

文章编号: 2095-4980(2024)09-0952-07

基于 Q -learning 的变电站无线传感器网络路由算法

赵 锴, 沙 杰, 丛 尤嘉

(国网上海市电力公司 嘉定供电公司, 上海 201800)

摘 要: 电力系统中的无线传感器网络(WSN)可以对工作中设备的状态和环境数据进行实时感知采集, 是一种推动智能电网发展的重要技术。针对变电站场景中 WSN 的网络存活时间、传输时延、传输丢包率上的特殊要求, 提出了一种基于强化学习的 WSN 路由方案。将数据包在 WSN 的发送过程抽象为一个马尔科夫决策过程(MDP), 根据优化目标合理设置奖励, 并给出了基于 Q -learning 的最优路由求解方法。仿真结果与数值分析表明, 所提方案在网络存活时间、传输时延、丢包率等方面的性能均优于基准方案。

关键词: 变电站无线传感网; 路由策略; 马尔科夫决策过程; Q -learning 算法; 网络性能优化

中图分类号: TN92

文献标志码: A

doi: 10.11805/TKYDA2024034

Q -learning based routing algorithm for substation wireless sensor networks

ZHAO Kai, SHA Jie, CONG Youjia

(Jiading Power Supply Company, State Grid Shanghai Electric Power Company, Shanghai 201800, China)

Abstract: Wireless Sensor Network(WSN) in the power system can sense and collect the status of the working equipment and environmental data in real time, which is an important technology to promote the development of smart grid. Aiming at the special requirements of network survival time, transmission delay, and transmission packet loss rate of WSN in substation scenarios, a WSN routing scheme based on reinforcement learning is proposed. The sending process of packets in WSN is abstracted as a Markov Decision Process(MDP), the rewards are reasonably set according to the optimization objective, and the optimal routing solution method based on Q -learning is given. Simulation results and numerical analysis show that the proposed scheme outperforms the benchmark scheme in terms of network survival time, transmission delay, and packet loss rate.

Keywords: Wireless Sensor Networks for substations; routing policy; Markov Decision Process (MDP); Q -learning; network performance optimization

现如今, 电网的快速发展和变电设施数量及规模的不断扩大, 对大量设备的状态感知、监测以及控制具有重大挑战。WSN 因其成本低廉、实时感知、数据安全性高的特点, 具有广阔的应用前景, 被广泛应用于变电站的设备状态感知及数据传输中, 做到了无人值守, 实时监控。变电站的场景一般是块状区域, 近似方形或者圆形, 覆盖范围是 100~500 m 左右, 各类传感器, 如温湿度传感器、水位传感器等, 往往被部署于变电设备上^[1], 传感器具有数据采集模块、数据处理模块和无线通信模块, 可以采集数据和转发数据。传感器将采集到的数据直接发往汇聚节点或通过多个传感器构成的路由路径发往汇聚节点, 汇聚节点又称作无线传感器网络中的网关节点, 它将从网络中接收到的数据汇总并发送至电力网络的接入节点, 从而实现电力网络系统对局部的监测。变电站场景下的无线传感器网络中的传感器具有结构简单、易于部署的特点, 在电力系统环境下, 为了避免线路问题, 同时提高传感的高效性, 传感器与传感器之间或传感器与汇聚节点间都采用无线通信的方式。由于传感器往往采用电池供电, 考虑到节点能量有限的特点, 如何在提升网络性能的同时做到均衡能耗, 延长网络寿命, 保证传输时的网络质量一直是 WSN 有关研究中的主要问题。

许多学者在这一方面做了大量的研究, 层次路由是一种良好的解决方案, 该类方案大多基于分簇的思想,

收稿日期: 2013-04-24; 修回日期: 2013-07-02

基金项目: 国网上海市电力公司科技项目(B30931230003)

经典的分簇算法有低功耗自适应分簇路由算法(Low Energy Adaptive Clustering Hierarchy, LEACH)、传感器信息系统能量高效聚集协议(Power-Efficient Gathering in Sensor Information Systems, PEGASIS)等, 现有的许多方案也是基于这些算法进行了改进。Harun 等将传统的遗传算法与LEACH算法结合^[2], 使用遗传算法来完成最优的簇头选择, 大大增加了网络的存活时间与传输数量。文献[3]中提出一种能耗均衡的分簇算法, 有效地缓解了节点能耗。层次分簇路由算法在大规模WSN中有着良好的表现, 但是在变电站场景中, 由于变电站场合的规模适中, 且考虑到成本问题, 大多数节点的算力较低, 往往不能成为簇首节点的问题、传感器传输距离问题等, 在变电站的场景中往往使用平面路由协议。

在平面路由中, 网络中的节点往往具有相同的地位, 传感器节点在感知状态的同时, 也需要协同构成网络, 将某一时刻某一节点感知到的关键数据按照一定的路由规则上传到汇聚节点。在中小规模的网络中, 平面路由由成本更低, 相对层次路由更有优势。然而以泛洪路由为代表的路由策略虽然具有维护开销小的优点, 但是存在显著的资源浪费问题。为了对平面路由的性能进行优化, 文献[4]提出了一种将启发式蚁群算法应用到传统按需距离矢量路由协议(Ad hoc On-Demand Distance Vector, AODV)算法中, 有效地优化了通信效率。文献[5]将蜂群算法应用到改进的节能链路由(Improved Energy Efficient Chain Based Routing, IEECBR)中, 显著地延长了网络的寿命, 提高了传输效率。然而考虑到启发式算法需要大量的计算与迭代, 不具有计算特性的节点很难承担这类计算工作。

近年来, 人工智能领域的发展为优化问题提供了新的解决思路, 强化学习作为人工智能的重要分支, 已经越来越多被用于工业优化领域的研究中。已经有部分学者将强化学习Q-learning运用于WSN的优化研究, PK Donta等在文献[6]中研究了一种应用Q-learning解决网络拥塞的方法, 该方案是中心化的Q-learning, 由汇聚节点实时捕捉网络状态并更新Q表, 然而当网络状态发生改变时, 中心化的Q-learning难以捕捉到各个节点的状态改变, 并且频繁搜集状态数据必然会大量浪费网络的能量。文献[7]提出了一种去中心化的Q-learning路由算法, 但在其算法的设计中缺少探索, 容易收敛到局部最优。文献[8]将Q-learning应用于可充电传感器网络的路径规划中, 与传统的充电算法相比具有一定的优势。由于无线传感器网络中的数据转发过程容易被看做一个MDP, 且Q-learning算法在具有不同位置和数量的WSN中均有良好的收敛性能, 能够感知网络的状态从而实时改变路由策略, 提升网络性能, Q-learning算法简单直观易于实现, 传感器节点只需要采取简单的计算设备即可实现, 因此本文将结合变电站场景下的WSN的实际需求, 使用Q-learning路由作为路由方案。

尽管众多优秀学者在WSN的路由研究中已经完成了一些工作, 但是并没有考虑到变电站WSN场景下的实际需求和特性, 部分研究中采用的中心化Q-learning路由算法, 并不适用于变电站场景中配置有大量独立转发接收数据节点的WSN。本文将变电站中的WSN路由发送过程抽象为一个MDP, 并提出一种去中心化的Q-learning方案, 即每个节点能够根据环境的状态自己更新Q表。基于变电站无线传感器网络中的实际需求定义了一种结合能耗、时延, 网络丢包率的优化目标, 并根据优化目标合理设置奖励, 基于 ϵ -贪婪原则增大对动作的搜索, 对优化问题进行求解。仿真结果与数值分析表明, 该算法与基准算法相比, 在网络存活时间、网络时延、网络丢包率方面的性能均有所提高。

1 变电站WSN模型

1.1 WSN模型

在WSN的研究中, WSN可以抽象为图的形式, 表示为图 $G = \langle V, E \rangle$ 。假设整个网络中一共分布 N 个传感器节点, 每个节点具有感知收集数据、转发数据的功能, $V = \{v_1, v_2, \dots, v_N\}$ 表示所有传感器节点。 $E = \{e_{ij} \mid \|q(v_i) - q(v_j)\|_2 < R, v_i, v_j \in V\}$ 表示任意2个在通信距离内的传感器节点间的边的集合。其中 $q(v_i) = (x_i, y_i)$ 是节点 i 的位置向量, R 是节点的通信半径, 节点在被部署至网络后具有固定位置向量, 且每个节点被分配唯一的标识编号, 分别为1到 N , WSN中有一个汇聚节点, 编号为 $N+1$, 传感器采集数据后, 通过自组网将数据发送给汇聚节点。变电站WSN结构如图1所示。

1.2 能量模型

采用WSN中的一种被广泛使用的网络传输能量消耗模型, 在

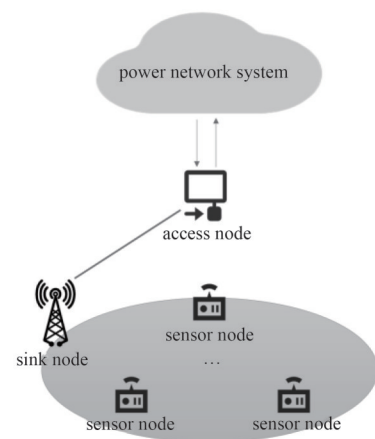


Fig.1 WSN structure of substation
图1 变电站WSN结构

一次数据传输中, 从节点 i 向节点 j 发送长度为 l 的数据需要的能量可以表示为^[9]:

$$E_T(i,j) = \begin{cases} lE_{\text{elec}} + l\epsilon_{\text{fs}}d_{i,j}^2, & d_{i,j} < d_0 \\ lE_{\text{elec}} + l\epsilon_{\text{mp}}d_{i,j}^4, & d_{i,j} \geq d_0 \end{cases} \quad (1)$$

式中: E_{elec} 为功耗系数, 表示发送或接收 1 bit 需要消耗的能量; ϵ_{fs} 、 ϵ_{mp} 为与传输设备有关的能耗系数; 节点 i 和 j 之间的传输距离 $d_{i,j} = \|\mathbf{q}(v_i) - \mathbf{q}(v_j)\|_2$; 传输距离阈值 $d_0 = \sqrt{\epsilon_{\text{fs}}/\epsilon_{\text{mp}}}$ 。节点 j 接收长度为 l 的数据需要的能量可以表示为:

$$E_R = lE_{\text{elec}} \quad (2)$$

假设每个节点的初始能量为 E_0 , 第 j 个节点的当前剩余能量为 E_j , 则使用已消耗能量的比例来表示节点的剩余能量为:

$$E_{\text{consume},j} = \frac{E_0 - E_j}{E_0} \quad (3)$$

在传输的过程中, 节点应当选择已消耗能量较少的节点发送, 避免网络过早死亡, 式(3)的取值在 0~1 之间。

1.3 时延模型

在某一条传输路径中, 将数据包从任意一节点发送到汇聚节点的总时延可以表示为^[10]:

$$D = k \times (D_{\text{hop}} + D_{\text{others}}) \quad (4)$$

式中: k 为该路径传输过程中数据的转发次数; D_{hop} 和 D_{others} 分别为传输的时延、计算排队等其他因素消耗的时间, 两者之和取定值 10 ms, 因此, 路径的传输时延与路径上的节点个数成正比。

1.4 丢包率模型

采用如下的路径丢包率计算模型: 假设从源节点到汇聚节点构成的传输路径的所有边构成的集合为 S , 若该路径中任一条边 e 的丢包率为 $lost(e)$, 则这条路径上的丢包率可以表示为^[11]:

$$Lost(S) = 1 - \prod_{e \in S} (1 - lost(e)) \quad (5)$$

即, 只要一条路径上任何一条边发生了丢包, 该路径就发生了丢包, 路径的丢包率可以由确定概率 1 减去每条边都没有丢包的概率。在无外界干扰的条件下, 每一条边上的初始丢包率相同, 故该丢包率取决于传输路径上的节点个数。

1.5 优化目标

优化目标使得网络的传输代价最小, 从源节点到汇聚节点构成一条路径, 在这条路径中, 任意节点 i 将数据发往 j 的一跳构成一条边 e_{ij} , 定义该跳的能量代价函数为:

$$cost(e_{ij}) = w_{\text{tran}} E_T(i,j) + w_{\text{consume}} E_{\text{consume},j} \quad (6)$$

$E_T(i,j)$ 是由式(1)计算的发送能量, 若发送的耗能越大, 其数值越大, 代价就越大。 $E_{\text{consume},j}$ 由式(3)计算, 为接收节点 j 已消耗能量, 如果选择了已经消耗能量过多的节点发送, 则代价就越大。 w_{tran} 、 w_{consume} 分别是 $E_T(i,j)$ 、 $E_{\text{consume},j}$ 对应的权重系数。

优化目标是让一条路径 S 的总代价最小, 优化目标可以表示为:

$$\min w_{\text{Delay}} Delay(S) + w_{\text{Lost}} Lost(S) + \sum_{e_{ij} \in S} cost(e_{ij}) \quad (7)$$

$$\text{s.t. } C_1: E_i, E_j > 0, \forall e_{ij} \in S, v_i, v_j \in V \quad (8)$$

$$C_2: \|\mathbf{q}(v_i) - \mathbf{q}(v_j)\|_2 \leq R, \forall e_{ij} \in S, v_i, v_j \in V \quad (9)$$

式中: $\sum_{e_{ij} \in S} cost(e_{ij})$ 表示传输路径上所有节点传输的代价和; w_{Delay} 、 w_{Lost} 分别为传输时延 $Delay(S)$, 丢包率 $Lost(S)$ 对应的权重系数。限制条件 C_1 表明所有传输节点和接收节点的能量都必须大于零, 其中 E_i 表示节点 i 已消耗的能量, E_j 表示节点 j 已消耗的能量; 限制条件 C_2 表明路径中两两传输的节点必须在传感器通信半径 R 内。

2 基于 Q-learning 的路由方案

2.1 MDP

为了求解上述优化问题，将数据包转发的过程抽象为一个 MDP，并使用强化学习中的 Q-learning 算法求解。将一次发送过程中的数据包看作智能体，将一次数据包的转发看作是智能体执行动作转向的下一节点，依据转发过程的代价设置奖励。整个 MDP 的过程可以表示为 $MDP = \{\mathcal{S}, \mathcal{A}, \mathcal{R}\}$ ，其中各个符号表示的含义如下：

- 1) 状态空间 $\mathcal{S} = \{s_1, s_2, \dots, s_N\}$ ，表示数据包当前到达的节点编号；
- 2) 动作空间 $\mathcal{A} = \{a_1, a_2, \dots, a_N\}$ ，表示数据包发送的下一个节点编号；
- 3) 概率空间 P 表示状态之间的转移概率，在大部分问题中，该空间难以直接预测，所以往往采取智能体与环境直接交互来代替概率空间的估算。
- 4) 奖励空间 $\mathcal{R} = \{r_{ij} | i \in V, j \in V, i \neq j\}$ ，数据包从节点 i 到达节点 j 的奖励可以设置为：

$$r_{ij} = -\eta_1 \frac{\|q(v_i) - q(v_j)\|_2}{R} - \eta_2 E_{\text{consume},j}(t) - \eta_3 \quad (10)$$

该奖励是 3 个部分的权重之和，由于传输能量仅与距离有关，故式(6)的第 1 项可以看作距离的归一化值乘权重 $-\eta_1$ ，即传输节点之间的距离除以最大通信半径乘上对应权重。式(6)的第 2 项是消耗能量占比，乘以权重 $-\eta_2$ 。由于式(7)中时延与丢包率只和网络的跳数有关，传输跳数越多，式(4)、式(5)的值就越大，故直接加入权重 $-\eta_3$ ，当一次传输路径中的跳数越多，该项累计值就越小，奖励越低。

2.2 Q-learning 路由

强化学习 Q-learning 通过计算 Q 表估算 Q 函数，可以用 $Q(s, a)$ 表示，它表示在状态 s 执行动作 a 可以产生的价值。Q 表的更新过程可以描述为：智能体在状态 s 执行了动作 a ，获得了奖励 r 并转移到了下一个状态 s' 。基于该过程，智能体由式(11)更新 Q 表：

$$Q(s, a) \leftarrow Q(s, a) + \alpha(r + \gamma \max_{a \in A'} Q(s', a) - Q(s, a)) \quad (11)$$

在当前状态下的动作的选择中，智能体会基于 Q 表，大概率选择使得 Q 值最大的动作来执行。本文采用的 Q-learning 算法的更新过程可以由图 2 描述，每个节点记录邻居信息生成一张 Q 表作为路由表并依据此选择下一跳节点，Q 表的更新过程由节点自身完成，这是一种去中心化的 Q-learning 算法^[12]。当数据包抵达节点 i 时，节点 i 将向它的所有邻居节点发送包装在控制数据包内的学习请求，邻居节点收到学习请求后，估算奖励值，根据自己的 Q 表计算 $\max_a Q(s, a)$ ，并将奖励值和 $\max_a Q(s, a)$ 封装成控制数据包内发回节点 i 。每收到一个邻居节点的返回数据，节点 i 就根据式(9)更新自己的 Q 表。于是，基于 Q-learning 的路由算法的一次数据转发如表 1 所示。

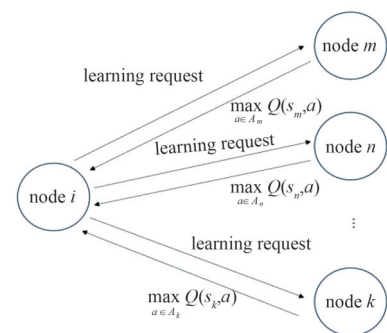


图 2 Q-learning routing
Fig.2 Q-learning 路由

表 1 WSN 中基于 Q-learning 的路由

Table1 Q-learning based routing in WSN

algorithm: Q-learning based routing in WSN

- 1) the source node i generates data with a length of l ;
- 2) state $s \leftarrow i$;
- 3) while $s \neq s_{\text{sink}}$;
- 4) if the neighboring nodes of the current node include the sink node, send the data to the sink node and turn to step 9);
- 5) send learning packets of length l_c to each neighboring node, after receiving the data returned by neighboring node s' , calculate reward r according to (10), obtain $\max_{a \in A} Q(s', a)$ by current Q-table, update the Q-table of node s according to equation(11);
- 6) select action a based on the ϵ -greedy principle in equation(12), send the data packet to the next node with the number a ;
- 7) update state $s \leftarrow a$, calculate $E_T(s, a)$ according to equation(1), calculate E_R according to equation(2), update node energy $E_s \leftarrow E_s - E_T(s, a), E_a \leftarrow E_a - E_R$;
- 8) end while;
- 9) end

在转发的过程中，动作选取依据 ϵ -贪婪原则，在每次智能体选取动作时，通常会选取 Q 表中使得动作价值

最大的动作, 但有 ε 的概率采取随机动作, 即随机选择下一跳节点, 这样做的目的是增大搜索范围, 防止算法陷入局部最优解, 基于该思想的动作选取可以表示为:

$$a = \begin{cases} \max_a Q(s, a), w.p. 1 - \varepsilon \\ \text{rand } a, a \in \mathcal{A}, w.p. \varepsilon \end{cases} \quad (12)$$

3 仿真与分析

3.1 仿真参数

为了验证本文提出的 Q -learning 算法路由的有效性, 使用 Matlab 构建仿真环境并进行仿真, 与基准算法进行性能对比。

仿真的环境可以描述为, 在 $100 \text{ m} \times 100 \text{ m}$ 的环境中, 随机分布 N 个无线传感器节点, 每个节点位置固定, 拥有唯一的标识号, 汇聚节点位于区域的中心, 在每一次的数据感知采集中, 某传感器采集了设备的状态数据, 并通过某一路由路径发送至汇聚节点, 当网络死亡时, 仿真结束, 本文将第一个节点死亡作为整个网络死亡的标志, 仿真所使用的参数见表 2。

表 2 仿真参数
Table 2 Simulation parameters

symbol	value	meaning
l/bit	4 000	packet size
l_l/bit	100	learning packet size
$E_{\text{elec}}/(\text{J} \cdot \text{bit}^{-1})$	50×10^{-9}	power consumption coefficient
$\varepsilon_{\beta}/(\text{J} \cdot \text{bit}^{-1} \cdot \text{m}^{-2})$	10×10^{-12}	amplifier coefficient
R/m	30	communication radius
α	0.8	learning rate
γ	0.9	discounted factor
ε	0.1	random action probability
η_1	0.5	weight1
η_2	0.5	weight2
η_3	0.9	weight3

3.2 仿真结果

为了验证本文算法的有效性, 在本文构建的仿真环境中, 同时使用本文构建的 Q -learning 路由与基准路由协议通过协商传输信息的传感器协议 (Sensor Protocols for Information via Negotiation, SPIN) 和最短路径路由 (Shortest Path Routing, SPR) 作对比。在 SPIN 路由协议中, 发送数据的源节点首先向邻居节点发送广播数据包, 广播数据包大小远小于传输数据大小, 若邻居节点不是汇聚节点, 则继续向邻居节点转发广播数据包, 当汇聚节点接收到广播数据包时, 会发送请求数据包给源节点, 源节点再将数据发送给汇聚节点。在 SPR 路由协议中, 传感器依据自身的位置信息, 以 Dijkstra 算法为基础建立自己到汇聚节点的最短路径, 依据此路径发送数据。

如图 3 所示, 取 $N=100$, 一共进行了 60 次仿真, 在每次仿真中, 传感器节点的位置都是随机分布的, 所以网络存活时间在一定范围内波动。可以看出, SPIN 算法下的网络存活时间是最短的, 这是因为 SPIN 算法在执行的过程中需要广播大量数据包, 导致发送能耗较大, 减少了网络存活时间。SPR 算法依据 Dijkstra 算法建立最短路径, 但是当最短路径建立后, 会出现一条路径被高频传输的情况, 该路径上的节点的转发次数就会更加频繁, 同样会造成网络存活时间的减少。由于 Q -learning 算法在进行计算时考虑了节点的剩余能量, 在发送数据包时既考虑了发送的能耗, 也尽量避免选择了剩余能量少的节点发送, 所以大大延长了网络存活时间, 采用本文提出的 Q -learning 路由算法下的网络在第一个节点死亡之前平均完成了 50 000 次数据发送, 而 SPR、SPIN 两种基准算法分别进行了平均 20 000、10 000 次数据发送后网络死亡, 也就是说, 采用 Q -learning 路由的网络具有更长的网络存活时间。

图 4 说明了完成相同的任务 3 种方法所需要的平均时间, 设定每次仿真网络采集并传输 5 000 次数据, 共进行 60 次仿真,

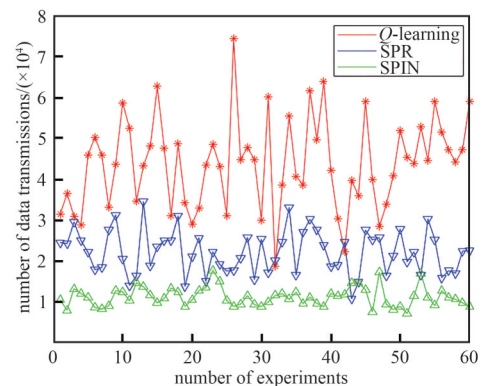


Fig.3 Network survival time

图 3 网络存活时间

由于本文的 Q-learning 路由减少了路由的转发次数，所以在相同的网络传输次数下，平均处理时间为 921.3 s，显著低于相对 2 种基准算法的平均处理时间，分别为 1 450 s 和 2 011 s。

由图 5 看出，随着节点数量的增加，采用 Q-learning 作为路由方案的网络寿命不断增加，而 2 种基准算法无法做到均衡网络的能量消耗，从而会造成某些频繁转发数据的节点提前耗尽能量，网络寿命被限制在一定范围内。

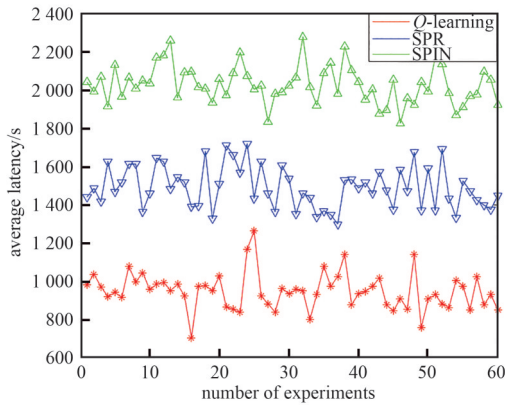


Fig.4 Average latency
图4 平均时延

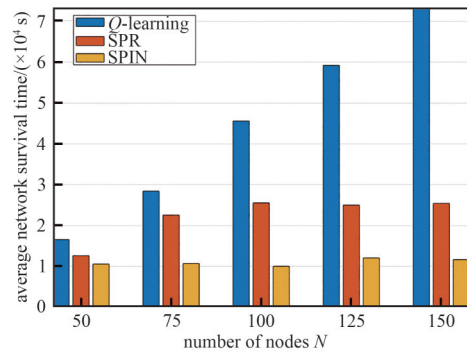


Fig.5 Network survival time with number of nodes
图5 网络存活时间随节点数量变化

5 000 次数据传输中，随节点数量变化下几种方法的平均处理时间如图 6 所示，可以看到，随着节点数量的增加，SPR 的处理时间增加十分迅速，这是因为最短路径算法在构建路由表时，节点数量增加了，导致传输数据需要的转发次数大量增加。而 SPIN 算法因为节点数量的增加导致需要广播的数据包的数量也会增加，所以平均时延也在增长。而采取 Q-learning 的路由仅在节点数量较少时性能劣于 SPR，在节点数量增加时由于考虑到了减少路由跳数而使得传输时延较低。

同样地，如图 7 所示，由于在发送数据包时考虑了减少传输时经过的节点，因此 Q-learning 的路由丢包率一直少于其他 2 种算法，且随着节点数量的改变，3 种算法的丢包率均没有较大变化。

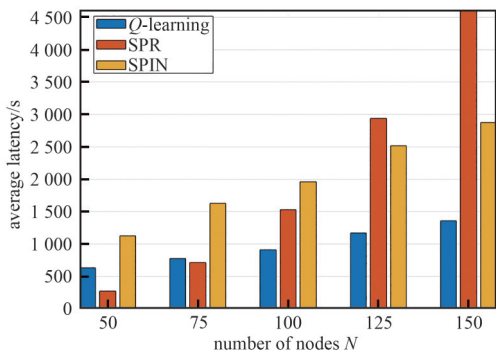


Fig.6 Latency varying with the number of nodes
图6 时延随节点数量变化

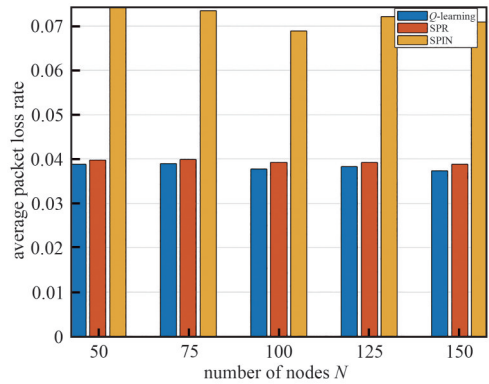


Fig.7 Variation of packet loss rate with number of nodes
图7 丢包率随节点数量变化

4 结论

本文针对变电站中无线传感网络的路由问题，结合变电站场景下无线传感网络的实际需求，提出了一种基于 Q-learning 的路由算法，以最小化网络的传输代价函数为目标，该算法能够很好地均衡网络中的能耗，减少传输时的路由跳数和传输距离，减少网络的传输时间。仿真结果表明，所提出的路由算法在网络存活时间、传输时延、丢包率方面均优于基准算法。

参考文献：

[1] 路永玲,王真,薛海,等. 面向输变电场景的无线传感网体系架构设计[J]. 信息技术与标准化, 2023(5):59-66. (LU Yongling, WANG Zhen,XUE Hai,et al. Architecture design of wireless sensor network for power transmission scenarios[J]. Information

- Technology and Standardization, 2023(5):59–66.)
- [2] HARUN H B, ISLAM M S, HANIF M. Genetic algorithm for efficient cluster head selection in LEACH protocol of Wireless Sensor Network[C]// 2022 International Conference on Advancement in Electrical and Electronic Engineering(ICAEEEE). Gazipur, Bangladesh: IEEE, 2022:1–6. doi:10.1109/ICAEEEE54957.2022.9836352.
- [3] 李虹飞, 申玉霞. 无线传感网络中一种能耗均衡的分簇路由算法[J]. 火力与指挥控制, 2022, 47(10):159–165. (LI Hongfei, SHEN Yuxia. An energy-equalizing cluster routing algorithm in wireless sensor networks[J]. Fire and Command and Control, 2022, 47(10):159–165.)
- [4] XU Jian. A modified AODV routing protocol using in WSN based on ant colony algorithm[C]// 2021 2nd International Conference on Electronics, Communications and Information Technology(CECIT). Sanya, China: IEEE, 2021:87–90. doi:10.1109/CECIT53797.2021.00023.
- [5] SHARMA D, KULKARNI S. Network lifetime enhancement using improved honey bee optimization based routing protocol for WSN[C]// 2018 Second International Conference on Inventive Communication and Computational Technologies(ICICCT). Coimbatore, India: IEEE, 2018:913–918. doi:10.1109/ICICCT.2018.8473267.
- [6] DONTA P K, AMGOTH T, ANNAVARAPU C S R. Congestion-aware data acquisition with Q-learning for wireless sensor networks[C]// 2020 IEEE International IOT, Electronics and Mechatronics Conference(IEMTRONICS). Vancouver, BC, Canada: IEEE, 2020:1–6. doi:10.1109/IEMTRONICS51293.2020.9216379.
- [7] SU Xing, REN Yiting, CAI Zhi, et al. A Q-learning based routing approach for energy efficient information transmission in wireless sensor network[J]. IEEE Transactions on Network and Service Management, 2022, 20(2): 1949–1961. doi: 10.1109/TNSM.2022.3218017.
- [8] 刘洋, 王军, 吴云鹏. 改进 Q-Learning 的 WRSN 充电路径规划算法[J]. 太赫兹科学与电子信息学报, 2022, 20(4):393–401. (LIU Yang, WANG Jun, WU Yunpeng. Improved Q-Learning WRSN charging path planning algorithm[J]. Journal of Terahertz Science and Electronic Information Technology, 2022, 20(4):393–401.) doi:10.11805/TKYDA2020729.
- [9] ZHANG Degan, LI Guang, ZHENG Ke, et al. An energy-balanced routing method based on forward-aware factor for wireless sensor networks[J]. IEEE Transactions on Industrial Informatics, 2013, 10(1):766–773.
- [10] 孙毅, 刘浩程, 曾璐琨, 等. 面向配电通信网的 WMSNs 多路径 QoS 路由算法[J]. 计算机应用研究, 2016, 33(11):3387–3390. (SUN Yi, LIU Haocheng, ZENG Lukun, et al. A multipath QoS routing algorithm for WMSNs for distribution communication networks[J]. Computer Application Research, 2016, 33(11):3387–3390.)
- [11] 杨佳, 段琪玥, 许强, 等. 一种面向配电通信网 WSN 分簇路由优化算法[J]. 重庆理工大学学报(自然科学), 2022, 36(9):187–194. (YANG Jia, DUAN Qiyue, XU Qiang, et al. A cluster routing optimization algorithm for WSN in power distribution communication network[J]. Journal of Chongqing University of Technology(Natural Science), 2022, 36(9):187–194.)
- [12] YUN W K, YOO S J. Q-learning-based data-aggregation-aware energy-efficient routing protocol for Wireless Sensor Networks[J]. IEEE Access, 2021(9):10737–10750. doi:10.1109/ACCESS.2021.3051360.

作者简介:

赵 锴(1989–), 男, 硕士, 工程师, 主要研究方向为电力数字孪生、电力数字化转型. email:zk999000@163.com.

丛 尤嘉(1993–), 男, 学士, 工程师, 主要研究方向为电力数字化转型、电网运维监控方向.

沙 杰(1976–), 男, 大专, 高级技师, 主要研究方向为变电站二次设备运维检修、电力数字化转型.