

Clustering Topology Control for Multiple Base Stations in WSNs with Q-Learning*

YAN Xinfang*, FENG Yan, WANG Xiaoxiao

(School of Information Engineering, Zhengzhou University, Zhengzhou 450001, China)

Abstract: This paper presents a multiple base stations clustering topology control algorithm to solve the issue of “energy hole” and high network delay near a single base station in wireless sensor networks (WSNs). The number of base stations is determined according to different scenarios. Our algorithm uses a method of graph theory and the gradient of directional diffusion to clustering. To achieve the clustering topology control, our method exploits Q-learning method, incrementally learning at each node’s sufficient network knowledge to choose the best path to base stations. Evaluation of the resulting our algorithm demonstrates its ability to significantly increase the lifetime of network in comparison to the single base station clustering algorithm.

Key words: wireless sensor networks; clustering topology control; multiple base stations; Q-learning

EEACC: 6150P

doi: 10.3969/j.issn.1004-1699.2016.04.019

基于Q学习的多基站分簇拓扑控制算法*

阎新芳*, 冯 岩, 王晓晓

(郑州大学信息工程学院, 郑州 450001)

摘 要: 为了解决无线传感器网络中单基站附近出现的“能量空洞”和网络时延过高等问题, 引入多基站分簇拓扑控制算法。算法根据不同的场景来选择基站数目, 结合图论和定向扩散中梯度的思想对网络进行分簇并运用 Q 学习算法对簇头节点进行周期性的学习训练, 比较到达不同基站的各不同路径上的 Q 值进行最优路径的选择。通过仿真分析表明, 该算法相对于单基站分簇算法可以有效延长网络的生命周期。

关键词: 无线传感器网络; 分簇拓扑控制; 多基站; Q 学习

中图分类号: TP393

文献标识码: A

文章编号: 1004-1699(2016)04-0578-05

无线传感器网络^[1] WSNs (Wireless Sensor Networks) 是由大量的传感器节点部署在监测区域内, 通过节点间的相互通信所组成的多跳网络来感知周围环境的各种信息。由于无线传感网络节点一次性播撒后, 节点能量不可再生, 因此降低网络能量消耗、延长网络生存期成为传感网路由协议的首要设计目标^[2-3]。目前无线传感器网络核心的路由协议是分簇路由协议, 而作为分簇路由协议的基础的多级簇树拓扑结构由于其能量高效和易于维护扩展等特点被广泛研究和应用^[4-7]。

其中基于梯度的分簇拓扑控制算法^[4] ETBG (Energy-Aware Topology Protocol Based on Gradient) 根据节点的通信半径把网络建成一个梯度场, 对

同梯度等级的节点进行分簇。但该算法中节点竞选簇头的的能力仅考虑能量和距离, 生成簇树的过程仅考虑单个簇头节点的权值, 没有综合考虑各方面的因素, 未能找到最优成树路径。而且在大规模的无线传感器网络中, 单基站的 ETBG 算法由于靠近基站的节点承担大量的数据转发任务而造成过多的能量消耗, 从而出现“能量空洞”的问题, 严重的缩短了网络的生命周期。而章等人提出了 Q 学习算法^[8-11] 的路由选择机制, 使数据始终沿着能量消耗代价最小的路径进行数据传输, 在一定程度上避免了使用剩余能量少的节点转发数据, 但是该算法适用于平面路由, 在大规模网络中的分层路由中并不适用。

项目来源: 河南省科技厅基础与前沿研究计划项目 (152300410023)

收稿日期: 2015-10-18

修改日期: 2016-01-18

本文在 ETBG 算法和 Q 学习算法的基础上,提出了基于 Q 学习的多基站分簇拓扑控制算法 (CTQL-MB)。该算法首先根据不同的场景确定相应的基站数目,然后利用 OWA 算子多属性决策的方法^[12]将节点的能量、距离、节点间的相互作用等因素融合为一个属性值来决定节点竞选簇头的能力,利用定向扩散中的梯度思想将网络划分为一个不均匀的梯度图并结合图论中独立集的概念对网络进行分簇,最后运用 Q 学习算法对簇头节点进行周期性的学习训练,比较到达不同基站的不同路径上的 Q 值,选择整条路径上 Q 值最大的路径为最优路径。

1 基于Q学习的单基站分簇算法

1.1 Q 学习

Q-learning 是强化学习的一种常用算法。强化学习系统接受环境状态的输入 s , 根据内部的推理机制, 系统输出相应的行为动作 a 。环境在系统动作作用 a 下, 变迁到新的状态 s' 。系统接受环境新状态的输入, 同时得到环境对于系统的瞬时奖惩反馈 r 。对于强化学习系统来讲, 其目标是学习一个行为策略 $\pi: S \rightarrow A$, 使系统选择的动作能够获得环境奖励的累计值最大。

Q 值的更新公式如式(1)所示:

$$Q(i)^{t+1}(s^t, a^t) = r(i)(s^t, a^t, s^{t+1}) + \gamma \max_{j \in \text{Neighbor}(i)} Q(j)^t(s^{t+1}, a) \quad (1)$$

式中: γ 为折扣因子; $r(i)$ 是回报函数^[5], 具体定义如下:

$$r(i) = \frac{\text{node}(i).w}{\text{node}(j).w} * \frac{1}{\text{node}(i).\text{cost}(j)} \quad (2)$$

式中: $\text{node}(i).w$ 是收到学习消息的任意节点 i 的权值; $\text{node}(j).w$ 是指发送学习消息的节点 j 的权值; $\text{node}(j).\text{cost}(i)$ 是指发送学习消息的节点 j 到收到消息的节点 i 的路径能量消耗。这样就可以把节点的能量、距离、邻居节点数目以及两节点间的链路通信耗能全部考虑进去, 更能反映出网络的节点动态, 回报函数越大, 说明该节点路由的“趋势”就越强。在该算法中假定 $\gamma=1$ 以加快学习速度。系统产生该动作的趋势主要决定于环境的奖赏值, 奖赏值如果为正则趋势会越来越强。换言之, 系统要使得(1)式最大化。

1.2 单基站算法性能分析

采用文献[5]中的算法进行仿真。为了对 Q 学习算法的性能进行评估, 本文首先对该算法在单基站的情况下和 ETBG 算法进行比较。定义从算法开

始运行到第一个节点死亡之间的时间为网络生存期, 网络生存期同样可以以数据采集总轮数表示。仿真的参数设置为的范围内 200 个节点, 在二十个不同的场景下进行仿真, 然后取其平均值, 可得不同通信半径下两种算法的生存期对比图, 如图 1 所示。

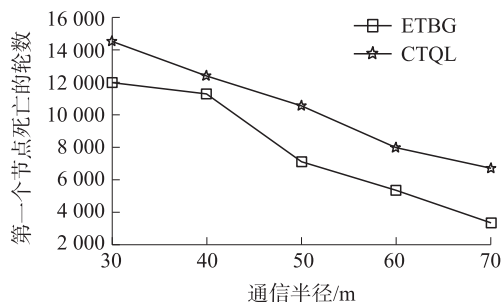


图1 不同通信半径下的生存期对比图

可以看到, 当 R 大于 40 时, 在 ETBG 算法中生命周期会出现明显的下降, 这是因为随着 R 增大, 梯度上限增大, 邻居节点数目增多, 使得簇的个数减少, 簇头之间距离增大, 导致簇头能量消耗更大, 数据传输的轮数就会明显减少。而引入 Q 学习算法之后, 对每个节点进行周期性的学习训练, 根据每条路径上的 Q 值选择最佳的路径, 就解决了 ETBG 算法生成簇树过程中未能找到最佳路径造成的能量在传输过程损耗过大的问题, 故其生命周期下降速度较 ETBG 算法更为平缓, R 越大 CTQL 算法的优势就体现的越明显。

CTQL 算法仍还存在一些问题, 如簇树过高会造成网络时延更长, 网络能量不均衡会导致部分节点过早死亡等, 在大规模的分层路由中并不适用。本文在 CTQL 算法的基础上, 提出基于 Q 学习的多基站分簇拓扑控制 (CTQL-MB) 算法。

2 基于Q学习的多基站分簇算法

多基站的策略将使簇树更矮, 网络的鲁棒性和扩展性更强, 能有效解决 CTQL 算法中时延过高、能量不均衡的问题, 使得生命周期得到有效延长。基站的能量容易补充, 多个基站不仅可以根据不同的场景分布在不同位置, 也可以在必要的情况进行移动, 以照顾能量损耗大的节点, 同时多个基站之间也可以进行直接的信息交互。

基站位置选择的原则是尽量以最小的梯度覆盖整个网络, 从而减少由于远距离传输时延过大, 以及部分节点转发过多引起耗能过大导致过早死亡的问题。

算法总体包括以下几个过程: 构建唯一梯度值、簇的确定、建立簇树。

2.1 构建唯一梯度值

假设网络中有 $n(n=1,2,3,\dots,n$ 为整数) 个基站, 分别为 BS1, BS2, \dots , BS n 。首先以基站 BS $n(n=1,2,3,\dots)$ 为中心, 以节点的通信半径 $mR(m=1,2,3,\dots,D/m, D/m$ 为整数) 为半径发送梯度划分消息。任意节点的初始唯一梯度值 $L=0$ 。节点在收到第一个梯度划分消息后将其置为自己的唯一梯度值, 当节点继续收到其他的梯度划分消息时, 将所收到的梯度值与当前的梯度值比较, 并按以下规则处理: ①如果当前梯度值与所收到的梯度值不同, 则将唯一梯度值更新为小的梯度值。②如果当前梯度值与所收到的梯度值相同, 则计算该节点到这两个梯度值所对应的基站的距离 d to BS $k(k \in n)$, 然后选择距离基站距离最小的所对应的梯度值作为该节点的唯一梯度值。

2.2 簇的确定

在确定各个节点的唯一梯度值后节点间进行第一次信息交互, 以 R 为功率半径向其邻居节点广播当前状态消息, 其中包括节点 ID、当前剩余能量、唯一梯度值 L 、状态 status。每个节点将得到的邻居信息保存在自己的邻居集中, 包括邻居节点的状态信息并计算本节点的邻居节点数目。节点根据自己邻居集中的信息, 运用 OWA 多属性决策方法^[5]确定自己的权值。

各个节点将自己邻居中唯一梯度值相同节点构成同梯度等级集合, 然后利用图论中独立集的概念将同一梯度等级内的节点进行比较, 如果其权值最大, 则宣布自己成为簇头节点, 具体方法见文献^[5]。

簇头确定后, 网络中的非簇头节点按照 ETBG 算法^[4]的策略加入不同的簇, 从而完成网络的分簇。

2.3 建立簇树

分簇阶段完成后, 基站 BS $n(n=1,2,3,\dots)$ 周期性地向其邻居节点发送学习消息 learnBS $n(n=1,2,3,\dots)$, 启动路径建立。学习消息中记录了节点的 Q 值、回报函数以及节点的能量信息和权值, 各个节点的初始 Q 值为 0。邻居节点继续向下一梯度的邻居簇头节点发送学习消息直到网络中所有簇头节点均进行学习训练。任意节点如果首次收到学习消息则建立 Q 表储存学习消息中的信息, 而收到学习消息的节点只有当到相应基站的距离大于发送消息的节点时才进行学习训练并按照规则 1~规则 4 处理, 否则放弃学习, 避免形成回路。

规则 1 如果簇头节点收到来自邻居的簇成员节点的学习消息, 根据式(2)来计算节点的回报函数, 再根据式(1)更新节点 Q 值, 并储存在自己的 Q 表中, 继续转发消息, 等待所有基站的学习任务进

行完后执行规则 4。

规则 2 如果簇成员节点收到簇头节点的学习消息, 根据式(2)计算回报函数, 再根据式(1)来更新节点的 Q 值, 并储存在自己的 Q 表中, 继续向下在两跳范围内转发该学习消息, 等待所有基站的学习任务都进行完则进入规则 4。

规则 3 如果簇成员节点收到非簇头节点的学习消息, 根据式(2)计算回报函数, 再根据式(1)更新 Q 值, 并储存并继续向下一跳范围内转发该学习消息, 等待所有基站的学习任务都进行完则进入规则 4。

规则 4 在各个基站到该节点的所有路径的 Q 值逐步迭代出来后, 选出该节点 Q 表中 BS $nQ(n=1,2,3,\dots)$ 的最大值所对应的节点作为自己的父亲节点, 如果该父亲节点为簇成员节点则其该簇成员节点声明自己为网关节点。

整个网络的簇头节点都遍历后, 算法结束。每个节点都建立到达基站的最优的传输路径。在该算法中, 多基站解决了“能量空洞”和时延问题, 生成簇树的过程中节点间通信的每一步选择都综合考虑了节点的通信能力、跳数、剩余能量, 从中选择到达各个基站最能均衡能量、节省能量、延长节点寿命的路径选择路由, 因而针对大规模的无线传感器网络具有一定的实用价值和现实意义。

2.4 特例分析

特例: 假设在一个的区域内随机抛洒 50 个传感器节点, 设定两个基站位置分别为 (100, 0), (100, 200)。采用文献^[4]所示能量模型, 各个参数的设定如下: 网络中所有节点的初始能量为 0.5 J, 通信半径为 50 m, 每接收一位消息耗能 50 nJ, 每发送一位消息传输 1 m 距离耗能 0.1 nJ。消息包固定长度为 128 bit, 基站初始 Q 值为 50。假定节点位置不变且相互间通信正常, 不考虑重传问题并且节点间不存在单向链路。如图 2 所示, 为运行该算法后的簇树图, 其中“方块”代表簇头, “菱形”代表网关, 其余为普通节点。

图 2 中可以看出基站 BS1, BS2 均发送一个学习消息, 其邻居节点进行转发。其中基站 BS2 的邻居中有簇头节点 37, 该节点在进行学习训练后继续想起邻居簇头转发学习消息, 其邻居节点中的簇头节点按照规则①进行更新, 并继续向下转发, 非簇头节点按照规则②、规则③、规则④进行更新; 基站 BS1 的邻居节点中无簇头节点在, 则节点 22 作为网关节点继续转发学习消息, 如果作为网关节点连接簇树, 则退出其所属的簇, 直接与基站通信。

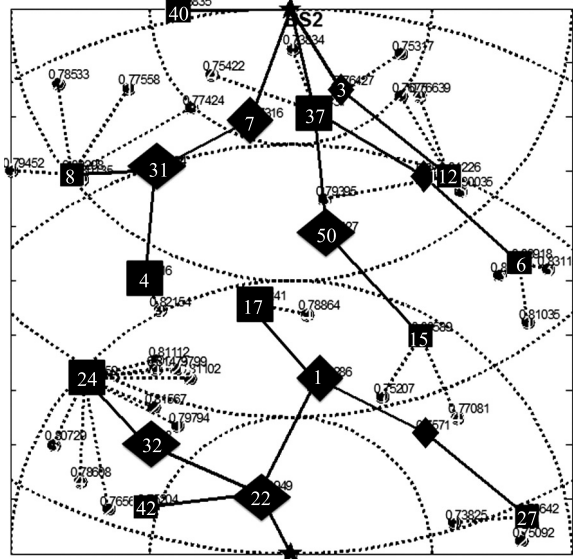


图2 算法运行后生成的簇树图

例如簇头节点17的可能路径及各个路径的 Q 值如表1所示:

表1 节点17的 Q 表

路径	Q 值
BS2-7-41-4-17	23.340
BS2-37-50-17	27.939
BS1-22-32-24-17	27.580
BS1-22-1-17	34.408

可以看到,其中BS1-22-1-17这条路径上的 Q 值最大,则选择该路径传输数据。

3 算法性能分析

对算法的性能进行评估,把该算法和单基站的Q学习算法进行比较,查看基站个数对网络生命周期的影响。在仿真的参数设置为 $200\text{ m}\times 200\text{ m}$ 的区域里随机抛洒200个节点,设定两基站时基站的位置分别是 $(100,0)(100,200)$,三基站时基站的位置分别是 $(0,0)(100,200)(200,0)$,检测整个网络的生命周期在二十次不同的场景下进行仿真,然后取其平均值,可得不同通信半径下不同基站个数下网络生命周期的对比图的如图3所示。

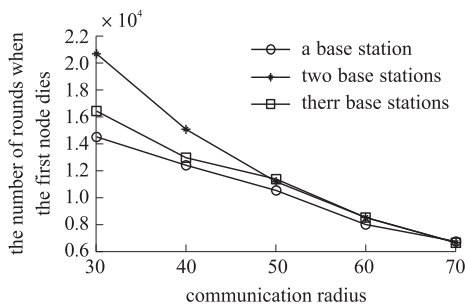


图3 不同基站个数在不同通信半径下的生存期对比图

从图3可以看到,当 $R=30$ 时单基站的生命周期最短,两基站的生命周期最长;当 $R=50$ 时双基站和三基站情况下生命周期无明显差别,但均比单基站的生命周期长。

理论上基站越多,网络能量消耗越小。而由于本论文中仿真参数选择的区域较小,三个基站与两个基站相比,每个节点的学习任务更重,计算所消耗的能量更多,反而使生命周期减小。

总体上,随着节点通信半径的增加,在网络的生命周期内,数据传输的轮数在减小。这是因为随着 R 增大,梯度的上限增大,邻居节点的数目就会增多节点间交互信息的能耗必然增加,且形成簇数也会减少,同时簇数减少还会造成簇成员数增加,簇头能量消耗更大。簇头之间距离增大导致信息交互所消耗的能量也越大,从而数据传输的轮数就会减少。

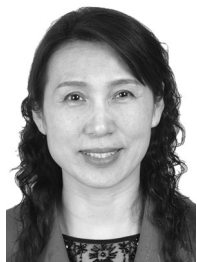
4 结论

针对单基站CTQL分簇拓扑控制算法中存在的“能量空洞”以及网络能量不均衡,网络时延过高,生成簇树的路径不优化的问题,提出了CTQL-MB算法。该算法引入了多个基站,在生成簇树的过程中通过在簇头间运行Q-learning算法,综合考虑节点的剩余能量、路径通信消耗等因素寻找簇头节点到达哪个基站才是最优路径,优化了数据传输时的路径选择,节省了整个网络的能量消耗。仿真结果表明,CTQL-MB算法相比较CTQL算法有效地延长了网络的生命周期。

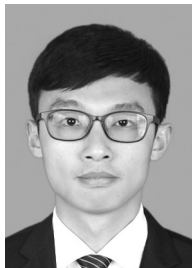
参考文献:

- [1] Tubaishat M, Madria S. Sensor Networks: An Overview[J]. IEEE Potentials, 2003, 22(2): 20-23.
- [2] Heinzelman W B, Chandrakasan A P, Balakrishnan H. An Application-Specific Protocol Architecture for Wireless Microsensor Network[J]. Wireless Communications, 2002, 1(4): 660-670.
- [3] Manjeshwar A, Agrawal D P. TEEN: A Routing Protocol for Enhanced Efficiency in Wireless Sensor Networks[C]//IEEE International Proceedings of 15th Parallel and Distributed Processing Symposium. IEEE Conference Proceedings, 2001: 2009-2015.
- [4] 阎新芳,朱玉芳,安娜. 无线传感器网络中一种分级簇的优化算法[J]. 传感技术学报, 2009, 22(3): 401-406.
- [5] 阎新芳,王晓晓,冯岩. 基于Q学习的无线传感网分簇拓扑控制算法[J]. 郑州大学学报, 2015, 36(2): 85-88.
- [6] Yan Xinfang, Zhang Yongkun, Tang Hailing. An ETBG Optimization Algorithm Based on Analytic Hierarchy Process in WSN[C]// Proceedings of ICCSEE'2013: 2013.3 The 2nd International Conference on Computer Science and Electronics Engineering. China, 2013: 1687-1690.
- [7] 何延杰,李腊元,邢明彦. WSN中一种能量均衡的分簇路由协

- 议的设计[J]. 传感技术学报, 2009, 22(10): 1510-1514.
- [8] Stephen S, Thiel M. A Q-Learning Strategy for LTE Mobility Load Balancing[C]//Andreas Personal Indoor and Mobile Radio Communications (PIMRC) IEEE 24th International Symposium 2013: 2154-2158.
- [9] Xie Ya, Huang Zhonghua. Study on Statistics Based Q-Learning Algorithm for Multi-Agent System[C]//Intelligent Systems Design and Engineering Applications, 2013 Fourth International Conference on DOI: 10.1109/ISDEA.2013.541 Publication Year: 2013: 595-600.
- [10] 章韵, 王静玉, 陈志. 基于 Q 学习的无线传感器网络自组织方法研究[J]. 传感器学报, 2012, 23(11): 1623-1626.
- [11] 苏彬庭, 方禾, 许力. 基于 Q-Learning 的无线传感器网络生命周期平衡路由[J]. 信息安全学报, 2015(4): 74-77.
- [12] 吴坚. 基于 OWA 算子理论的混合型多属性群决策研究[D]. 合肥: 合肥工业大学, 2008.



阎新芳(1958-), 女, 教授, 博士, 硕士生导师, 主要从事无线传感网等方面研究, iexfyan@zzu.edu.cn;



冯 岩(1990-), 男, 硕士生, 主要研究无线传感网络路由算法, 120628416@qq.com。