节点承担着簇头之间数据转发的任务，簇头只需负责维护簇内拓扑信息，数据融合以及收集簇内节点的剩余能量值，而无需直接转发数据。这样就避免了传统算法只通过簇头转发数据，从而造成簇头节点能量消耗过快的的问题。由于重叠节

点有多个，可以从中轮流选取能量较大的节点作为中转节点向下一个簇头转发数据。

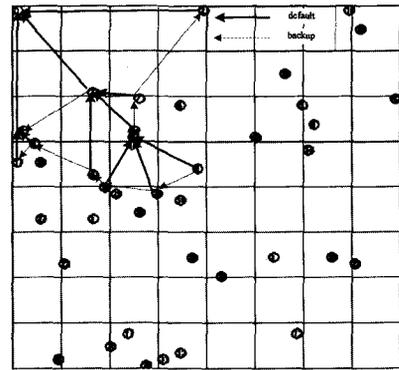


图3 簇间多路径路由树建立示意图

#### 3.2.2 簇内通讯

当数据传递到簇边界的重叠节点时，根据该节点本地存储着簇内拓扑信息及该簇的剩余能量信息。利用 CMMBCR 算法<sup>[4]</sup>找到到达簇首的消耗能量值之和最小的路径，然后经过簇首节点进行数据融合后，由簇首再次运行 CMMBCR 算法选择一条能量最小的路径，通过该路径把数据发送到下一个簇的重叠节点。而如果路径上某节点的剩余能量值小于节点电量的临界值，则在选择路由时尽量避开该节点以延长该簇的生命周期。

## 4 仿真实验

### 4.1 仿真说明

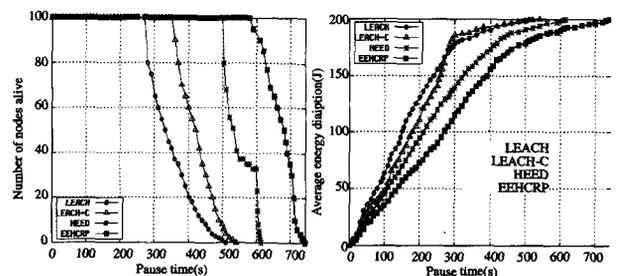
本文采用 J-Sim 仿真平台<sup>[9]</sup>对 EEHCRP, LEACH, LEACH-C, HEED 进行了性能比较。为了获得更加准确的比较结果，均采用文献<sup>[2]</sup>的仿真实验参数，如表2所示。

表2 仿真实验参数表

节点数	100
区域面积	100m×100m
基站坐标	(0,0)
信道带宽	1 Mbps
节点初始能量	2J
发送/接受 1bit 能量消耗	50 nJ/bit
传输放大器的能量消耗	10 pJ/bit/m <sup>2</sup>   0.0013 pJ/bit/m <sup>4</sup>
数据融合的能量	5 nJ/bit/signal
簇的半径	75m~100m
阈值 d <sub>0</sub>	75m
数据包大小	100bytes
广播包大小	25 bytes
包头大小	25 bytes
采样周期	10s

目标每次采样生成的数据量为 4096bytes，分别采用这四种分簇路由协议对数据进行传送。实验中对全网的能耗、生存节点个数、基站收到的数据量在每个采样周期进行了统计。

### 4.2 仿真结果



(a) 生存节点个数

(b) 全网的能耗

NSSA 是一项新兴的研究内容, 仍有许多未知领域需要做更加深入的探讨。就目前而言, 更加有效的 NSSA 的异质多传感器融合算法仍然需要去探索; 安全态势生成算法是单一的态势量化值, 仍然需要讨论更多态势要素, 找到更多的态势生成和威胁因子赋值算法, 从而能够更细粒度地生成网络安全态势; 安全态势生成需要综合分析更多的网络安全要素, 比如系统漏洞等, 而不只是入侵攻击; 除此之外 NSSA 的预测技术也是下一步的重要研究目标。

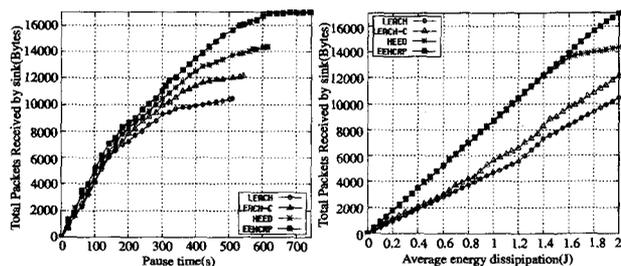
## 参考文献

- [1] Bass T. Multi-sensor Data Fusion for Next Generation Distributed Intrusion Detection Systems//Proceeding of IRIS National Symposium on Sensor and Data Fusion, 1999
- [2] Mica R E. Design and Evaluation for Situation Awareness Enhancement//Proceeding of Human Factors Society 32nd Annual Meeting, Santa Monica, 1988
- [3] Steinburg A N, Bowman C L, White F E. Revisions to the JDL Data Fusion Model//Joint NATO/IRIS Conference, Quebec, 1998
- [4] John J S, Michael L H, Douglas M B. A Situation Awareness Model Applied to Multiple Domains//Proceedings of SPIE, 2005, 5813, 65-74
- [5] Stephen L. The Spinning Cube of Potential Doom. Communications of ACM, 2004, 47(6): 25-26
- [6] Yurcik W. Visualizing NetFlows for Security at Line Speed: The SIFT Tool Suit//19th Usenix Large Installation System Admin-

istration Conference, San Diego, 2005

- [7] Carnegie Mellon's SEL. System for Internet Level Knowledge (SILK). <http://silktools.sourceforge.net>, 2005
- [8] 王慧强, 赖积保, 朱亮, 等. 网络态势感知系统研究综述. 计算机科学, 2006, 33(10): 5-10
- [9] 陈秀真, 等. 层次化网络安全态势量化评估方法. 软件学报, 2006, 17(4): 885-897
- [10] 张慧敏, 等. 集成化网络安全监控平台的研究与实现. 通信学报, 2003, 24(7): 155-163
- [11] 刘伟, 刘鲁. 基于模糊模式识别和 D-S 证据理论的安全态势估计. 计算机工程与应用, 2006, 22: 20-26
- [12] 崔荣一, 洪炳熔. 关于前馈神经网络隐层构建问题的研究. 计算机研究与发展, 2004, 41(4): 524-530
- [13] 周晓东, 邓伟, 陆建德. 多层前馈神经网络的面向对象的程序设计框架. 苏州大学学报, 2006, 26(1): 57-61
- [14] Richard P L, Robert K C. Improving Intrusion Detection Performance Using Keyword Selection and Neural Networks. Computer Networks, 2000, 34: 597-603
- [15] Zhang M, Yao J T. A Rough Sets Based Approach to Feature Selection//Proceeding of the 23rd International Conference of NAFIPS, Banff, 2004
- [16] 陈继军. 多传感器管理及信息融合. 硕士论文. 西安: 西北工业大学, 2002
- [17] Lincoln Laboratory. Darpa Intrusion Detection Evaluation. <http://www.ll.mit.edu>, 1999
- [18] Liu X W, Wang H Q, Liang Y, et al. Heterogeneous Multisensor Data Fusion with Neural Network; Creating Network Security Situation Awareness//Proceeding of ICAIA'07, Hong Kong, 2007

(上接第 34 页)



(c) 基站收到的数据量—时间 (d) 基站收到的数据量—能耗

图 4

图 4(a) 对网络中生存节点的个数进行了仿真, 图 4(b) 对全网的能量消耗进行了仿真, 从图中可以看到, EEHCRP 使得网络的生存周期显著增长, 单位时间内全网的能耗更低。从而验证了本文使用混合分簇可以减少簇重构带来的开销, 引入胞腔划分算法及多网关机制可以更好地均衡节点之间的能量消耗, 从而延长网络的生命周期这一思想。图 4(c) 和 (d) 从时间和能量角度分别仿真了基站收到的数据量, 以此来表征协议的鲁棒性。从图 4(c) 中可以看出单位时间内使用 EEHCRP, 基站收到的数据量更大; 从图 4(d) 中可以看出单位能耗内, 使用 EEHCRP 进行数据传输时, 其性能明显优于 LEACH 和 LEACH-C, 略优于 HEED。这表明 EEHCRP 具有较强的鲁棒性, 这主要是由于本文使用多路径路由树建立簇间路由, 使其在链路发生故障时可以在很短的时间内切换到备用路径继续传输采集到的数据。此外频繁的簇重构也会对基站收到的数据量有影响, 显然 EEHCRP 所采用的混合分簇减少了簇重构的次数, 提高了基站接收到的数据量。

**结束语** 分簇路由协议是目前无线传感器网络路由协议研究的主流方向, 而如何选举簇头, 建立一个什么样的簇以及在生成簇的基础上如何建立簇间、簇内路由是影响网络生命

周期的关键因素, 也是目前研究的热点。本文针对上述问题, 提出一种基于混合分簇的无线传感器网络路由协议 (EEHCRP)。EEHCRP 使用胞腔算法初始化簇, 之后, 周期性地根据簇的状态和全网的状态进行局部簇更新和全局簇更新。然后在生成簇的基础上, 利用多路径路由树及多网关机制建立簇间路由, 使用 CMBCR 算法建立簇内路由。仿真实验表明, 该方法与同类算法相比不仅有效地均衡了网络能量消耗, 延长网络生命周期, 而且具有较强的鲁棒性。

## 参考文献

- [1] Heinzelman W R, Chandrakasan A, Balakrishnan H. Energy-efficient communication protocol for wireless micro networks // IEEE Proceedings of the Hawaii International Conference on System Sciences, 2000: 1-10
- [2] Younis O, Fahmy S. Distributed clustering in ad-hoc sensor networks: A hybrid energy efficient approach // Proc. 13th Joint Conf on IEEE Computer and Communications Societies (INFOCOM), 2004, 3(4): 660-669
- [3] Heinzelman W. Application - Specific protocol architectures for wireless networks. Ph. D. Thesis, Boston: Massachusetts Institute of Technology, 2000
- [4] Toh CK. Maximum battery life routing to support ubiquitous mobile computing in wireless ad hoc networks [J]. IEEE Communication Magazine, 2001, 39(6): 138-147
- [5] Nasrabadi N M, King R A. Image coding using vector quantization: a review [J]. IEEE Transaction on Communications, 1988, 36(8): 957-971
- [6] Wang Y-H, Tsai C-H, Mao H-J, et al. An Energy-Efficient Hierarchical Multiple-Choice Routing Path Protocol for Wireless Sensor Networks [C] // Proceedings of the IEEE International Conference on SUTC'06
- [7] 刘明, 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
- [8] 于鹏程, 张华忠, 刘志杰. 基于聚簇的多跳路由协议的研究 [J]. 计算机应用, 2007, 27(2): 351-354
- [9] Sobehi A, Chen W-P, Hou J C, et al. J-sim: a simulation environment for wireless sensor networks [C] // Proceedings of the 38th IEEE Annual Simulation Symposium (ANSS'05). [S. I.]: IEEE Press, 2005: 175-187