

# 基于 Q 学习的无线传感网分簇拓扑控制算法

阎新芳, 王晓晓, 冯 岩, 严晶晶

(郑州大学 信息工程学院, 河南 郑州 450001)

**摘 要:** 为了延长大规模无线传感器网络的生命周期, 在 ETBG 算法的基础上提出基于 Q 学习的分簇拓扑控制算法. 该算法利用有序加权平均(OWA)算子多属性决策的方法确定节点的权值, 利用 Q 学习算法对节点进行周期性的学习训练, 按照每条路径的 Q 值进行最优路径的选择, 然后就可以实现网络的拓扑控制. 仿真分析表明, 基于 Q 学习算法形成的簇树机制解决了 ETBG 算法在生成簇树过程中未能寻找到最佳路径而造成数据传输时能量损耗过多的问题, 从而达到延长网络生命周期的目的.

**关键词:** 无线传感器网络; OWA; Q 学习

**中图分类号:** TP393

**文献标志码:** A

**doi:**10.3969/j.issn.1671-6833.2015.02.019

## 0 引言

无线传感器网络(Wireless Sensor Networks, WSNs)针对不同的检测环境, 随机部署大量的传感器节点并在该区域内组成多跳网络 WSNs, 主要用来感知周围环境的各种信息<sup>[1]</sup>. 在环境监测和预报方面, 无线传感器网络可用于监视农作物灌溉情况、土壤空气情况、家畜和家禽的环境和迁移状况、无线土壤生态学、大面积的地表监测、行星探测、气象和地理研究、洪水监测等. 基于无线传感器网络, 可以通过数种传感器来监测降雨量、河水水位和土壤水分, 并依此预测山洪爆发, 描述生态多样性, 进行动物栖息地生态监测. 无线传感器网络的特点有: 节点硬件资源有限(能量、计算和存储能力)、能量效率要求高、节点数量众多且分布密集、自组织、多跳路由、动态拓扑. 由于无线传感网络节点的能量限制问题, 尤其是整个网络节点在一次性随机播撒后, 节点能量不可再生, 所以延长网络生存期、降低网络能量消耗成为传感网路由协议的首要设计任务<sup>[2-3]</sup>.

基于分簇的层次路由算法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<sup>[4-7]</sup>. 其中, 基于梯度的分簇拓扑控制算法<sup>[5]</sup>(Energy-Aware Topology Protocol Based on Gradient, ETBG)参考定向扩散协议中梯度的思想, 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结合图论的概念进行分簇, 然后在不同的梯度等级中结合极大权独立集的概念生成簇树. 但 ETBG 算法中生成簇树的过程仅考虑单个簇头节点的权值, 并没有综合考虑各方面的因素, 未能找到最优成树路径. 为解决这个问题, 笔者在 ETBG 算法的基础上, 用 Q 学习的方法生成簇树, 综合考虑整条路径上的最终 Q 值, 来选择最优路由路径.

## 1 相关工作

### 1.1 OWA 多属性决策

多属性决策是决策信息通过对一组现有方案进行排序并选择最优方式的决策方法. 主要包括决策信息的获取, 排序和选择最优信息的方法. 美国著名学者 Yager 教授<sup>[8]</sup>提出了有序加权平均算子(Ordered Weighted Average operator, OWA), 把各种信息按照从大到小重新排序, 并通过数据或证据的位置进行加权汇总, 可以消除不合理的情况. OWA 算子是一种较好的属性融合方法, 所以笔者采用基于最大熵的 OWA 多属性决策方法来对影响权值的各种因素进行集合处理, 这样可以让决策者的主观喜好带来的影响相对减小.

### 1.2 Q 学习

Q 学习属于强化学习中的一种常用算法. 而

收稿日期: 2014-10-30; 修订日期: 2014-12-27

基金项目: 郑州市科技攻关基金资助项目(20120555)

作者简介: 阎新芳(1958-), 女, 河南嵩县人, 郑州大学教授, 博士, 主要从事无线传感网等方面研究, E-mail: iexfyang@zzu.edu.cn.

强化学习<sup>[9]</sup>是指对环境状况进行映射的学习,为了从环境中获得系统行为累积奖励值的最大值,强化学习系统的输出行为动作  $a$  是根据内部的推理机制,在环境状态输入  $s$  作用下所得到的. 然后环境改变到一个新的状态  $s'$ , 同时得到从环境反馈的瞬时回报值  $r$ , 强化学习系统的目标是使环境中的智能体学习一个行为策略  $\pi: S \rightarrow A$ . Q 学习是强化学习中的一种常用算法,使系统中的智能体所选择的下一步的动作能够获得 Q 值累计值最大.

其中  $Q$  值得更新公式如式(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:  $\gamma$  为折扣因子,在本算法中令  $\gamma = 1$  以加快学习速度;  $r(i)$  是回报函数.

## 2 基于 Q 学习的分簇算法

基于 Q 学习的分簇拓扑控制(CTQL)算法首先以基站为中心,以节点的通信半径为半径将整个网络划分为同等大小的梯度,然后利用图论中独立集的概念对同一梯度内的簇头节点进行分簇,最后运行 Q 学习算法使整个网络成簇树.

### 2.1 综合权值的确定

在该算法开始时,基站依次以  $nR(n=1,2,\dots, D/R, D/R$  为整数)为通信半径发送梯度信息,各个节点确定自己的梯度值. 然后各个节点以  $R$  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、梯度  $L$ 、当前剩余能量  $node(i). Er$ 、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,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 多属性决策方法确定自己的权值,其中综合权值定义如下,

$$W = w_1 H_1 + w_2 H_2 + w_3 H_3, \quad (2)$$

式中:  $H_1$  为节点的剩余能量因素;  $H_2$  为节点到基站的距离因素;  $H_3$  为节点的邻居节点的数目因素;  $\{w_1, w_2, w_3\}$  是各个属性的权重集合.

各个属性的归一化表达式如下所示,

$$H_1 = \frac{node(i). E_r}{node(i). E_{\max}}, \quad (3)$$

$$H_2 = \frac{node(i). dist(BS) - dtoBSmin}{dtoBSmax - dtoBSmin}, \quad (4)$$

$$H_3 = \frac{node(i). number - nodenumbermin}{nodenumbermax - nodenumbermin}, \quad (5)$$

式中:  $node(i). E_{\max}$  是任意节点  $i$  初始的最大能量值;  $node(i). E_r$  是任意节点  $i$  当前的剩余能量;

$node(i). dist(BS)$  是任意节点  $i$  到基站的距离;  $dtoBSmax$  是整个网络中的节点与基站最远的距离;  $dtoBSmin$  是整个网络中的节点与基站最近的距离;  $node(i). number$  是任意节点  $i$  的邻居节点的数目;  $nodenumbermax$  是所有节点中邻居节点数的最大值,  $nodenumbermin$  是所有节点中邻居节点数目最小值.

针对各个属性的权重集合  $\{w_1, w_2, w_3\}$ , 采用基于最大熵的 OWA 决策方法<sup>[10]</sup> 决定权重系数, 其中在该算法中把决策者的乐观系数  $\alpha$  定义为 0.5, 通过这样的决策方法可以把各种因素归一化到一个权值上, 而且集结了决策者的主观偏好以减少决策中不合理因素, 定性和定量地分析表达了各个因素对簇头选举所占的比重.

### 2.2 节点成簇

在各个节点确定自己的综合权值后, 与其同梯度等级的邻居节点相比较, 如果其权值最大, 则宣布自己成为簇头节点, 网络中的其它节点按照 ETBG 中的方法加入不同的簇, 从而完成网络的分簇.

### 2.3 建立簇树

#### 2.3.1 改进回报函数

在生成簇树之前, 首先确定 Q 学习的回报函数  $r(i)$ . 传统的应用于无线传感器网络的 Q 学习算法的回报函数只考虑单一因素影响, 如两节点间的跳数最少或距离最短, 这样就会忽略了节点的剩余能量以及节点间通信路径上能量消耗的影响. 笔者为了更加节省和均衡网络的能量消耗, 定义回报函数

$$r(i) = \frac{node(i). w}{node(j). w} \cdot \frac{1}{node(i). cost(j)}, \quad (6)$$

式中:  $node(i). w$  是收到学习消息的任意节点  $i$  的权值;  $node(j). w$  是发送学习消息的节点  $j$  的权值;  $node(j). cost(i)$  是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. 综合考虑节点的能量、距离、邻居节点数目以及两节点间的链路通信耗能.

#### 2.3.2 算法描述

当网络中的节点均确定分簇状态后, 进入生成簇树的算法. 算法开始时, 基站周期性地向其邻居节点发送学习消息, 学习消息中包含节点  $Q$  值、 $r$ 、 $E_r$ 、 $w$ , 各个节点的初始  $Q$  值为 0, 以启动路径建立. 收到学习消息的节点继续向其邻居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. 收到学习消息的节点只有距离大于发送

消息的节点时才进行学习训练按照规则 1、2 处理,并建立  $Q$  表储存学习消息中的信息,否则直接抛弃该消息.

规则 1 收到学习消息节点根据公式(6)来计算节点的回报函数,根据公式(1)来更新节点的  $Q$  值,并储存在自己的  $Q$  表中,继续转发学习消息.簇头节点在两跳范围内转发.非簇头节点在一跳范围内转发,在转发学习消息的同时转发节点的  $Q$  表,这样可以让其邻居节点在此基础上计算  $Q$  值从而减少计算量.

规则 2 等待到达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  $Q$  最大值所对应的节点作为自己的父亲节点,并向该父亲节点发送一个 *father* 消息.收到 *father* 消息的节点如果是簇头节点则将发送给它学习消息的节点置为网关节点,如果收到消息的节点是簇成员节点,则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.这样就可以得到各个节点到基站进行数据传输的最优路径.在该算法中,每一次节点间的通信路径选择都考虑节点的跳数、节点间的通信能力、剩余能量等因素,因此选择的路径可以平衡网络能量消耗,延长网络的生命周期,具有一定的实用价值和现实意义.

2.4 特例分析

特例:在一个  $200 \times 200 \text{ m}^2$  的的树林区域内随机抛洒 50 个传感器节点,用来监测树林的湿度以防范火灾.采用文献[4]所示的能量模型,各个参数的设定如下:网络中所有节点初始能量  $0.5 \text{ J}$ ,每接受一位消息消耗能量  $50 \text{ nJ/bit}$ ,每发送一位消息传输一米距离的能量消耗为  $0.1 \text{ nJ/bit} \cdot \text{m}^2$ ,消息包固定长度  $128 \text{ bits}$ ,通信半径  $R = 50 \text{ m}$ ,  $n = 4$ ,基站的  $Q = 50$ .假定节点位置不变,节点间通信正常.不考虑消息发送过程中的冲突的重传且节点间不存在单向链路.

图 1 所示为运行该算法后的簇树图,可以看出基站 BS1 发送一个学习消息,其邻居节点进行转发,其中簇头节点 35 继续向其邻居节点发送学习消息,其邻居节点中的节点按照规则(1)、(2)进行更新;基站 BS1 的邻居节点中的非簇头节点也转发学习消息,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.其中,例如节点的 24 的通信路径  $24 \rightarrow 35 \rightarrow \text{BS1}$  的叠加的  $Q$  值为  $58.12$ .

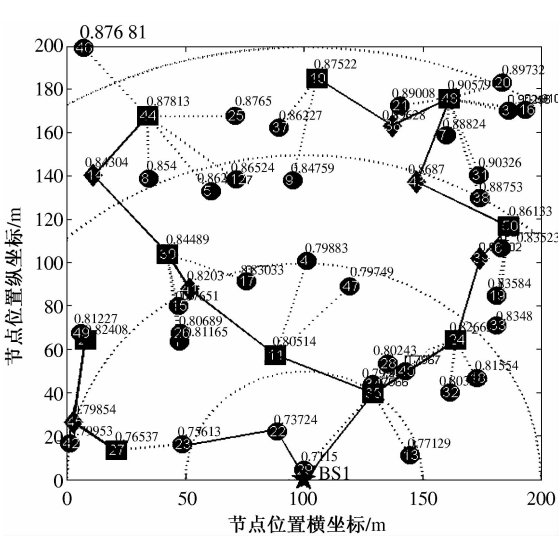


图 1 算法运行后生成的簇树图  
Fig.1 the cluster tree after running the algorithm

3 算法性能分析

为了对算法的性能进行评估,把该算法和 ETBG 算法进行比较.定义网络生命周期为从算法开始运行到第一个节点死亡之间的时间,以数据采集总轮数来表示.在仿真参数为  $200 \times 200 \text{ m}^2$  的树林里随机抛洒 200 个传感器节点,监测整个网络的生命周期.在 20 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两种算法的生存期对比图,如图 2 所示.

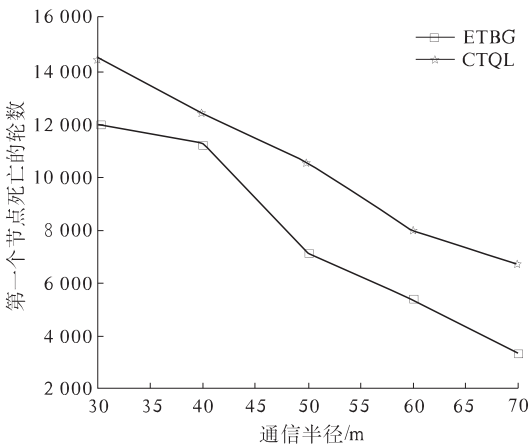


图 2 不同通信半径下的生存期对比图  
Fig.2 Comparison of lifetime under different communication radius

可以看到,CTQL 算法的生存期较 ETBG 算法更平稳且网络的生命周期更长,传感器节点的可监测时间也更长,当  $R = 40$  时,两者差距最小.总体上,随着节点通信半径的增加,在网络的生存周期内,数据传输的轮数在减小.这是因为随着通信

半径  $R$  增大,形成簇数减少而成员数增加,同时簇头之间的距离增大,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

通过研究 EBTG 算法中存在的网络能量不均衡,生成簇树的路径没有优化的问题,提出了 CTQL 算法.该算法和 ETBG 算法相比,主要利用了基于极大熵的 OWA 多属性决策的方法来决定综合权值,更加全面的反映了节点在竞选簇头时的能力.利用 Q 学习算法生成簇树,综合考虑节点的剩余能量,路径通信耗能因素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.通过在树林环境中仿真结果表明,该算法与 ETBG 算法相比较有效地延长了网络的生命周期.

#### 参考文献:

- [1] TUBAISHAT M, MADRIA. SENSOR S. 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] AN Na, YAN Xin-fang, ZHU Yu-ang. A virtual backbone network algorithm based on the multilevel cluster tree with gateway for wireless sensor networks [C]//Proceedings of CCWMSN 2007 Shanghai, China: 2007. 12 The IET International Communication Conference on Wireless Mobile and Sensor Networks. Shanghai, 2007: 462 - 465.
- [5] 阎新芳, 段磊, 李腾. 无线传感网中基于梯度的拓扑控制算法 [J]. 计算机工程与应用, 2011, 47 (2): 95 - 98.
- [6] 阎新芳, 王志龙, 闫新生. WSN 中基于梯度场拓扑控制算法的维护更新 [J]. 传感器与微系统, 2011, 30 (8): 56 - 58.
- [7] YAN Xin-fang, ZHANG Yong-kun, TANG Hai-liang. An ETBG optimization algorithm based on analytic hierarchy process in WSS [C]//Proceedings of ICCSEE' 2013: 2013. 3 The 2nd International Conference on Computer Science and Electronics Engineering. China, 2013: 1687 - 1690.
- [8] YAGER R R. Weighted maximum entropy OWA aggregation with applications to decision making under risk [J]. IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans, 2009, 39 (3): 183 - 189.
- [9] MEHTA N, NATARAJAN S, TADEPALLI P, et al. Transfer in variable-reward hierarchical reinforcement learning [J]. Machine Learning. 2008 (3): 156 - 172.
- [10] YAGER R R. "Centered OWA operators" Soft Computing [J]. soft Computing, 2007, 11 (7): 631 - 639.

### A Clustering Topology Algorithm Based on Q-learning in WSN

YAN Xin-fang, WANG Xiao-xiao, FENG Yan, YAN Jing-jing

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

**Abstract:** To prolong the lifetime of wireless sensor network, a Clustering Topology Algorithm Based on Q-learning in WSN (CTQL) is proposed on the basis of classical clustering algorithms such as ETBG. The Ordered Weighted Average (OWA) operator multi-attribute decision making method is used to determine the weight of the nodes, and Q learning algorithm is used to periodically train the cluster heads. So the Q value of the optimal path is selected of this algorithm and the topology control is realized. Through simulation study shows that the use of Q-learning algorithm to resolve the problem that much energy consumption of ETBG algorithm fails to find the best path and CTQL effectively extend the network lifetime.

**Key words:** wireless sensor network; Ordered Weighted Average (OWA) operator; Q-learning