

# 基于增强学习的非均匀分簇水声传感器网络能耗研究

侯睿,何柳婷

(中南民族大学 计算机科学学院 武汉 430074)

**摘要** 近年来,水声传感器网络越来越成为研究的热点,但由于水下环境复杂多变,导致网络中能量消耗不均的问题.针对此问题提出了一种基于增强学习的非均匀分簇的水声传感器网络路径优化算法.该算法首先根据水声传感器网络中节点的深度和剩余能量把传感器节点分成大小不同的簇;然后根据节点的综合属性值选出最佳簇头;最后在数据传输阶段利用增强学习和  $\epsilon$ -greedy 策略对簇间的传输路径进行决策和学习,寻找最优路由.实验结果表明:本文方法可以有效均衡能耗,并延长网络寿命.

**关键词** 水声传感器网络;聚类算法;增强学习;路径优化

中图分类号 TP393 文献标志码 A 文章编号 1672-4321(2020)02-0205-05

doi: 10.12130/znmzdk.20200216

引用格式 侯睿,何柳婷.基于增强学习的非均匀分簇水声传感器网络能耗研究[J].中南民族大学学报(自然科学版) 2020,39(2):205-209.

HOU Rui, HE Liuting. Study on energy-consumption of unequal clustering in Underwater Acoustic Sensor Networks based on reinforcement learning [J]. Journal of South-Central University for Nationalities( Natural Science Edition) 2020, 39( 2) : 205-209.

## Study on energy-consumption of unequal clustering in Underwater Acoustic Sensor Networks based on reinforcement learning

HOU Rui, HE Liuting

(College of Computer Science, South-Central University for Nationalities, Wuhan 430074, China)

**Abstract** In recent years, Underwater Acoustic Sensor Network has become a hot research topic, but due to the complex and changeable underwater environment, the energy consumption in the network is uneven. In order to solve this problem, a path optimization algorithm of Underwater Acoustic Sensor Network based on reinforcement learning is proposed. The algorithm firstly divides the sensor nodes into clusters of different sizes according to the depth and residual energy of the nodes in the network. Secondly, the best cluster head is selected according to the comprehensive attribute value of nodes. At last, the transmission paths between the cluster-head and the cluster-head are determined and learned by using reinforcement learning and  $\epsilon$ -greedy strategy in the data transmission stage, and the global optimal path is selected. Experimental results show that the proposed method can effectively balance energy consumption and prolong network life.

**Keywords** Underwater Acoustic Sensor Network; clustering algorithm; reinforcement learning; path optimization

由于全世界海洋面积约占地球表面积的 71%, 并且在海洋当中蕴含着丰富的资源, 近些年, 水声通信越来越成为了一个新的研究热点. 水声传感器网

络( UASN, Underwater Acoustic Sensor Network) [1, 2] 是水下声学通信技术结合无线传感器网络( WSN, Wireless Sensor Network) [3] 所产生的一个新的研究方向. 目前该技术广泛应用于海洋的定期监测以及

收稿日期 2019-10-11

作者简介 侯睿(1977-) 男, 教授, 博士, 研究方向: 互联网体系结构、无线传感器网络, E-mail: hourui@mail.scuec.edu.cn

基金项目 国家自然科学基金资助项目(61972424; 60841001); 中央高校基本科研业务费专项资金资助项目(CZT19011); 中南民族大学研究生学术创新基金资助项目(2019sycxj116)

海洋数据采集等方面.但是,由于水下环境复杂多变,使得传感器节点更换新的电池或者再次充电非常困难,因此如何降低 UASN 能量消耗,已经成为该领域的热点问题.目前针对 UASN 能耗优化方面,已经有不少关于聚类算法以及路由算法方面的研究.

## 1 相关研究工作

近些年,聚类算法已经广泛应用于无线传感器网络中.文献[4]提出了一种低能耗自适应集群分层型协议(Low Energy Adaptive Clustering Hierarchy, LEACH),它是针对无线传感器网络的第一个分簇路由算法,其采用随机分簇策略和周期性簇头轮换机制,将整个传感器网络分成若干个大小均等的簇,使得网络中的所有节点能量消耗均衡,从而延长网络的生命周期.但是,由于簇头的选择是随机的,这样低能量的节点被选为簇头,很容易造成节点的过早死亡.文献[5]提出了一种能量高效的非均匀分簇算法(Energy-Efficient Unequal Clustering, EEUC).在 EEUC 算法中,每个节点先根据一个预先确定的门限,随机确定自身能否成为备选簇头.备选簇头节点根据接收 sink 节点信号强度,选出不同范围的簇,并在簇内选择能量最大的成为簇头节点.该算法虽然在一定程度上延长了网络生命周期,但实际操作中比较困难,不利于实施.文献[6]提出了一种能量平衡的不等层分簇(Energy-balanced Unequal Layering Clustering, EULC)路由算法,它根据节点的深度网络被划分为不同宽度的层,较浅的层的宽度比较深的层要小,并且通过根据节点与 sink 节点的距离来调整传输功率,从而缓解了“热点”问题.“热点”问题是指在网络中,一些节点可能在数据传输中过多地消耗能量而导致过早死亡,从而使整个网络瘫痪,无法正常工作的问题.

针对路由算法方面的研究,传统的路由算法虽然能很好地提高网络性能,但由于其处理多个约束的能力和较高的计算复杂度仍然受到限制.近年来,已有许多基于智能算法的路由方案被提出用于地面无线传感器网络,但这些算法很少被应用于 UASN 中.近几年,增强学习领域的研究有了很大的发展.目前,已经有一些针对单个节点的路径优化方案被提出,文献[7]发现基于增强学习的路由优化不仅要考虑到到达目的地的跳数,还要考虑流量阻塞的贡献,考虑每个节点的队列长度对延迟的影响,从而确定最佳路由路径.文献[8]提出了一种基于机器学习

的高效寿命扩展水下传感器网络自适应路由算法(Q-learning based Energy-efficient and Lifetime-extended Adaptive Routing, QELAR),该算法将 Q 学习技术应用于 UASN 的分布式路由算法中,以平衡传感器节点之间的工作负载,降低网络开销,提高能源效率,延长网络寿命.但是,目前基于增强学习的路由优化算法主要运用于单个节点的路径优化上,用在集群之间的路径优化研究非常少.

针对现有的水声传感器网络聚类算法在簇的形成以及路由优化方面的不足,本文设计了一种基于增强学习的能量消耗非均匀分簇算法(Energy-consumption of Unequal Clustering based on Reinforcement Learning, EUCRL).该算法首先根据水声传感器网络的深度和剩余能量把传感器节点分成大小不同的簇.同时在数据传输阶段利用增强学习和  $\epsilon$ -greedy 策略对簇间的传输路径进行决策和学习,选出全局最优路径.实验结果表明本文方法可以有效减少数据传输所带来的能量消耗.

## 2 通信模型

为了能够准确模拟水下环境,首先需要对水声通信进行建模.本文采用文献中经典的水声通信模型<sup>[9]</sup>模型中的参数定义如下:

发送节点的最低发送功率表示为:

$$P = P_t A,$$

其中  $P_t$  为一个数据包被节点接收的正常功率,  $A$  为传输功率随传输距离的衰减量,与传输距离、工作频率以及数据的发送方式有关,  $A$  可表示为:

$$a = \alpha^d d^k,$$

其中  $k$  为能量扩散因子,与信号传播条件有关(在实际应用中常取  $k = 1.5$ ),  $d$  是传输距离,衰减系数  $\alpha = 10^{a(f)/10}$ ,它由吸收损耗  $a(f)$  决定,与频率有关,  $a(f)$  表示为:

$$a(f) = \frac{0.11f^2}{1+f^2} + \frac{44f^2}{4100+f^2} + \frac{2.75f^2}{10000} + 0.003,$$

其中  $f$  为节点的工作频率.

当传输距离为  $d$  时,节点发送长度为  $l$  bit 的数据包的能耗为:

$$E_T(l, d) = lP_t \alpha^d d^k.$$

节点接收长度为  $l$  bit 的数据包的能耗为:

$$E_R(l, d) = lP_r,$$

其中  $P_r$  为常数.

### 3 EUCRL 路由算法

#### 3.1 计算节点的竞争半径

在聚类算法中, 簇头节点距离基站越近, 需要转发其他簇头节点的数据就越多, 能耗越大. 为了平衡簇头节点的能耗, 本文使靠近基站的簇范围更小, 也就是在靠近基站的区域选举更多的簇头. 因此引入竞争半径  $R_i$ , 它综合考虑了节点的剩余能量、与基站之间的距离等因素, 以控制簇头节点的分布. 其计算公式如下:

$$R_i = (1 - \lambda(\mu \frac{E_{avg} - E_{res(i)}}{E_{avg}} + (1 - \mu) \cdot \frac{d_{max} - d_{i-sink}}{d_{max} - d_{min}})) R_0,$$

式中  $d_{max}$ ,  $d_{min}$  分别为网络中节点距离基站的极大、最小距离,  $R_0$  为预先设定的极大竞争半径,  $E_{avg}$  为网络中存活节点的平均剩余能量,  $\lambda, \mu \in (0, 1)$  是自适应系数.

#### 3.2 簇头的选举

本文在选举簇头时, 综合考虑节点的剩余能量、节点到基站的距离以及节点度因素, 选出综合属性值最高的节点为簇头, 具体计算公式如下:

$$W_i = \begin{cases} \alpha \frac{E_{res(i)}}{E_{int(i)}} + \beta(1 - \frac{d_{i-sink}}{d_{max}}) + \gamma N_i, & E_{res(i)} \geq E_{avg} \\ 0 & E_{res(i)} < E_{avg} \end{cases},$$

其中  $E_{res}$  为节点的剩余能量,  $E_{int}$  为节点的初始能量,  $N_i$  为节点度.

#### 3.3 簇间数据传输

在本算法中, 簇内节点采用单跳的方式进行数据传输, 簇间采用单跳和多跳结合的方式进行数据传输. 在簇间数据传输中, 使用增强学习中的  $Q$  学习方法对路由进行选择, 每个簇头从候选中继簇头中选择  $Q$  值最大的作为下一跳, 以此不断地将数据转发至基站. 增强学习具体工作原理如图 1 所示, 每个传感器节点相当于一个智能体 Agent, 当节点做出路由动作 action 后, 环境将返回 Agent 一个奖励值, 发送方节点将会更新自己的状态.

簇头  $i$  的候选中继节点集合定义如下:

$$s_{i-j} = \{s_j \mid d(s_j, sink) < d(s_i, sink)\},$$

其中  $d(s_i, sink)$ ,  $d(s_j, sink)$  分别表示簇头  $s_i, s_j$  到基站的距离, 若  $d(s_j, sink) < d(s_i, sink)$ , 则  $s_j$  为候选中继簇头节点; 若  $s_{i-j}$  为空集, 则  $s_i$  直接通过单跳的方式将数据传输给基站.

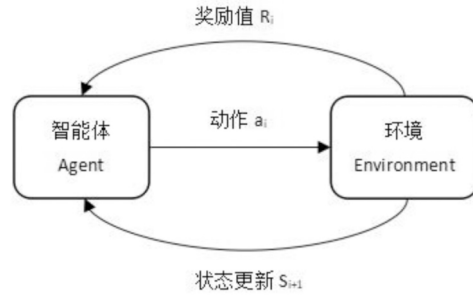


图 1 增强学习工作原理

Fig.1 Working principle of reinforcement learning

在路由选择中, 每个节点在做出动作后将被赋予一个  $Q$  值, 其计算公式如下:

$$Q(s_i, a_i) = R(s_i, a_i) + \gamma \max_{a_{i+1}} Q(s_{i+1}, a_{i+1}).$$

同时, 因为水下环境复杂多变, 拓扑结构可能随时间而不断变化, 本文通过根据如下  $Q$  值更新公式调整数据传输路线:

$$Q(s_i, a_i) \leftarrow Q(s_i, a_i) + \alpha [R(s_i, a_i) + \gamma \max_{a_{i+1}} Q(s_{i+1}, a_{i+1}) - Q(s_i, a_i)],$$

其中  $\alpha$  为学习率, 是一个权衡上一次学习结果和这一次学习结果的量,  $\gamma \in (0, 1)$  为折损因子.

在该算法中, 定义了节点的奖励函数, 其中节点间距离和邻居节点剩余能量都被考虑用于适当的路由决策中. 奖励函数  $R(s_i, a_i)$  定义如下:

$$R(s_i, a_i) = \alpha \frac{d_{own(j)}}{d_{max}} + \beta \frac{E_{res(j)}}{E_{int}} + c, \quad (1)$$

其中  $d_{own(j)}$  为邻居节点  $j$  与 sink 节点间的距离,  $E_{res(j)}$  为邻居节点  $j$  的剩余能量.

在簇间数据传输中, 簇头节点首先利用增强  $Q$  学习算法从备选中继簇头中选择  $Q$  值最大的节点为下一跳节点. 除此之外, 本文在增强学习的基础上采用一种基于  $\epsilon$ -greedy 的策略. 该策略的使用, 使节点保持  $(1 - \epsilon)$  的概率直接选择具有最大  $Q$  值的节点作为中继节点, 同时有  $\epsilon$  的概率随机选择下一跳节点. 这样做的目的是: 水下环境复杂多变, 所以选择的具有最大  $Q$  值的路径在数据传输中不一定是全局最优路径. 通过利用  $\epsilon$ -greedy 策略可以很好地平衡随机和贪婪的比率, 让网络跳出局部最优, 实现真正意义上的全局最优, 并加快网络收敛速度.

基于增强学习的  $\epsilon$ -greedy 策略路由决策  $\pi(s)$  选择定义如下:

$$\pi(s) = \begin{cases} \arg \max_a Q & \text{概率 } (1 - \epsilon) \\ \text{随机选择一个动作} & \text{概率 } \epsilon \end{cases}$$

## 4 实验与分析

### 4.1 实验配置

为了验证 EUCRL 算法的有效性,本文利用 Matlab<sup>[10]</sup> 构建水声传感器网络仿真环境,设定节点在网络中随机分布,sink 节点为所有数据包传输的目的地.实验具体参数配置见表 1.

表 1 仿真参数的配置

Tab.1 Configuration of simulation parameters

参数	取值
节点个数	500
网络覆盖范围/m	200×200×200
最大传输半径/m	50
sink 节点位置	(100,100,ρ)
数据包大小/bits	200
初始功率 $P_0$ /mW	3
频率 $f$ /kHz	10
节点初始能量 $E_0$ /J	5

### 4.2 实验结果与分析

本文实验分为两部分.第一部分实验,是对(1)式奖励函数中的加权因子  $\alpha$   $\beta$  对网络性能的影响进行实验,并分析加权因子  $\alpha$   $\beta$  的设置对网络中包的到达率以及能量消耗的影响.第二部分实验是本文算法与 LEACH, QELAR 的对比实验,通过实验结果验证本文算法的有效性.

#### 4.2.1 $\alpha$ $\beta$ 对网络性能的影响

图 2 为加权因子  $\alpha$  对网络中数据包到达率以及能量消耗的影响.如图所示,随着  $\alpha$  的增加,数据包的到达率逐渐增高.但是,由于选择出的下一跳时没有充分考虑节点的剩余能量,因此网络中的能量消耗越来越高.

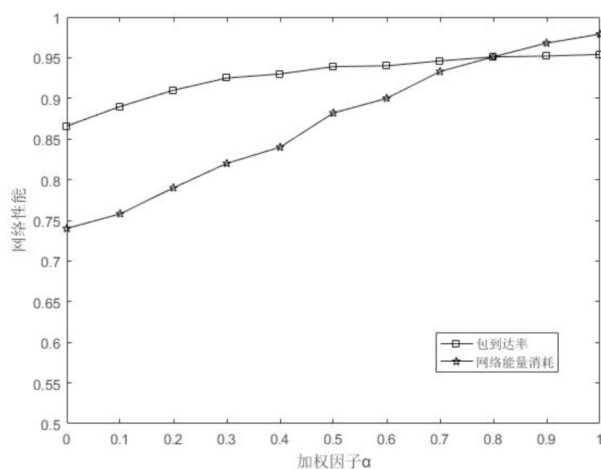


图 2  $\alpha$  对网络性能的影响

Fig.2 Effect of  $\alpha$  on network performance

如图 3 所示,加权因子  $\beta$  反应了节点剩余能量对于路由选择的重要程度.随着  $\beta$  的增加,节点逐渐倾向于选择剩余能量较大的节点作为下一跳,而忽略了节点间距离这一约束条件,这样将增加所选路径不是最短路径的可能.因此,根据实验结果可以看出,随着  $\beta$  的增加,节点的能量消耗逐渐变大,而数据包的到达率却逐渐减小.

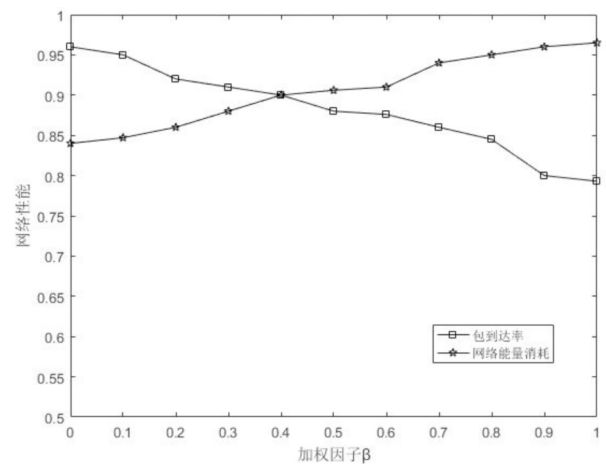


图 3  $\beta$  对网络性能的影响

Fig.3 Effect of  $\beta$  on network performance

#### 4.2.2 算法比较

如图 4 所示,通过对比 LEACH, QELAR 和 EUCRL 三种算法,可以看出本文所提出的能耗优化 EUCRL 算法可以有效平衡水下传感器网络的能量消耗,并延长了网络生命周期.

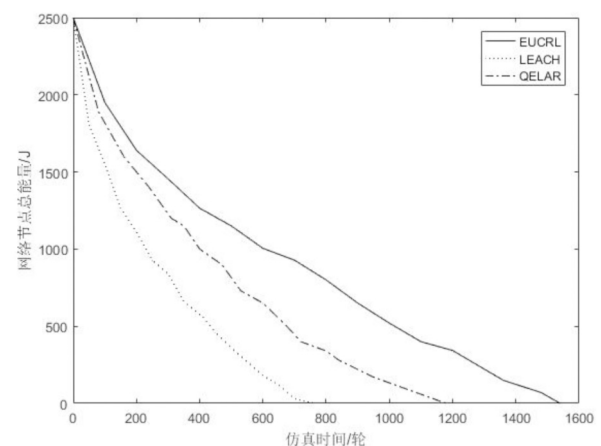


图 4 网络能耗情况对比

Fig.4 Comparison of network energy consumption

图 5 为 EUCRL, LEACH, QELAR 三种算法基站接收节点数据包个数的对比情况.如图所示,在节点未出现死亡之前,三种算法节点发送到基站的数据包数相同,但由于本文 EUCRL 算法使用非均匀分簇使得均衡能耗效果优于其他两种算法,并且节点

死亡轮数晚于其他两种算法, 所以由实验结果可以看出本文所提出的 EUCRL 算法可以有效提高网络中数据包的接收率.

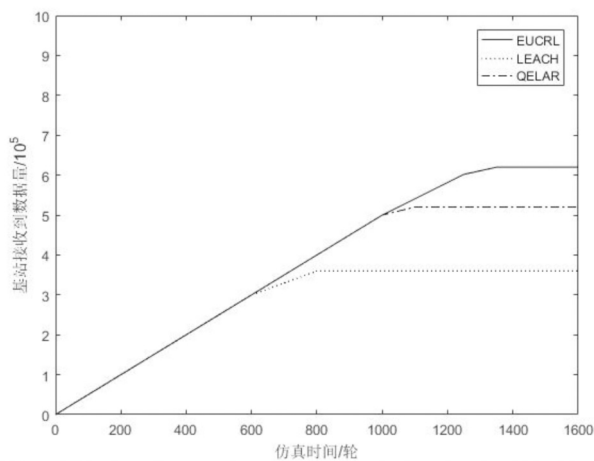


图5 sink 节点接收数据量情况对比

Fig.5 Comparison of the amount of data received by sink node

## 5 结语

本文针对水下通信提出了一种基于增强学习的非均匀分簇的水声传感器网络能耗优化算法.该方法通过对水下传感器进行非均匀分簇,使簇头分布更加合理,有效平衡了水下传感器网络的能量消耗;同时,通过利用增强学习和  $\epsilon$ -greedy 策略对簇间路径的学习及预测,显著降低了数据传输路径的复杂度,减少了数据传输时的能量消耗,延长了网络寿命.

### 参考文献

- [1] ZHANG S, LI D, CHEN J. A link-state based adaptive feed-back routing for Underwater acoustic sensor networks [J]. IEEE Sensors Journal, 2013, 13( 11): 4402-4412.
- [2] YU H, YAO N, LIU J. A adaptive routing protocol in underwater sparse acoustic sensor networks [J]. Ad Hoc Networks, 2015, 34: 121-143.
- [3] JANNU S, JANA P. Energy efficient unequal clustering and routing algorithms for wireless sensor networks [C]// IEEE. International Conference on Advances in Computing, Communications and Informatics( ICACCI). Calcutta: IEEE, 2014: 2091-2097.
- [4] HEINZELMAN W, CHANDRAKASAN A, BALAKRISHNAN H. Energy-efficient communication protocol for wireless microsensor networks [C]//IEEE. The 33rd Annual Hawaii International Conference on System Sciences. Hawaii: IEEE, 2000: 1-10.
- [5] LI C, YE M, CHEN G, et al. An energy-efficient unequal clustering mechanism for wireless sensor networks [C]//IEEE. The 2nd IEEE International Conference on Mobile Ad-hoc and Sensor Systems( MASS). Washington, D C: IEEE, 2005.
- [6] HOU R, HE L, HU S, et al. Energy-balanced unequal layering clustering in underwater acoustic sensor networks [J]. IEEE Access, 2018, 6: 39685-39691.
- [7] BOYAN J, LITTMAN M. Packet routing in dynamically changing networks: a reinforcement learning approach [C]//NIPS. Advances in Neural Information Processing Systems. San Francisco: NIPS, 1994: 671-678.
- [8] HU T, FEI Y. QELAR: A machine-learning-based adaptive routing protocol for energy-efficient and lifetime-extended underwater sensor networks [J]. IEEE Trans Mobile Comput, 2010, 9( 6): 796-809.
- [9] SOZER E, STOJANOVIC M, PROAKIS J. Underwater acoustic network [J]. IEEE J Ocean Eng, 2000, 25( 1): 72-83.
- [10] HIGHAM D, HIGHAM N. MATLAB guide [M]. Philadelphia: Society for Industrial and Applied Mathematics, 2000.

(责任编辑 曹东)