

无线传感器网络自适应分布式聚簇路由协议

洪榛^{1,2} 俞立^{1,2} 张贵军^{1,2}

摘要 提出了一种适用于无线传感器网络的自适应分布式聚簇路由协议 (Adaptive distributed clustering routing, ADCR). ADCR 协议基于 N 阶近邻理论, 在传感节点部署及网络拓扑动态变化时, 通过对节点分布离散度及曲率变化的分析自动确定当前最优的聚簇数量, 进而根据最优簇数计算最佳簇头占有比率并结合节点剩余能量选取簇头集; 同时引入 Hausdorff 距离调整初始快速形成的簇结构. 仿真结果验证了 ADCR 的有效性, 比 DT, LEACH 和 PEGASIS 拥有更长的生存时间.

关键词 无线传感器网络, 聚簇, N 阶近邻, Hausdorff 距离, 路由

DOI 10.3724/SP.J.1004.2011.01197

An Adaptive Distributed Clustering Routing Protocol for Wireless Sensor Networks

HONG Zhen^{1,2} YU Li^{1,2} ZHANG Gui-Jun^{1,2}

Abstract In this paper, an adaptive distributed clustering routing protocol (ADCR) is proposed for wireless sensor networks. ADCR determines the current optimal cluster number adaptive through dispersion of node distribution and curvature change based on N^{th} -order nearest-neighbor theory when position of sensor node or network topology is changed. Furthermore, with the residual energy of sensor nodes, the optimal proportion of cluster head in network is calculated to elect the cluster head set according to optimal cluster number. Simultaneously, Hausdorff distance is introduced to adjust the initial rapid formation clusters. Simulation results have proved the efficiency of ADCR. It prolongs lifetime more than DT, LEACH and PEGASIS.

Key words Wireless sensor networks (WSNs), clustering, N^{th} -order nearest-neighbor, Hausdorff distance, routing

路由协议是无线传感器网络 (Wireless sensor networks, WSNs) 研究的基本问题, 传统基于 IP 协议的路由机制以及 Ad Hoc 网络中针对移动性较强的多跳路由协议 (如 AODV (Ad hoc on-demand distance vector)) 并不适用于 WSNs^[1]. 因此, 研究适合 WSNs 中特定的路由协议是一项极具挑战的课题^[2]. 现有文献按照网路结构可大致分为三类^[2-3]: 平面路由、基于位置的路由和层次型路由协议. 层次型路由协议被认为比平面路由和基于位置的路由更适用于在 WSNs 中应用, 结合应用分布式聚类算法^[4], 有效解决了传统多跳路由协议的弊端, 如 Heinzelman 等^[5] 提出 LEACH (Low-energy

adaptive clustering hierarchy) 协议以轮机制周期性地执行簇建立和稳定数据通信过程, 等概率随机选取簇头并实施数据融合策略, 在一定程度上降低数据通信量, 有效减少能量消耗, 延长了网络的生存时间.

LEACH 协议引入的簇头融合机制降低了数据传输的冗余度, 改变了传统短距离多跳路由存在的问题, 但也带来新的问题: 1) 未将节点的剩余能量纳入考虑范畴, 导致部分节点很快耗尽能量而缩短网络生存时间; 2) 簇头选取分布不均, 未考虑节点的实际分布, 从而导致网络能耗的不均; 3) 随机选取簇头及成簇方式容易带来大量不确定因素, 导致当前的通信网络并非最优, 进而造成生存时间大幅缩减; 4) 伴随网络规模的扩大, 容易出现能耗不均的现象 (离簇头越远消耗的能量越多, 反之则越少), 从而进一步影响了网络的整体覆盖和生存时间. LEACH-C 和 LEACH-F 这两种涵盖集中式簇形成算法的路由^[5] 解决了上述的部分问题, 但是 LEACH-C 依赖 GPS 与查找算法, 效率较 LEACH 有所降低; 而 LEACH-F 则依赖簇数量的设置以及簇大小的能量成本, 一旦网络发生改变 (新节点的加入或者节点死亡), 其性能指标就迅速下降.

收稿日期 2010-09-10 录用日期 2011-06-05
Manuscript received September 10, 2010; accepted June 5, 2011
国家自然科学基金 (60974017, 61075062, 50908213), 浙江省自然科学基金 (Y1100891), 浙江省科技厅创新团队子项目 (2011R09007-09) 资助

Supported by National Natural Science Foundation of China (60974017, 61075062, 50908213), Zhejiang Provincial Natural Science Foundation of China (Y1100891), and Zhejiang Provincial Science and Technology Innovation Team project (2011R09007-09)

1. 浙江工业大学信息工程学院 杭州 310023 2. 浙江省嵌入式系统联合重点实验室 杭州 310023

1. College of Information Engineering, Zhejiang University of Technology, Hangzhou 310023 2. Zhejiang Provincial United Key Laboratory of Embedded System, Hangzhou 310023

为此, 其他学者在研究 LEACH 的基础上提出了许多有针对性的改进. Lindsey 等^[6] 提出 PEGASIS (Power-efficient gathering in sensor information systems) 协议只保留一个簇头, 利用贪婪策略按照离基站的远近及相邻关系成链, 通过令牌 (Token) 控制方式沿链以最小功率传输融合后的数据, 并最终由簇头把融合后的数据发送给基站. PEGASIS 避免了 LEACH 协议因频繁选取簇头、重构簇带来的通信开销, 较之 LEACH 提高了网络的生存时间, 但也存在诸如单簇遇到网络失败的机率提高、传输时延增大等问题. Manjeshwar 等^[7] 提出了适用于响应型 WSNs 的 TEEN (Threshold sensitive energy efficient sensor network protocol) 协议, 通过设立硬、软两阈值来调节数据的发送量, 达到精度与系统能耗之间的平衡, 实现对突发事件及热点区域的监视, 但也存在阈值设置敏感、容易产生噪声干扰等问题, 不适用于周期性的数据上报. Younis 等^[8] 提出的 HEED (Hybrid, energy-efficient, distributed clustering approach) 协议考虑节点剩余能量, 并依据主从关系引入多个约束条件作用于簇头选取, 同时首次提出将簇内平均可达能量 (Average minimum reach-ability power, AMRP) 作为衡量簇内通信成本, 相比 LEACH 网路拓扑合理, 能耗更均匀, 但却无法保证所选簇头是当前最优 (节点能量最多), 簇头选取的可靠性不高. 刘明等^[9] 提出的基于簇结构的 EADEEG (An energy-aware data gathering protocol for wireless sensor networks) 协议采用簇头竞争机制及簇内节点调度算法, 通过控制活动节点的开闭, 最小化网络通信及达到能量负载均衡, 进而实现网络寿命的延长. 康一梅等^[10] 提出的 LPH (Low-power hierarchical), 利用动静态地址结合的方式提高网络层次及维护性, 同时通过优化拓扑的方式降低簇内通信能耗, 达到延长网络生命的作用. Melese 等^[11] 提出已消耗的能量代替剩余能量作为簇头选取的影响因子以减少功率损耗. 此外, 文献 [12–13] 针对异构 WSNs 给出各自的可行改进算法, 而文献 [14–15] 也根据具体的应用场合提出了各自有针对性的解决方法.

本文在研究各类路由协议的基础上, 针对层次型路由存在的问题, 提出一种适用于 WSNs 的自适应分布式聚簇路由协议 (Adaptive distributed clustering routing, ADCR). ADCR 以 N 阶邻近分析理论为依据, 通过在网络构建及拓扑发生变化时分析节点离散度及寻找曲率变化的拐点, 自适应确定最佳簇数量; 在此基础上结合节点剩余能量选取簇头, 快速成簇后根据 Hausdorff 距离调整簇结构, 最终构建稳定的、低能耗且负载均衡的 WSNs. 最

后, 给出了 ADCR 与直接传输 (Direct transmission, DT), LEACH, 和 PEGASIS 协议的仿真比较结果.

1 网络与能耗模型

1.1 网络模型描述

本文提出的自适应分布式聚簇路由协议隶属于层次型路由, 具体的网络结构如图 1 所示. 假设 N 个传感节点随机地被部署在一个 $M \times M$ 的二维正方形平面区域内, 同时假定该无线传感器网络具有如下性质:

- 1) 属于密度较高的静态网络, 即传感节点部署后静止或微移动, 除非节点失败或者能量耗尽死亡, 其网络拓扑才会改变;
- 2) 基站部署于固定位置, 且是唯一的, 其无线发射功率可控, 能量不受限制 (假设其有外接电源);
- 3) 传感节点的初始能量 E_0 均相同, 且为固定, 在网络通信过程中不能补充;
- 4) 传感节点每轮消耗的能量不一定相同;
- 5) 传感节点有一定的存储空间, 在网络中的作用和地位都相同, 如可以轮流成为簇头, 具备一定的数据融合能力, 且保有唯一标识;
- 6) 网络链路对称, 即节点 $v_1 \rightarrow v_2, v_2 \rightarrow v_1$ 的发送功率是相同的;
- 7) 传感节点未装备 GPS, 因此节点位置是未知的;
- 8) 传感节点发射功率可控, 可以通过调节接收信号强度 (Received signal strength indicator, RSSI) 计算出发送的近似距离, 如 Berkeley Motes, 这里假设如果需要所有节点都能与基站直接通信 (保证协议对比的网络环境相同).

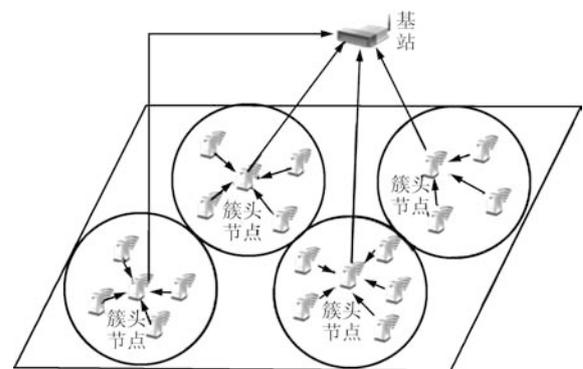


图 1 WSNs 网络结构

Fig.1 The network architecture of WSNs

1.2 能耗模型

本文选用文献 [5] 的无线能耗模型. 当节点 v_1

向距离 d 外的 v_2 发送 L 字节数据包时的能耗为

$$E_{Tx}(L, d) = \begin{cases} 8L \times E_{elec} + 8L \times \xi_{fs} \times d^2, & \text{若 } d \leq d_{\text{threshold}} \\ 8L \times E_{elec} + 8L \times \xi_{mp} \times d^4, & \text{若 } d > d_{\text{threshold}} \end{cases} \quad (1)$$

$E_{Tx}(L, d)$ 包含发射电路及功率放大损耗两部分, 其中, E_{elec} 为节点发射每位 (bit) 数据消耗的能量, 参数 ξ_{fs} 和 ξ_{mp} 依赖于具体的传输放大模型^[5], $d_{\text{threshold}}$ 为阈值, 当传感节点 v_1 与 v_2 的距离 $d \leq d_{\text{threshold}}$ 时, 则使用 Friiss 自由空间模型, 反之当 $d > d_{\text{threshold}}$ 时则采用多路径衰减模型. 特别地, 当 $d = d_{\text{threshold}}$ 时, 则有

$$d_{\text{threshold}} = \sqrt{\frac{\xi_{fs}}{\xi_{mp}}} \quad (2)$$

同时, v_2 接收 v_1 发送过来的数据包消耗的能量为

$$E_{Rx}(L) = 8L \times E_{elec} \quad (3)$$

2 聚簇理论分析

对于随机部署的无线传感节点, 寻找最优的 k 个聚簇是经典的 NP 难问题^[5], 本文提出利用 N 阶近邻分析理论解决此类问题. 假设在指定二维区域内部署 N 个传感节点, 则节点组成的集合记为 S , 且满足 $S = \{x_i | x_i = (a_i, b_i), x_i \in \mathbf{R}^2, i = 1, 2, \dots, N\}$, 通过分析节点集关系, 可以找到满足一定条件的 k 个簇.

定义 1. 部署区域内与节点 x_i 最近的节点的距离称为一阶最短近邻距离, 记作 d_1^i , 与 x_i 次近的节点的距离称为二阶最短近邻距离, 记作 d_2^i , 以此类推, 节点 x_i 的 j 阶最短近邻距离即为 d_j^i , 且 $1 < j < N$. 按照节点 x_i 的最短近邻距离从小到大的排列有:

$$d_1^i < d_2^i < \dots < d_{N-1}^i, \quad \forall i = 1, 2, \dots, N-1 \quad (4)$$

定义 2. 假设 d_1, d_2, \dots, d_{N-1} 分别为 N 个节点的 j 阶最短近邻距离之和, 且 $j \in [1, N-1]$, 那么 j 阶最短近邻距离 d_j 的平均值记作 \bar{d}_j , 平方均值记作 \bar{d}_j^2 , 分别对应于:

$$\bar{d}_j = \frac{\sum_{i=1}^N d_j^i}{N}, \quad \forall j = 1, 2, \dots, N-1 \quad (5)$$

$$\bar{d}_j^2 = \frac{\sum_{i=1}^N (d_j^i)^2}{N}, \quad \forall j = 1, 2, \dots, N-1 \quad (6)$$

定义 3. 网络中衡量传感节点实际分布结构的度量称为节点离散度, 记作 $\sigma_{N,m}^2$:

$$\sigma_{N,m}^2 = \frac{1}{m} \sum_{j=1}^m (\bar{d}_j^2 - (\bar{d}_j)^2), \quad \forall j, m = 1, 2, \dots, N-1 \quad (7)$$

其中 m 为变量, 表示第 m 个节点.

定理 1. 假设随机抛撒的传感节点服从均匀分布的条件, 则节点离散度 $\sigma_{N,m}^2$ 只与网络部署的节点数量 N 和节点 m 有关, 且满足下列关系式:

$$\sigma_{N,m}^2 = \alpha(N)(m+1)^2 + \beta(N)(m+1) \quad (8)$$

其中, $\alpha(N)$ 和 $\beta(N)$ 分别为

$$\alpha(N) = -\frac{1}{3N^2(N+1)} \quad (9)$$

$$\beta(N) = \frac{3N+1}{6N^2(N+1)}$$

证明. 假定 y_i 为原传感节点 x_i 的次序排列, 且满足 $y_1 \leq y_2 \leq \dots \leq y_{N-1}$ 的关系, 同时令 \tilde{y}_i 为其对应独立的且具有相同概率的概率密度函数 $f(\tilde{y}_i)$, 且必须满足 $\tilde{y}_i \geq 0$. 根据次序统计的原理^[16], 计算得到概率密度的 K 阶统计 $f_{N-1,K(y)}$:

$$f_{N-1,K(y)} = \frac{d}{dy} F_{N-1,K(y)} = \frac{d}{dy} P(K(y) \leq y) \quad (10)$$

其中, $F_{N-1,K(y)}$ 为 y_i 的累计概率分布函数, $P(K(y) \leq y)$ 为随机试验 K 阶次序传感节点集中小于等于随机变量 y 的概率. 相应的, 上述过程可以看成在 $N-1$ 次试验中大于等于 K 次成功的事件概率, 由于 y_i 的概率密度函数相同, 且满足关系 $y_1 \leq y_2 \leq \dots \leq y_{N-1}$, 则式 (10) 可转化写成

$$f_{N-1,K(y)} = \frac{d}{dy} \sum_{j=K}^{N-1} \frac{(N-1)!}{j!(N-1-j)!} F(y)^j \times (1-F(y))^{N-1-j} \quad (11)$$

将式 (11) 进一步推导可得 K 阶统计的概率密度 $f_{N-1,K(y)}$ 为

$$f_{N-1,K(y)} = \frac{(N-1)!}{(K-1)!(N-K-1)!} F(y)^{K-1} \times (1-F(y))^{N-K-1} f(y) \quad (12)$$

其中, $F(y)$ 是对应的累积概率分布函数:

$$F(y) = \int_0^y f(t) dt \quad (13)$$

此时, 定义 2 中的式 (5) 和式 (6) 可等价转化成下式:

$$\bar{d}_j \equiv \bar{y}_K = \int_0^\infty y f_{N-1,K}(y) dy \quad (14)$$

$$\bar{d}_j^2 \equiv \bar{y}_K^2 = \int_0^\infty y^2 f_{N-1,K}(y) dy \quad (15)$$

分别将式 (14) 与式 (15) 代入定义 3 中的式 (7), 同时结合式 (12) 和 (13), 在当前部署的传感节点满足均匀分布的条件下 ($f(y) = 1$), 则得

$$\sigma_{N,m}^2 = \frac{1}{m} \sum_{K=1}^m \left\{ \frac{(N-1)!}{(K-1)!(N-K-1)!} \times \int_0^1 y^{K+1}(1-y)^{N-K-1} dy - \left[\frac{(N-1)!}{(K-1)!(N-K-1)!} \times \int_0^1 y^{K+1}(1-y)^{N-K-1} dy \right]^2 \right\} \quad (16)$$

进一步, 由 Beta 函数及其性质 $B(p, q) = \int_0^1 x^{p-1} \times (1-x)^{q-1} dx = \frac{(p-1)!(q-1)!}{(p+q-1)!}$, 代入式 (16) 可得:

$$\sigma_{N,m}^2 = \frac{1}{N^2(N+1)} \left(\frac{N}{m} \sum_{K=1}^m K - \frac{1}{m} \sum_{K=1}^m K^2 \right) \quad (17)$$

利用求和公式与平方求和式, 可将式 (7) 转化为下式:

$$\sigma_{N,m}^2 = -\frac{1}{3N^2(N+1)}(m+1)^2 + \frac{3N+1}{6N^2(N+1)}(m+1) \quad (18)$$

显然, 从上式可得式 (8) 及式 (9). \square

定义 4. 设 $m \geq 2$, 在统计传感节点曲率变化时, 如果在某一节点上曲率变化满足

$$\left| \frac{\sigma_{N,m+1}^2}{m+2} + \frac{\sigma_{N,m-1}^2}{m} - 2 \frac{\sigma_{N,m}^2}{m+1} \right| < \varepsilon \left| \frac{\sigma_{N,m}^2}{m+1} \right| \quad (19)$$

则认为在这一节点上发生了跃变, 该节点也被称为跃变节点, 记为 m^* . 其中的阈值 ε 是在 $(0, 1)$ 区间中的足够小的常数, 且上式必须满足 $m \geq 2$ 的情况, 否则将不成立.

定理 2. 若一个网络存在跃变现象, 则该网络中必定存在多个簇 ($k \geq 2$), 且簇数量 k 随着跃变节点个数的增加而单调递增.

证明. 由定理 1 可知节点分布离散度只与节点个数 N 和节点 m 有关且满足式 (8), 对其做线性化处理得到:

$$\frac{\sigma_{N,m}^2}{m+1} = \alpha(N)(m+1) + \beta(N) \quad (20)$$

根据定义 4, 通过判定近邻节点曲率变化关系, 可确定跃变的发生及获取跃变节点.

不失一般性, 假定网络初始簇数为 $k = 1$, 当计算某一节点曲率时满足式 (19), 则意味着在该节点发生跃变, 簇数在原来的基础上加 1, 即当前必定存在两个簇, 使得在计算统计节点曲率时发生跃变. 迭代遍历节点集, 统计所有的跃变节点 m^* , 就可得到最终的簇数量, 其值必大于等于 2. 而且随着跃变节点的增加, 簇数 k 也必定为单调递增. \square

定理 3. 在 t 时刻, 如果网络节点不发生死亡 (或者网络拓扑不发生改变), 那么统计曲率变化获取的最终簇数量 k 即为这一时刻的最优簇数 k_{opt} .

证明. 定义 3 说明节点离散度是衡量当前网络中所有节点的分布情况的度量, 而定理 2 表明通过研究近邻点曲率变化可判定跃变信息, 进而统计网络中可划分簇的数量 k . 假设在 t 时刻确未出现节点失败或者死亡, 即网络拓扑未发生任何改变, 可认定在这一时刻 $k \equiv k_{opt}$. \square

3 ADCR 协议

基于定理 1~3, 本文提出一种自适应分布式聚簇路由协议 (ADCR). ADCR 协议仍然采用轮机制执行整个通信过程, 但不同的是每轮将被划分为最优聚簇分析阶段、簇头选取、簇形成以及数据传输四个阶段, 分别用 $T_{analysis}$, T_{select} , T_{form} 和 T_{data} 表示, 如图 2 所示, 四个阶段循环执行, 其中最优聚簇

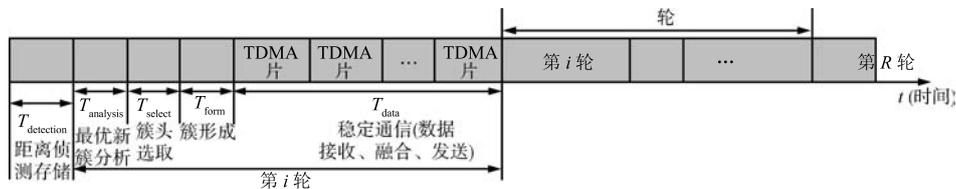


图 2 ADCR 协议通信执行周期

Fig. 2 The communication executable period of ADCR

分析只在网络初始化及拓扑动态改变时(节点失败或者死亡导致的网络拓扑改变)执行,其他情况下将跳过执行.因此,节点在死亡之前也会发送广播消息,基站接收到该消息后即再次触发最优聚簇分析过程.

最优聚簇分析阶段主要运用 N 阶近邻理论分析自动确定当前最优聚簇数量 k_{opt} ,进而引导后续的簇头选举、簇形成朝着期望的方向进行.簇头选取阶段则通过前一阶段得到的最优簇数 k_{opt} 来计算最佳簇头占有比率 p_{opt} ,同时考虑节点的剩余能量,选举当前最优的簇头集.簇形成阶段是在簇头已定的情况下快速成簇,利用 Hausdorff 距离判定簇与簇之间结构的合理性,以此来调整网络的簇划分.数据传输阶段是本轮通信周期的最后一阶段,将被划分成 L 个 TDMA (Time division multiple access) 时间片,节点分别在空闲、发送、接收、睡眠四个状态之间流转,即节点大部分时间仍保持睡眠状态,在一定时间间隔内将被基站唤醒,这样能够减少不必要的能耗.为了充分保证整个通信网络的工作状态和效率,满足关系 $T_{\text{analysis}} + T_{\text{select}} + T_{\text{form}} \ll T_{\text{data}}$.此外,由于网络模型中未设定节点带有 GPS,为了保证最优聚簇分析阶段的正常执行,还必须执行距离侦测存储 $T_{\text{detection}}$,即基站通过发送请求包,并按照接收节点信号强度确定近似距离,形成距离存储表后向各节点转发,该过程只在节点初始化部署后执行一次.

3.1 簇头选取

簇头的选取在本协议中起着关键作用,本文沿用 LEACH 中判定簇头的阈值 $T(n)$. $T(n)$ 决定传感节点成为簇头的概率,而簇头在所有节点中所占百分比 p 则决定着 $T(n)$ 大小,其设定值也直接影响着整个网络的性能. p 值的确定相当困难,往往与当前的网络规模与节点密度相关. LEACH 协议中假设 p 值是由一定的先验知识事先确定的,其随机不稳定性很难适用于实际的 WSNs. ADCR 协议通过前述对传感节点的最优聚簇分析可以得到当前最优的簇数量 k_{opt} ,进而计算当前最优的簇头占用比 p_{opt} :

$$p_{\text{opt}} = \frac{k_{\text{opt}}}{N} \quad (21)$$

则 $T(n)$ 可表示为

$$T(n) = \begin{cases} \frac{k_{\text{opt}}}{N - k_{\text{opt}} \times \left[r \bmod \left(\frac{N}{k_{\text{opt}}} \right) \right]}, & \text{若 } n \in G \\ 0, & \text{否则} \end{cases} \quad (22)$$

其中, r 为执行的轮数, N 为传感节点集的个数, G 为本周期内未当选为簇头的节点集合.当 r 到达一个周期执行时间,未当选为簇头的节点集合 G 将被重置.

考虑传感节点剩余能量对簇头选取乃至整个网络生存周期的影响, ADCR 协议将节点剩余能量纳入簇头选取评价标准,则 $T(n)$ 进一步转变为

$$T(n) = \begin{cases} \frac{k_{\text{opt}}}{N - k_{\text{opt}} \times \left[r \bmod \left(\frac{N}{k_{\text{opt}}} \right) \right]} \times \frac{E_{\text{residual}}}{E_{\text{max}}}, & \text{若 } n \in G \\ 0, & \text{否则} \end{cases} \quad (23)$$

其中, E_{residual} 为节点当前评估的剩余能量, E_{max} 则为节点的最大能量,即节点初始能量. ADCR 协议在引入节点选取簇头最优概率的同时辅以剩余能量决策,使得簇头选举过程更趋于合理,也可推广到异源节点网络中应用.

3.2 簇形成

一旦簇头选取完成,各簇头节点采用基于载波监听多路访问/防止冲突 (Carrier-sense multiple access/collision avoidance, CSMA/CA) MAC (Media access control) 协议及时向整个网络通告自己当选的消息,该消息包内容只包含簇头节点的 ID,同时簇头节点等待其他非簇头节点请求加入. ADCR 协议中的其他非簇头节点无需再依据 RSSI 的强弱来判定到底该加入哪个簇,直接根据初始状态下基站测定并转发的近似距离存储表迅速选择加入最近的簇头,仍按照 CSMA/CA 协议向簇头发回加入请求消息,该消息包含当前的非簇头节点 ID 以及欲加入的请求指令.

快速成簇过程存在许多缺陷,即使前述已经实行了最优聚簇分析和最优簇头选取策略,也无法保证当前簇就是最优的.尤其遇到某一区域内非簇头节点集中、聚簇结构不十分合理的情况下,会导致部分簇头因为承担过大的通信压力而缩短整个网络的生存时间.因此,为使网络划分的簇更趋均匀合理, ADCR 协议引入 Hausdorff 距离^[17]来调整已经初步形成的簇,使得调整后的簇是最优的.

假设 N 个传感节点围绕 k 个簇头初步形成的簇集为 $Clusters = \{C_1, C_2, \dots, C_k\}$, 节点 $v_1 \in C_i$, $v_2 \in C_j$, 且满足 $v_1 \neq v_2$, $C_i \neq C_j$, $1 \leq i \leq k$, $1 \leq j \leq k$, 则簇 C_i 到 C_j 的 Hausdorff 距离为

$$h(C_i, C_j) = \max_{v_1 \in C_i} \min_{v_2 \in C_j} \|v_1 - v_2\| \quad (24)$$

其中, $\|\cdot\|$ 为簇 C_i 中的节点到 C_j 中节点的范数,如 L_2 范数.反过来 C_j 到 C_i 的 Hausdorff 距离不一定

与 C_i 到 C_j 的 Hausdorff 距离相同. 因此, 最终两簇之间的 Hausdorff 距离可以用下式来表示:

$$H(C_i, C_j) = \max(h(C_i, C_j), h(C_j, C_i)) \quad (25)$$

簇头节点在确定簇间的 Hausdorff 距离后, 会将该信息发给各自簇内非簇头节点, 非簇头节点则依据该信息判断自身是否满足 Hausdorff 距离条件, 如果不满足则告诉当前簇头撤销加入. 假设簇 C_i 与 C_j 的 Hausdorff 距离 $H(C_i, C_j) = R_h$, 那么簇 C_i 中每个非簇头节点到 C_j 的节点必须满足 $d(v) < R_h$ (v 代表簇 C_i 中任一非簇头节点), 反之亦然. 当 C_i 中存在一个或者多个节点大于或者等于 R_h , 则微调簇 C_i 与 C_j , 将簇 C_i 中离簇 C_j 最近的节点划归为 C_j , 并且被划分出去的节点同时向 C_j 发送请求加入消息和向 C_i 发送撤销加入消息.

3.3 数据传输

数据传输阶段则发生在簇形成之后, 主要是将各传感节点收集的数据信息通过簇头进一步发送至基站. 非簇头节点基于不同的 TDMA 时间片向各自的簇头节点发送数据, 簇头节点在接收不同的数据后采取融合策略以降低冗余度或还原关键原始数据. ADCR 协议采用的多簇头机制使得整个网络能够分布式同步执行数据传输任务 (网络中多个簇可以同时收发数据), 以提高网络的整体效率. 通常节点会在空闲、发送、接收和睡眠四个状态下流转, 鉴于节省能耗的设计原则, 其大部分的时间处于睡眠状态, 一旦非簇头节点被授予 TDMA 时间片, 即被唤醒转入空闲状态等待发送数据到相应的簇头节点, 而簇头节点也将被唤醒处于准备接收数据的状态. 由于节点主要的能耗都在数据通信上, 数据融合及睡眠状态时消耗极小部分能量, 因此能保证尽可能地延长网络生存时间.

采用分布式聚簇协议也会带来一定的簇间通信干扰, 尤其是邻近簇之间的干扰最为严重. ADCR 协议为减少簇间干扰, 按照特定的网络编码方式, 即基站在获取各簇头向全网广播的消息包后及时分配给簇头节点在 t 时刻的唯一网络 ID, 其他非簇头节点与簇头节点通信时必须在数据包头包含此网络 ID, 簇头节点通过网络 ID 判定数据包是否发送给自身, 这样可以过滤掉大量邻簇的数据包. 同时, ADCR 使用的 CSMA/CA MAC 协议同样可以避免大量通信冲突干扰的发生, 使得全网能够保持稳定的数据通信传输.

4 仿真研究

4.1 仿真参数及评价指标

为验证 ADCR 的有效性, 本文在 Matlab 仿真平台下对其进行评价, 并与 DT, LEACH 以及 PEGASIS 协议进行比较. 网络仿真的具体参数设置详见表 1, 本文假定的场景为 $100\text{ m} \times 100\text{ m}$ 二维正方形区域, 网络节点规模为 $100 \sim 500$ 个节点, 并随时可扩展到更大规模网络; 初始化操作是在该二维正方形区域内随机生成 $100 \sim 500$ 个拥有相同功能和初始能量的节点, 唯一基站固定于坐标 $(50, 50)$, $(50, 125)$ 或者 $(50, 175)$, 这样做是为了评估基站的远近对于协议本身的影响, 以及进一步验证 ADCR 协议在延长网络生存周期上的优势.

表 1 网络仿真实验参数设置

Table 1 Network simulation parameters setup

Parameters setup	Value
Network size	$(100 \times 100)\text{ m}^2$
Node number	100, 200, 300, 400, 500
Base station (Sink node)	$(50, 50)$, $(50, 125)$, $(50, 175)$
Initial energy (E_0)	0.5 J
E_{elec}	50 nJ/bit
E_{fusion}	5 nJ/bit
ξ_{fs}	10 pJ/bit/m ²
ξ_{mp}	0.0013 pJ/bit/m ⁴
$d_{threshold}$	87.7 m
Data packet size	500 bytes
Packet header size	25 bytes
Broadcast packet size	25 bytes
Request packet size	25 bytes
Round	5 TDMA

由于无线传感节点一般被部署应用于特别恶劣的环境, 通常采用一次性电池供电的方式工作, 工作期间能量无法得到补充, 因此如何使得在有限能量的情况下最大化网络的生存周期是评价协议好坏的主要性能指标之一, 即 ADCR 与 DT, LEACH, PEGASIS 协议在相同网络条件下网络生存的最长时间 (轮数), 分别考虑第一个节点死亡时间、半数节点死亡时间以及全部节点死亡时间. 本文设计的 ADCR 协议不仅要延缓第一个节点的死亡时间, 而且还应考虑在第一个节点死亡后随之带来的网络拓扑发生改变的情形下如何有效延长整个网络生存时间, 即如何在剩余有限节点能量下组织起有效的网络, 这也是协议考虑的内容. 同时, ADCR 必须经

受住不同网络节点规模 (节点部署密度) 以及基站远近对于协议的影响, 以体现在不同网络参数条件下 ADCR 与 DT, LEACH, PEGASIS 的对比优势. 此外, 整个网络的平均能量消耗也将是四类协议比较的对象, 因为网络能耗越是均匀就越能体现整个网络的负载均衡, 这也将最终反映到网络生存周期的延长上.

4.2 性能分析

为了保证结果的客观有效性, 消除因为随机生成部署传感节点带来的不确定因素, 本文的仿真结果均取 100 次实验的平均结果. 图 3~5 给出了 ADCR 与 DT, LEACH, PEGASIS 协议在网络规模为 100 节点的环境下分别对应基站位置 (50, 50), (50, 125) 和 (50, 175) 的网络生存周期. ADCR 在基站位置为 (50, 50) 时, 第一个节点死亡时间为 374 轮, 是 DT 的 170%, LEACH 的 277% 以及 PEG-

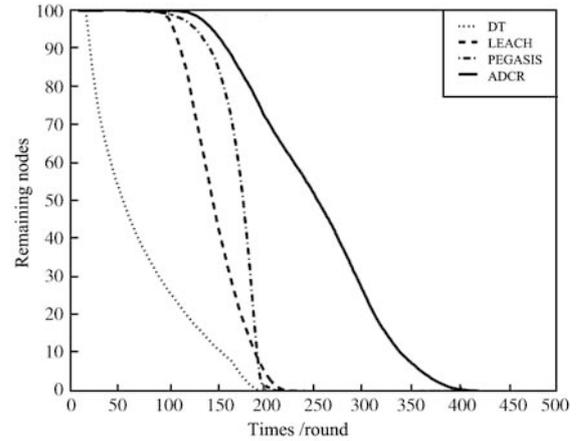


图 5 网络生存周期 — 基站 (50, 175)

Fig. 5 Lifetime at BS (50,175)

ASIS 的 385%; 当全部节点死亡, ADCR 总的生存时间分别为 DT, LEACH 以及 PEGASIS 时的 120%, 201% 和 165%. 而当基站分别位于 (50, 125) 和 (50, 175) 时, 所得到的结果仍与 (50, 50) 相类似. 显然, 从四类协议的网络生存周期结果可以看出, ADCR 协议比 DT, LEACH 与 PEGASIS 拥有更长的网络生存时间. 此外, 基站的位置对于网络生存时间有很大的影响, 甚至能直接导致协议的失败, 这主要是因为离基站越远, 节点发送数据包到基站需要消耗的能量也就越大. ADCR 协议也受到基站远近的影响, 离基站越近生存时间越长, 反之则短, 但生存时间仍然要好于 DT, LEACH 与 PEGASIS.

针对节点部署密度对协议的影响, 分别考虑同一网络情景模式下节点规模 100~500 的死亡情况, 如图 6 和图 7 给出在基站位置 (50, 175) 下第一个节点死亡及全部节点死亡时间统计结果. 不同基站位

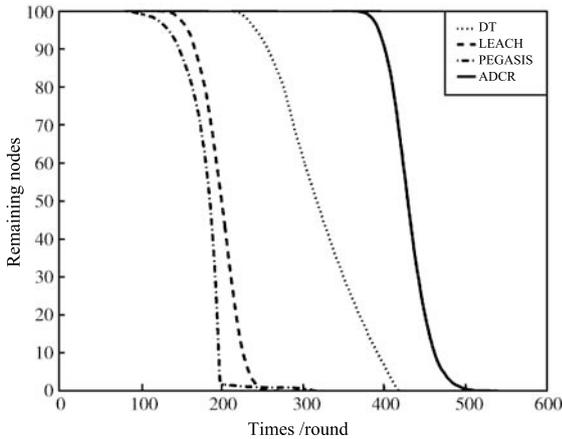


图 3 网络生存周期 — 基站 (50, 50)

Fig. 3 Lifetime at BS (50, 50)

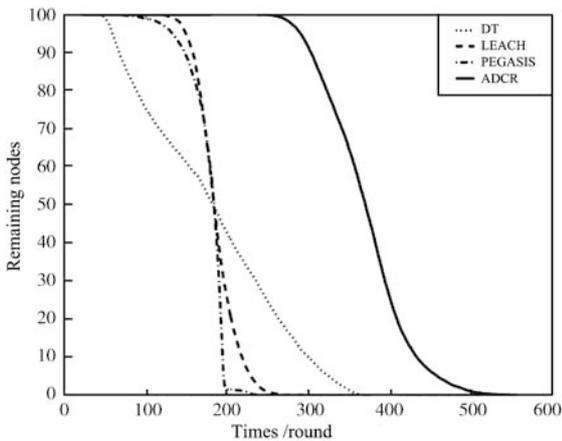


图 4 网络生存周期 — 基站 (50, 125)

Fig. 4 Lifetime at BS (50, 125)

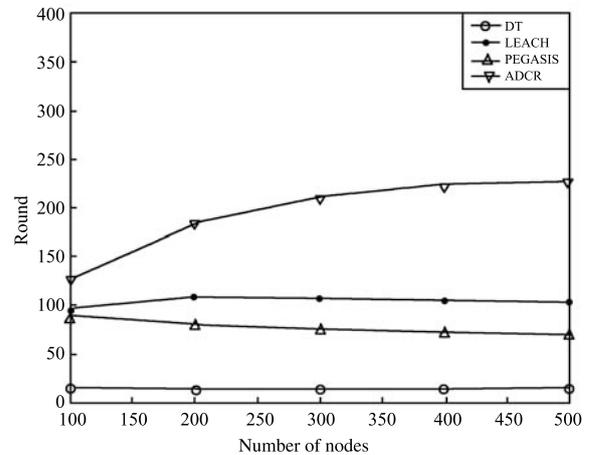


图 6 不同网络规模下第一个节点死亡时间

Fig. 6 First node dead in different node densities

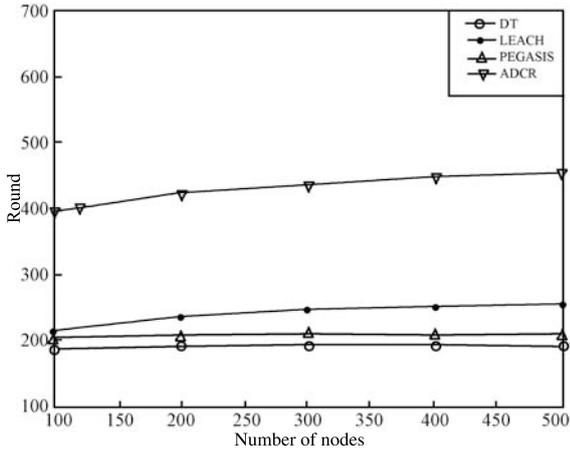


图 7 不同网络规模下全部节点死亡时间

Fig. 7 All nodes dead in different node densities

置的仿真结果表明, 随着网络节点部署的数量增加, 网络生存时间基本保持在一个稳定的浮动空间. 这也证明节点的增加不会加重某些簇头的通信负担, 网络节点部署的疏密与否对于网络生存时间影响较少, 该结果可进一步推广到更大网络场景和更大规模节点的网络.

图 8 给出了在 100 节点网络环境下, 基站位置为 (50, 50) 的基站数据包接收的比较情况 (不同基站位置的情况下得到的结果类似). 从中可以发现, DT 由于没有使用任何融合算法, 导致大量数据包直接发送至基站, 若在实际应用中往往会造成网络拥塞; PEGASIS 由于采用唯一簇头机制, 数据一层一层被融合压缩, 因此最后基站接收的数据包在四类协议中是最少的; ADCR 与 LEACH 基站接收的数据处于中间水平, 因为两者都采用了多簇通信和数据融合算法, ADCR 通过最优聚簇分析及最优簇的形成来保证基站接收的数据包始终处于较优水平.

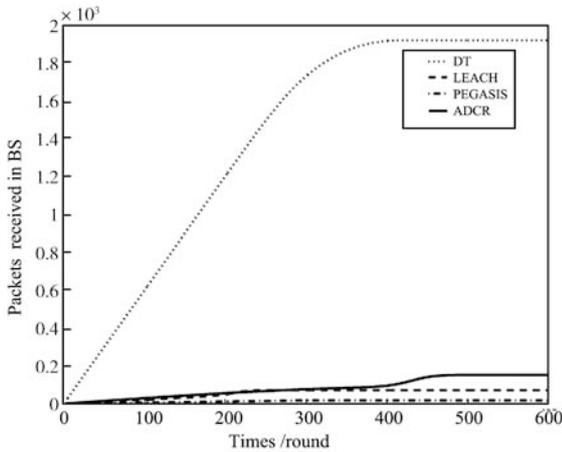


图 8 基站接收的数据包

Fig. 8 Packets received in BS

除此之外, 本文对 ADCR 协议与 DT, LEACH, PEGASIS 在剩余能量及平均能量消耗上也进行对比研究, 如图 9 和图 10 所示, ADCR 在能量消耗方面要远远小于 DT, LEACH, PEGASIS, 这也是 ADCR 能够在相同网络环境下延长生存时间的原因. 相对于 LEACH 协议的不确定成簇与 PEGASIS 的唯一成簇, ADCR 采用的最优聚簇机制有助于优化簇结构, 进而从根本上减少能量的消耗, 从而达到延长网络生存时间的目的. 此外, 在网络节点死亡拓扑发生改变之际, ADCR 同样能够利用最优聚簇机制分析当前最新节点分布, 进而重新决定最优簇数量及重构簇, 使得整个网络始终处于最优通信状态. 上述结果已经得到仿真验证, ADCR 能够在节点发生死亡的情况下, 继续保持良好的工作状态, 有效地延长网络的生存周期, 这在实际应用中显得尤为重要. 后续将运用可行的概率模型以减少 N 阶近邻分析的计算代价, 即通过计算小部分传感节点分布情况预测估计整个网络的分布, 从而进一步降低开销.

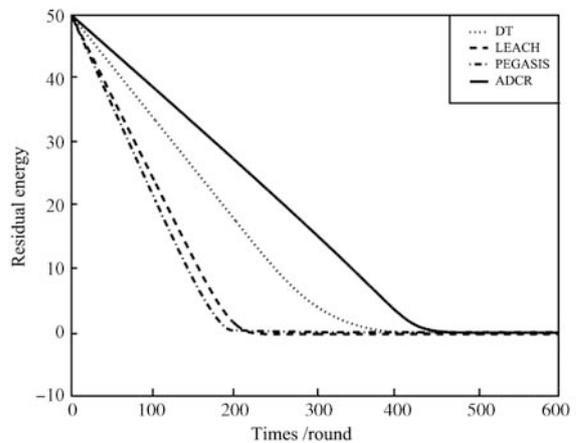


图 9 剩余能量对比

Fig. 9 Residual energy comparison

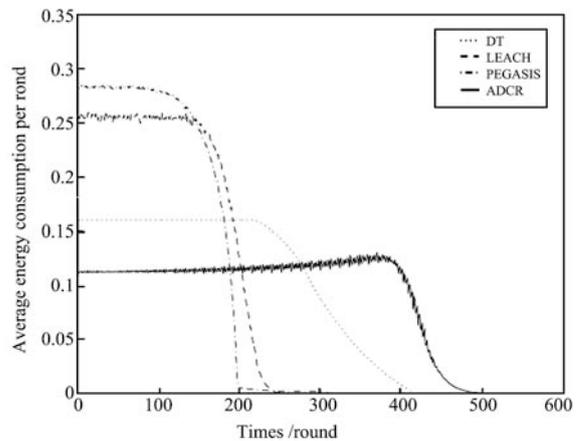


图 10 每轮平均能量消耗

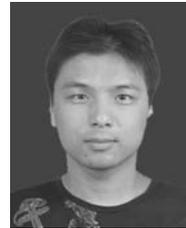
Fig. 10 Average energy consumption per round

5 结论

本文提出的无线传感器网络自适应分布式聚簇路由协议 (ADCR) 基于 N 阶近邻理论, 通过节点分布离散度及曲率变化的理论分析自适应确定最优聚簇数量, 以最优簇数量计算簇头最佳占有比率并结合考虑节点剩余能量确定簇头集, 引入 Hausdorff 距离优化簇形成结构, 减少了通信能量消耗, 延长了网络的生存周期. 仿真实验结果也验证了该协议的有效性, 为进一步投入实际应用打下良好的基础.

References

- 1 Yick J, Mukherjee B, Ghosal D. Wireless sensor network survey. *Computer Networks*, 2008, **52**(12): 2292–2330
- 2 Al-Karaki J N, Kamal A E. Routing techniques in wireless sensor networks: a survey. *IEEE Wireless Communications*, 2004, **11**(6): 6–28
- 3 Sun Li-Min, Li Jian-Zhong, Chen Yu, Zhu Hong-Song. *Wireless Sensor Networks*. Beijing: Tsinghua University Press, 2005
(孙利民, 李建中, 陈渝, 朱红松. 无线传感器网络. 北京: 清华大学出版社, 2005)
- 4 Abbasi A A, Younis M. A survey on clustering algorithms for wireless sensor networks. *Computer Communications*, 2007, **30**(14–15): 2826–2841
- 5 Heinzelman W B, Chandrakasan A P, Balakrishnan H. An application-specific protocol architecture for wireless microsensor networks. *IEEE Transactions on Wireless Communications*, 2002, **1**(4): 660–670
- 6 Lindsey S, Raghavendra C S. PEGASIS: power-efficient gathering in sensor information systems. In: Proceedings of the IEEE Aerospace Conference. Montana, USA: IEEE, 2002. 1125–1130
- 7 Manjeshwar A, Agrawal D P. TEEN: a routing protocol for enhanced efficiency in wireless sensor networks. In: Proceedings of the 15th International Parallel and Distributed Processing Symposium Workshops. San Francisco, USA: IEEE, 2001. 2009–2015
- 8 Younis O, Fahmy S. HEED: a hybrid, energy-efficient, distributed clustering approach for Ad Hoc sensor networks. *IEEE Transactions on Mobile Computing*, 2004, **3**(4): 366–379
- 9 Liu Ming, Cao Jian-Nong, Chen Gui-Hai, Chen Li-Jun, Wang Xiao-Min, Gong Hai-Gang. EADEEG: an energy-aware data gathering protocol for wireless sensor networks. *Journal of Software*, 2007, **18**(5): 1092–1109
(刘明, 曹建农, 陈贵海, 陈力军, 王晓敏, 龚海刚. EADEEG: 能量感知的无线传感器网络数据收集协议. 软件学报, 2007, **18**(5): 1092–1109)
- 10 Kang Yi-Mei, Li Zhi-Jun, Hu Jiang, Dong Ji-Chang. A low-power hierarchical wireless sensor network topology control algorithm. *Acta Automatica Sinica*, 2010, **36**(4): 543–549
(康一梅, 李志军, 胡江, 董吉昌. 一种低能耗层状型无线传感器网络拓扑控制算法. 自动化学报, 2010, **36**(4): 543–549)
- 11 Melese D G, Xiong H G, Gao Q. Consumed energy as a factor for cluster head selection in wireless sensor networks. In: Proceedings of the 6th International Conference on Wireless Communications, Networking and Mobile Computing. Chengdu, China: IEEE, 2010. 1–4
- 12 Kumar D, Aseri T C, Patel R B. EEHC: energy efficient heterogeneous clustered scheme for wireless sensor networks. *Computer Communications*, 2009, **32**(4): 662–667
- 13 Zhou H B, Wu Y M, Hu Y Q, Xie G Z. A novel stable selection and reliable transmission protocol for clustered heterogeneous wireless sensor networks. *Computer Communications*, 2010, **33**(15): 1843–1849
- 14 Liu T, Li F. Power-efficient clustering routing protocol based on applications in wireless sensor network. In: Proceedings of the 5th International Conference on Wireless Communications, Networking and Mobile Computing. Beijing, China: IEEE, 2009. 1–6
- 15 Yan L S, Pan W, Luo B, Li X Y, Liu J T. Modified energy-efficient protocol for wireless sensor networks in the presence of distributed optical fiber sensor link. *IEEE Sensors*, 2011, **11**(9): 1815–1819
- 16 Blekas K, Lagaris I E. Newtonian clustering: an approach based on molecular dynamics and global optimization. *Pattern Recognition*, 2007, **40**(6): 1734–1744
- 17 Zhu X R, Shen L F, Yum T S P. Hausdorff clustering and minimum energy routing for wireless sensor networks. *IEEE Transactions on Vehicular Technology*, 2009, **58**(2): 990–997



洪榛 浙江工业大学信息工程学院博士研究生. 2006 年获浙江工业大学计算机科学与技术学士学位. 主要研究方向为无线传感器网络、优化和聚类分析.

E-mail: hongzhen614@126.com

(HONG Zhen Ph.D. candidate at the College of Information Engineering, Zhejiang University of Technology. He

received his bachelor degree from Zhejiang University of Technology in 2006. His research interest covers wireless sensor networks, optimization, and clustering analysis.)



俞立 浙江工业大学信息工程学院教授. 主要研究方向为鲁棒控制, 时滞系统, 分布式控制和网络控制系统. 本文通信作者. E-mail: lyu@zjut.edu.cn

(YU Li Professor at the College of Information Engineering, Zhejiang University of Technology. His research interest covers robust control, time-delay

systems, decentralized control, and networked control systems. Corresponding author of this paper.)



张贵军 浙江工业大学信息工程学院副教授. 主要研究方向为优化调度和智能交通. E-mail: zgj@zjut.edu.cn

(ZHANG Gui-Jun Associate professor at the College of Information Engineering, Zhejiang University of Technology. His research interest covers optimization schedule and intelligent

transportation.)