

文章编号:1001-9081(2007)01-0001-03

传感器网络中一种基于质心的分布式成簇算法

姜少峰, 杨明花, 宋瀚涛, 吴正宇, 王捷民

(北京理工大学 计算机科学技术学院, 北京 100081)

(jiangshf@bit.edu.cn; mcl413@163.com)

摘要: 在 LEACH 的基础上, 提出了一种适合无线传感器网络的基于质心的分布式成簇算法——CDCS。在 CDCS 中, 每一个节点首先基于最优簇首概率 p_{opt} 自主确定自己是否为临时簇首。然后临时簇首根据收集到的簇内节点信息, 确定簇内近似质心, 并由此动态调整簇内结构, 使得调整后的簇内通信总能耗尽可能小。理论分析和模拟实验表明, CDCS 在保持 LEACH 算法简单性的同时, 可以获得比 LEACH 更好的性能, 优化后的簇首选择策略可以在不同场景下有效延长网络生存时间达 32% ~ 38%。

关键词: 无线传感器网络; 成簇算法; 质心

中图分类号: TP393 **文献标识码:** A

Centroid-based distributed clustering scheme for wireless sensor networks

JIANG Shao-feng, YANG Ming-hua, SONG Han-tao, WU Zheng-yu, WANG Jie-min

(School of Computer Science and Technology, Beijing Institute of Technology, Beijing 100081, China)

Abstract: Based on LEACH, we proposed a novel clustering algorithm Centroid-based Distributed Clustering Scheme (CDCS) for Wireless Sensor Networks (WSNs). In CDCS, each sensor firstly decided whether it was local tentative cluster-heads on its own at any given time with a certain probability p_{opt} . The tentative cluster-head computed the centroid of cluster based on information of sensors within cluster; and then dynamically adjusted the structure of cluster, so that the total energy dissipation within the cluster was minimized. Theoretical analysis and simulation results show that CDCS prolong the lifetime of a sensor network by 32% ~ 38% over that of LEACH in different scenes while still maintaining the simplicity of LEACH.

Key words: wireless sensor network; clustering algorithm; centroid

0 引言

随着传感器技术、嵌入式技术以及低功耗无线通信技术的发展, 生产具备感应、无线通信以及信息处理能力的微型无线传感器已成为可能。这些廉价的、低功耗的传感器节点共同组织成无线传感器网络(WSNs), 通过节点间的相互协作, 将其监测和感应的多种环境信息(如温度、湿度等)传回到远程基站进行处理。无线传感器网络的随机布设、自组织、环境适应等特点使其在环境监测、办公室智能化、战场监视以及交通流量监测等领域具有广阔的应用前景和很高的应用价值^[1]。由于受到成本、体积等因素的限制, 传感器节点的处理能力、无线带宽, 以及电池容量等资源非常有限。特别是在很多应用中, 由于传感器节点被部署在敌对或者环境恶劣的地区, 节点的能量无法得到补充, 这就使得如何延长传感器网络的生存时间成为设计上需要考虑的关键因素之一。将传感器节点分簇是目前解决这一问题的重要方法。分簇的基本思想是通过簇首对簇内节点间的相关信息融合及转发机制减少数据的传输量和距离, 进而降低通信能量, 达到网络节能的目的。其中, 如何优化簇首的选择是算法设计中所必须考虑的重要问题之一, 它是影响网络能量损耗分布、提高网络生存时间的关键因素。

1 相关工作及评价

传感器网络的主要任务是将网络中传感器节点收集的数据传送给基站。实现该任务的一种最简单方法是使用直接传送^[2], 即网络中的每个节点把收集到的数据直接传送给基站。然而, 对于远离基站的传感器节点, 节点直接传送数据消耗的能量代价太高将使节点很快死亡。为解决这个问题, 目前已经提出了许多以节约能量为目的的分布式成簇算法, 如 LEACH^[3], PEGASIS^[4]等。LEACH 采用所有节点周期性地轮流担任簇首节点的办法来达到使所有节点均匀消耗能量的目的。在 LEACH 中, 每一个节点均根据网络所决定的最优簇首概率 p_{opt} 自主决定是否成为簇首, 并保证每轮有相当数量的簇首节点。随机簇首选择机制具有实现简单、操作灵活和可扩展性好等优点, 但是却不能保证簇内结构的质量, 因此容易形成簇内节点能量损耗不均衡的状态, 从而降低了网络的整体性能。PEGASIS 则将节点组织成链的形式, 链的形成由每一个节点或者基站计算得到, 因此需要知道网络拓扑的全局知识。HEED^[5]也是一种完全分布式的成簇算法, 它随机选择簇首节点, 选举概率与该节点的剩余能量直接相关, 通过降低低能量节点成为簇首的概率来保证网络内能量负载的平均分布, 从而进一步延长网络生存时间。但 HEED 在选择簇首时需要在一定的迭代次数内与周围邻居节点不断地进行信息交互, 因此算法的实现需要额外的通信代价, 不适合大型无

收稿日期:2006-07-10; 修订日期:2006-09-25

作者简介: 姜少峰(1975-), 男, 甘肃兰州人, 博士研究生, 主要研究方向: 无线传感器网络; 杨明花(1977-), 女, 山东临沂人, 博士研究生, 主要研究方向: 数据挖掘; 宋瀚涛(1941-), 男, 吉林大安人, 教授, 博士生导师, 主要研究方向: 无线自组网; 吴正宇(1980-), 男, 博士研究生, 主要研究方向: 无线自组网; 王捷民(1980-), 男, 博士研究生, 主要研究方向: 无线传感器网络。

线传感器网络。DEEC^[6] 基于节点剩余能量与网络节点的平均能量的比例来选举簇头节点。较高初始能量和剩余能量的节点比低能量节点拥有更多的机会成为簇头节点。但和 LEACH 一样,该算法同样不能优化簇内结构。文献[7]提出了一种能量自适应的簇首选择机制,算法通过综合考虑候选节点的剩余能量级、通信能量损耗等参数来优化簇首的选择,从而有效地避免了低能量节点被选为簇首的可能性,但该算法也未能考虑优化簇内结构。

显然,为了尽可能的延长网络寿命,不仅需要实现网络能耗的均匀分布,而且在每一轮的数据收集过程中,也需要尽可能的降低簇内节点和簇首通信的总能耗,从而有效延长网络寿命。不同于上述成簇算法,本文提出的 CDSC 算法是在 LEACH 基础上设计的。它借鉴了 LEACH 的簇首轮转思想,同时优化 LEACH 算法每一轮生成的簇内结构,使得每一轮的能耗尽可能小。

2 系统模型和问题描述

与文献[3]等所提出的算法一致,本文提出的基于簇的 CDSC 算法同样按轮运行。每轮分为簇生成和数据收集两个阶段,分别用 $T_{cluster}$ 和 T_{data} 表示,其中 T_{data} 是网络进行环境监测的阶段,在此阶段中节点收集的数据被传送到基站进行处理。为了保证网络的有效工作时间,算法需要具有 $T_{data} \gg T_{cluster}$ 的性质。

2.1 网络模型

假定 N 个传感器节点随机均匀分布在一个 $L \times L$ 的正方形区域内,并且该传感器网络具有如下性质:(1) 在离传感区域较远处有一个固定的基站;(2) 节点能量有限,具有相似的能力(处理/通信),并对网络具有同等重要性;(3) 每个节点都可以通过单跳方式与基站直接通信;(4) 节点一旦配置完成即保持静止不动;(5) 节点周期性地将收集到的数据传回到远程基站;(6) 通信是对称的;(7) 无线发射功率可控;(8) 节点装备有 GPS,因此每个节点可以感知其具体位置。

2.2 无线通信模型

节点的能量损耗采用文献[3] 中的无线通信模型进行计算。该模型由发送电路、功率放大器和接收电路构成。在该无线通信模型中给出了一个阈值 d_0 (d_0 为常数,具体数值取决于使用环境),当发送节点与接收节点的距离小于 d_0 时,发送方发送数据的能量损耗与距离的平方成正比,否则与距离的四次方成正比。由 d_0 确定的这两种能量衰减模型分别称为自由空间模型和多径衰减模型。因此,根据发送节点和接收节点之间的距离,发送节点可以使用不同的能耗模型计算发送数据所需要的能量。当发送方传输 l 比特信息到距离为 d 的接收方时,可以使用下面的公式计算其能耗:

$$E_{Tx}(l, d) = E_{Tx-elec}(l) + E_{Tx-amp}(l, d) \\ = \begin{cases} lE_{elec} + l\epsilon_{js}d^2, & d < d_0 \\ lE_{elec} + l\epsilon_{em}d^4, & d \geq d_0 \end{cases} \quad (1)$$

相应地,接收方无线接收装置产生的能耗为:

$$E_{Rx}(l) = E_{Rx-elec}(l) = lE_{elec} \quad (2)$$

式中, E_{elec} 表示无线收发电路所消耗的能量。 E_{Tx-amp} 表示放大器消耗的能量,其大小取决于发送节点与接收节点间的距离以及可接受的位错误率。此外,大部分协议和算法都采用了数据融合技术来减少发送和接收的数据量,从而达到节省能量的目的。CDSC 算法同样采用数据融合技术来减少能

量损耗。与 PEGASIS 协议一样,CDSC 算法假定数据融合的能力为 $NK = K$, 其中 N 表示簇内所有节点的数量,而 K 表示数据包长度。数据融合所消耗的能量用 E_{DA} 表示。

2.3 问题描述

在 LEACH 算法中,节能的基本思想是通过随机循环地选择簇首将整个网络的能量负载均匀分配到每个传感器节点中,从而达到降低网络能量消耗、提高网络整体生存时间的目的。LEACH 假设所有节点配置相同的初始能量,并将网络的操作时间分成许多时隙,称为轮(round)。为了在所有网络节点中平衡负载,LEACH 在每一轮中选举大约 $N_{p_{opt}}$ 个簇首节点。

LEACH 的随机簇首选择机制具有算法简单、自适应的特点,而且能够保证簇首所负载的额外能耗平均分配到网络内的所有节点。但是算法并没有说明簇首节点在网络内如何分布,这就容易形成簇内节点与簇首之间通信距离的不平衡。图 1(a) 描述了网络在运行 LEACH 过程中某一轮的情况。从图中可以看出,簇内节点与簇首的通信距离并不均衡,这就使得距离簇首最远的节点需要耗费大得多的能量与簇首进行通信,加剧了簇内能量耗散的不均衡。显然,为了使得每一轮簇内能量损耗尽可能小,簇首应该位于簇内的质心位置。图 1(b) 显示了经过优化后的簇内结构。从图 1(b) 可以看出,优化后的簇内结构不仅使得每一轮的簇内通信代价大大减少,而且使得距离簇首最远的节点和距离簇首最近的节点,其能耗差值也大为减少,由此导致整个网络的能耗进一步减少,网络的生存时间也进一步延长。这也是本文提出 CDSC 算法的动机。

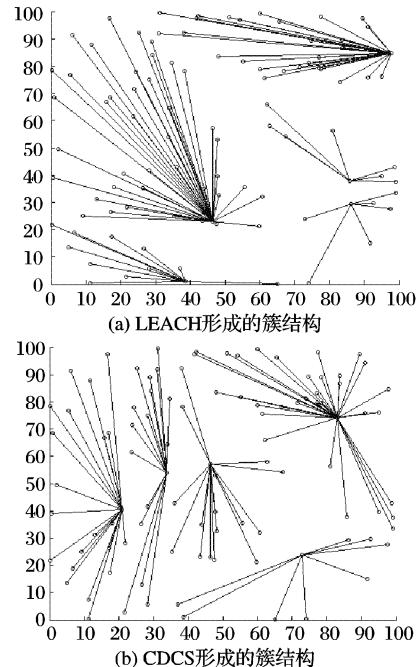


图 1 100 个节点的传感器网络下 LEACH 和 CDSC 形成的簇结构

3 基于质心的 CDSC 成簇算法

CDSC 是一个类似 LEACH 的算法,其基本思想是: 网络中的每个节点在每一轮首先根据最优簇首概率 p_{opt} 自主确定是否为临时簇首, 其他节点选择距离最近的临时簇首作为自己的簇, 然后临时簇首根据簇内总能耗最小化原则, 选择一个最靠近簇内质心的节点作为最终簇首, 最后, 由最终簇首负责把簇内信息传回到远程基站。

定义1 给定有界区域 $L \times L$ 上的一个有限点集 $S = \{p_1, p_2, \dots, p_n\}$, $L \times L$ 中与 S 中的点 p_i 接近超过 S 中任何其他点的点集, 定义了一个多边形区域 $V(p_i)$, 称为 p_i 的 Voronoi 区域。点集 $\{p_i\}_{i=1}^n$ 称为 Voronoi 产生点, Voronoi 区域的边称为 Voronoi 边, 其顶点称为 Voronoi 顶点。集合 $\{V_i\}_{i=1}^n$ 称为 $L \times L$ 的 Voronoi 划分^[8]。

考虑部署在有界区域 $L \times L$ 上的传感器网络 $S = \{s_1, s_2, \dots, s_n\}$ 。由 CDSC 在某一轮生成的临时簇首集合记为 $H = \{h_1, h_2, \dots, h_m\}$, 则将临时簇首节点作为 Voronoi 产生点, 可以得到 $L \times L$ 区域上的唯一 Voronoi 划分。在下文中, 用 BVH_i 表示位于 $V(h_i)$ 内的传感器节点集合, 包括 h_i 本身。

定义2 BVH_i 中任意两个节点间的最大距离称为 BVH_i 的直径, 记为 $Diam(BVH_i)$ 。假设 $d(p_k, p_j) = Diam(BVH_i)$, 则称点 $p = (p_k + p_j)/2$ 为 BVH_i 的近似质心。其中 d 为欧氏距离, p_k 和 p_j 为传感器节点 s_k 和 s_j 的坐标。

3.1 临时簇首选举

设 p_{opt} 表示最优簇首概率, p_i 表示节点 s_i 的簇首轮转周期。在每一轮的开始, 每个节点均根据最优簇首概率 p_{opt} 自主决定是否成为临时簇首。具体的选择办法是: 每个传感器节点随机选择 $0 \sim 1$ 之间的一个值, 如果选定的值小于某一个阈值 $T(s_i)$, 那么这个节点成为临时簇首。 $T(s_i)$ 值计算如下:

$$T(s_i) = \begin{cases} \frac{p_{opt}}{1 - p_{opt}(r \bmod \frac{1}{p_{opt}})}, & s_i \in G \\ 0, & \text{其他} \end{cases} \quad (3)$$

其中, r 为当前的轮数, G 为在最近 $r \bmod (1/p_{opt})$ 轮中没有成为簇首的节点集合。当临时簇首选举完成, 每一个临时簇首向全网广播一个 $TENTATIVE_CH$ 消息, 网络中的其他节点根据接收信息的信号强度确定距离自己最近的临时簇首作为自己从属的簇, 并发送 $JOIN$ 消息通知相应的临时簇首, 消息中包含该节点 ID 和其位置信息 (x, y) 。设 $H = \{h_1, h_2, \dots, h_m\}$ 为在某一轮生成的临时簇首集合, 则将 H 作为 Voronoi 产生点, 可以得到 $L \times L$ 区域上的唯一 Voronoi 划分。图 2 是 CDSC 在某一轮生成的临时簇首集合形成的 Voronoi 划分, 其中黑心实圆圈表示临时簇首, 位于与其关联的 Voronoi 区域内的其他节点为簇内成员。在有界 Voronoi 划分中, 每个 Voronoi 多边形都是凸多边形。由此, 每个临时簇首均通过其 $V(h_i)$ 可以得到其相应的 BVH_i 。需要指出的是, 临时簇首在簇内的位置并不重要, 因为我们随后将根据收集到的簇内信息选择一个最终簇首, 并由最终簇首负责簇内信息的数据融合和转发。

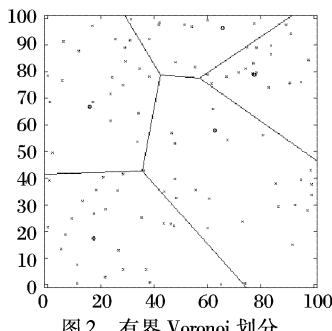


图 2 有界 Voronoi 划分

3.2 基于质心的簇首选举

当临时簇首得到其相应的 BVH_i 值后, 它根据 BVH_i 计算簇的近似质心。这里首先利用文献[9]中计算凸包直径的算

法得到 $Diam(BVH_i)$, 由此可以计算出每个簇的近似质心 p 。然后, 临时簇首根据簇内节点到质心 p 的距离, 确定具有到质心 p 距离最小且在最近 $r \bmod (1/p_{opt})$ 轮中没有成为簇首的节点作为最终簇首, 并向选择好的节点发送 $FINAL_CH$ 消息, 该消息指出最终选择的簇首。此后, 收到 $FINAL_CH$ 消息的节点向簇内成员广播一个 $COMPLETE_CH$ 消息, 簇内成员收到该消息后, 根据接收到的信号强度调整向最终簇首发送信息需要的发送功率, 此时成簇过程结束。在随后的数据传输过程中, 簇内成员按照簇首通过 TDMA 方式为簇中每个节点分配的时间片, 向该簇首发送收集到的传感信息。具体的簇内信息传送过程和 LEACH 类似, 这里不再赘述。

3.3 性能分析

定理1 CDSC 算法最多需要两次迭代即可终止。

证明 当临时簇首选举完成后, 临时簇首需要向全网广播一个 $TENTATIVE_CH$ 消息来确定簇内成员。此后, 临时簇首根据簇内质心确定最终簇首后, 需要再发送一个 $FINAL_CH$ 消息, 因此, 整个算法只需要两次迭代即可终止。

定理2 CDSC 算法的消息复杂度为 $O(N)$, 其中 N 为网络中的节点数。

证明 CDSC 算法中, 临时簇首需要发送 $TENTATIVE_CH$ 和 $FINAL_CH$ 两个消息, 最终簇首需要发送一个 $SCHEDULE$ 消息为簇中每个节点分配相应的时间片, 而簇内其他节点只需要发送一个 $JOIN$ 消息。显然, 算法在每一轮中需要发送 $2Np_{opt} + Np_{opt} + N(1 - p_{opt})$, 即 $N(1 + 2p_{opt})$ 个消息, 因此算法的总的消息复杂度为 $O(N)$ 。

由上述分析可以看出, 算法只需要两次迭代即可终止, 因此收敛速度是很快的。和 LEACH 算法相比, LEACH 算法每轮需要发送 $N(1 + p_{opt})$ 个消息, 而 CDSC 算法每轮需要发送 $N(1 + 2p_{opt})$ 个消息, 因此和 LEACH 算法的消息复杂度也相同。而且, 由于 CDSC 算法优化了簇内结构, 使得簇内节点的能耗分布更为均匀。

4 模拟实验及结果分析

为了评估算法性能, 我们用 C++ 语言在 Linux 操作系统下设计了一个模拟程序。考虑 100 个节点的无线传感器网络, 所有节点随机地分布在一个 $100m \times 100m$ 的正方形区域内, 远程基站位于坐标 $(x = 50, y = -100)$ 处。在和 LEACH 进行比较时, 忽略了由信号冲突和无线信道干扰等随机因素带来的影响。为了把网络中节点随机配置的影响最小化, 以下所有模拟结果均取 100 次模拟的平均值。

为了考察不同初始能量对网络性能的影响, 根据每个节点携带的初始能量设计了三组实验。无线传感器网络模型的主要参数为: 模拟的开始阶段, 每个节点的初始能量均为 0.25J、0.5J 或 1J; 网络内簇首的最优簇首选择概率 p_{opt} 为 5%; 在实验中使用的其他参数见表 1。

表 1 实验参数

参数	值	参数	值
E_{elec}	50nJ/bit	E_{DA}	5nJ/bit/message
ϵ_{fs}	10pJ/bit/m ²	d_0	70m
ϵ_{mp}	0.0013pJ/bit/m ⁴	Message Size	2000bits

网络生存时间是评价传感器网络性能的一个关键指标, 因此在模拟实验中, 采用网络生存时间作为算法性能的评价

(下转第 6 页)

连通度簇头间的时候,它们往往因为无法与其中任何一个簇头通信而成为孤立簇头。通过采用公式(1)计算得出的 R 值,我们得到较为理想的拓扑结构。更进一步,我们研究了通信半径与簇头个数的关系。200 次仿真结果表明随着通信半径的增长,平均簇头个数呈递减趋势(如图 4)。

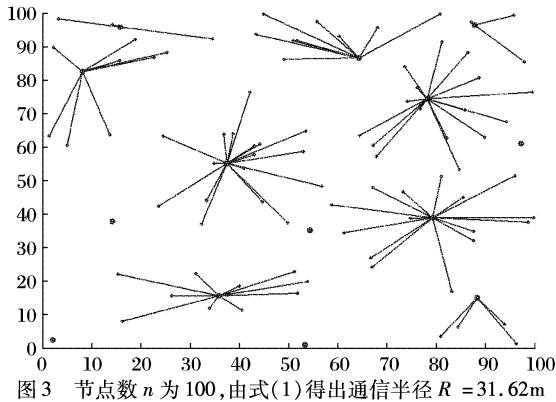


图 3 节点数 n 为 100,由式(1)得出通信半径 $R = 31.62\text{m}$

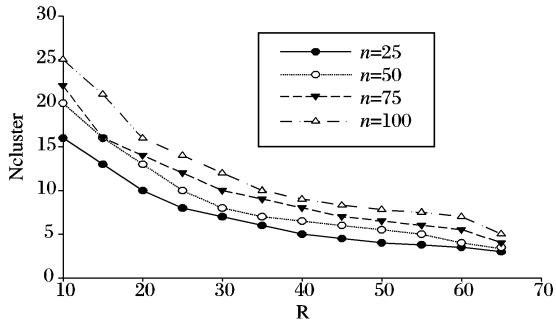


图 4 通信半径和簇头个数的关系

(上接第 3 页)

指标。这里主要考察从开始到第 1 个节点死亡所经过的轮数和整个网络,即最后一个节点死亡时所经过的轮数。模拟结果如表 2。

表 2 节点携带不同初始能量下网络的生存时间

初始能量/J	算法	第一个节点死亡时 网络经历的轮数	最后一个节点死亡 时网络经历的轮数
0.25	LEACH	389	672
	CDSCS	513	907
0.5	LEACH	942	1298
	CDSCS	1281	1765
1	LEACH	1796	2593
	CDSCS	2478	3630

从表 2 可以看出,经过 CDSCS 优化后的簇结构有效延长了网络生存时间,在初始能量为 0.25J、0.5J 和 1J 时,第一个节点死亡的轮数分别延长了 32%、36% 和 38%。此外,网络中最后一个节点死亡时网络经历的轮数也有相应的提高。这一方面证实了 CDSCS 确实实现了网络能量的均衡分布,同时也说明 CDSCS 基于质心的策略减少了每一轮簇内通信的总能耗,从而有效延长了网络的生存时间。

5 结语

在无线传感器网络中,成簇算法是减少能量消耗的一种关键技术,它能够增强网络的扩展性和延长网络的生存时间。在经典 LEACH 算法的基础上,提出了一种适合无线传感器网络的基于质心的成簇算法 CDSCS。由于优化后的簇首选择算法减少了每一轮的能耗,因此有效降低了网络总体能量损

4 结语

EDCC 是一种基于连通度的分布式簇划分算法,具有高效、收敛速度极快等特点,并且通过随机化初始时间的方法使得该算法不依赖于时间同步。实验表明 EDCC 能对网络进行有效分割,适用于大规模传感器网络。

参考文献:

- [1] POTTIE GJ, KAISER WJ. Wireless Integrated Network Sensors[J]. Communications of the ACM, 2000, 43(5): 51–58.
- [2] YU JY, CHONG PHJ. A Survey of Clustering Schemes for Mobile Ad Hoc Networks[J]. IEEE Communications Surveys and Tutorials, First Quarter, 2005, 7(1): 32–48.
- [3] HEINZELMAN WR, CHANDRAKASAN A, BALAKRISHNAN H. Energy-efficient communication protocol for wireless microsensor networks[A]. Proceeding 33rd Annual Hawaii International Conference on System Sciences (HICSS '00)[C]. 2000.
- [4] CHIASSERINI CF, CHLAMTAC I, MONTI P, et al. Energy Efficient design of Wireless Ad Hoc Networks[A]. Proceedings of European Wireless[C]. 2002.
- [5] CHATTERJEE M, DAS SK, TURGUT D. WCA: A Weighted Clustering Algorithm for Mobile Ad hoc Networks[J]. Journal of Cluster Computing, Special issue on Mobile Ad hoc Networking, 2002, (5): 193–204.
- [6] AMIS AD, PRAKASH R. Load-Balancing Clusters in Wireless Ad Hoc Networks[A]. Proceedings of ASSET 2000[C]. 2000.
- [7] SANTI P, BLOUGH DM, VAINSTEIN F. A probabilistic analysis for the range assignment problem in ad hoc networks[A]. Proceeding 2nd ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc '01)[C]. 2001. 212–220.

耗并进一步地延长了网络整体生存时间。模拟实验表明,CDSCS 与 LEACH 相比具有更好的性能,优化后的簇首选择策略在不同的场景下最大可以延长 38% 的网络生存时间。

参考文献:

- [1] AKYILDIZ I, SU W, SANKARASUBRAMANIAM Y, et al. A survey on sensor networks[J]. IEEE Communications Magazine, 2002, 40(8): 102–114.
- [2] ESTRIN D. Next century challenges: Scalable coordination in sensor networks[A]. Proceeding of the MobiCOM'99[C]. 1999. 263–270.
- [3] HEINZELMAN WR, CHANDRAKASAN AP, BALAKRISHNAN H. Energy-efficient communication protocol for wireless sensor networks[A]. Proceedings of the Hawaii International Conference on System Sciences[C]. 2000. 175–187.
- [4] LINDSEY S, RAGHAVENDA CS. PEGASIS: Power efficient gathering in sensor information systems[A]. WILLIAMSON DA, ed. Proceeding of the IEEE Aerospace Conference[C]. 2002. 1125–1130.
- [5] YOUNIS O, FAHMY S. HEED: A hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks[J]. IEEE Transaction on Mobile Computing, 2004, 3(4): 660–669.
- [6] 卿利, 朱清新, 王明文. 异构传感器网络的分布式能量有效成簇算法[J]. 软件学报, 2006, 17(3): 481–489.
- [7] 梁英, 曾鹏, 于海斌. 无线传感器网络中一种能量自适应的簇首选择机制[J]. 信息与控制, 2006, 35(2): 141–146.
- [8] OKABE A, BOOTS B, SUGIHARA K, et al. Spatial tessellations: concepts and applications of voronoi diagram[M]. 2nd. New York: John Wiley & Sons, 1999.
- [9] ALSUWAIYEL MH. Algorithms design techniques and analysis [M]. World Scientific Publishing, 1999.