

文章编号: 2095-4980(2025)05-0476-06

基于终端接入网的全时全域电力业务接入路由

王忠钰¹, 卢志鑫¹, 吕国远¹, 刘乙召¹, 李霜冰², 岳顺民², 韩雨阳³

(1. 国网天津市电力公司 信息通信公司, 天津 300140; 2. 国网天津市电力公司, 天津 300010;
3. 天津大学 智能电网教育部重点实验室, 天津 300072)

摘 要: 随着能源互联网的迅速发展, 车网互动等新型电力业务对服务质量的要求日益严格, 给电力终端接入网带来诸多挑战。针对终端接入网多种通信技术覆盖范围重叠, 通信方式可选范围广泛导致的资源浪费和网络性能下降问题, 提出一种基于随机森林的通信方式选择算法(RF-CMS)。通过随机森林对海量多样的新型电力业务进行智能分类, 为其选择最合适的通信方式; 然后, 从流量负载和通信质量角度出发, 利用多智能体近端策略优化(MAPPO)算法为电力业务动态分配路由, 确保各种终端业务数据(如测量信息、控制信息)能够在接入网中及时、准确地传输, 从而保证电力网络状态的全景全域可观可控。将所提算法与仅基于 MAPPO 的路由算法在平均端到端时延、负载均衡度等方面进行对比, 证明了所提算法的有效性。

关键词: 电力终端接入网; 随机森林; 多智能体近端策略优化; 通信方式; 路由优化

中图分类号: TM732

文献标志码: A

DOI: 10.11805/TKYDA2024247

Omni-directional power service access routing based on terminal access networks

WANG Zhongyu¹, LU Zhixin¹, LYU Guoyuan¹, LIU Yizhao¹, LI Shuangbing², YUE Shunmin², HAN Yuyang³

(1. Information & Telecommunication Company, State Grid Tianjin Electric Power Company, Tianjin 300140, China;

2. State Grid Tianjin Electric Power Company, Tianjin 300010, China;

3. Key Laboratory of Smart Grid (Ministry of Education), Tianjin University, Tianjin 300072, China)

Abstract: With the rapid development of the energy Internet, new power services such as vehicle-network interaction have increasingly stringent requirements on service quality, which brings many challenges to the power terminal access network. To address the problem of resource waste and network performance degradation caused by the overlapping coverage of multiple communication technologies and a wide range of communication methods in terminal access network, a Communication Method Selection algorithm based on Random Forest (RF-CMS) is proposed, which intelligently categorizes a large number of diverse new electric power services and selects the most suitable communication methods for them through Random Forest mode. Then, the Multi-Agent Proximal Policy Optimization (MAPPO) algorithm is employed to dynamically allocate routes for the power services from the viewpoint of traffic loading and communication quality to ensure that various terminal service data (e. g., measurement information, control information) can be transmitted timely and accurately in the access network. The effectiveness of the proposed algorithm is validated by comparing it with the routing algorithm based solely on MAPPO in terms of average end-to-end delay and load balancing degree.

Keywords: power terminal access network; Random Forest (RF); Multi-Agent Proximal Policy Optimization (MAPPO); communication method; routing optimization

收稿日期: 2024-05-29; 修回日期: 2024-07-11

引用格式: 王忠钰, 卢志鑫, 吕国远, 等. 基于终端接入网的全时全域电力业务接入路由[J]. 太赫兹科学与电子信息学报, 2025, 23(5): 476-481. DOI: 10.11805/TKYDA2024247.

Citation format: WANG Zhongyu, LU Zhixin, LYU Guoyuan, et al. Omni-directional power service access routing based on terminal access networks[J]. Journal of Terahertz Science and Electronic Information Technology, 2025, 23(5): 476-481. DOI: 10.11805/TKYDA2024247.

随着能源互联网的迅速发展和新型配电网建设的持续推进^[1]，新型电力系统中分布式新能源的大规模分散式接入给电力终端接入网带来诸多挑战^[2]，如，通信边界更加复杂、覆盖范围更加广泛、终端设备更加丰富，以及服务质量(Quality of Service, QoS)要求更加严格等。为应对这些挑战，须提高电力终端接入网对光纤、无线和电力线载波等多种可靠通信方式的利用效率，从而确保配用电业务信息的高效传输，满足各类应用需求。

近年来，基于强化学习的路由算法得到了广泛关注^[3-4]。LIN 等^[5]在 Q-learning 算法的基础上，提出一种基于强化学习的 QoS 感知自适应路由算法(QoS Aware Adaptive Routing, QAR)，实现了时间高效、自适应和网络服务质量优质的数据包转发，减少了信令时延。文献[6]利用深度确定性策略梯度(Deep Deterministic Policy Gradient, DDPG)算法实现动态网络的路由优化。张庚等^[7]提出一种基于强化学习的电力光通信网风险均衡路由算法，实现了电力光通信网风险均衡化。但这些算法缺乏对业务通信方式的考虑，在电力终端接入网中，由于多种通信技术覆盖范围重叠，业务终端的通信方式可选范围广泛，会导致资源浪费和网络性能下降。

为此，本文提出一种基于随机森林的通信方式选择算法(RF-CMS)。通过随机森林模型对海量多样的新型电力业务进行智能分类，为其选择最合适的通信方式；然后，从流量负载和通信质量角度出发，利用多智能体近端策略优化(MAPPO)算法为电力业务动态分配路由，确保各种终端业务数据(如测量信息、控制信息)能够在接入网上及时、准确地传输，从而保证电力网络状态的全时全域、可观可控。

1 基于随机森林的通信方式选择算法

电力业务最佳通信方式的选择可建模为具有连续特征值的多分类问题。在解决分类问题时，随机森林是一种简单而有效的方法。随机森林算法是一种基于决策树的集成学习算法^[8]，通过装袋集成算法对训练数据随机抽样，构建多棵决策树，并通过投票或取平均值的方式进行预测。与迭代二叉树三代(Iterative Dichotomiser 3, ID3)、C4.5 等其他决策树算法相比，分类与回归树(Classification and Regression Trees, CART)算法基于基尼不纯度准则进行分裂选择，能够使决策树更好地适应连续特征的数据分布，更适合处理具有连续特征值的多分类问题，因此每个决策树模型采用 CART 决策树算法对单个样本进行分类。

综合考虑终端接入网承载的光纤通信、无线通信等通信资源的传输特性，根据电力业务的 QoS 需求和电力设备的状态构建初始特征集 X 。与无线路径相比，有线路径不受电磁信号干扰、障碍物和传输距离的影响，通常更稳定，带宽更高，延迟更低，在远距离传输场景下的穿透能力更强，更适合业务重要度高、时延要求高的业务。而无线路径可为移动设备提供更大的灵活性和便利性，并且可覆盖到有线网络无法到达的地方。根据以上传输特性、业务性能要求和路由优化目标，通信方式的选择受 5 个因素的影响：业务重要度、业务的服务质量、设备接收到的信号强度、设备间距离(Equipment Distance, ED)和设备的移动性(Mobility, Mod)。

1) 业务重要度：目前针对电力通信业务重要度评价已有较为成熟的研究成果，根据文献[9]，各安全分区电力通信业务重要值如表 1 所示。

表1 业务重要值

Table1 Important values of power service

service attribute	security zone	service name	criticality
key operation zone services	zone I	transmission line relay protection	1.000
		power grid stability control	0.948
		dispatch telephone system	0.443
		dispatch automation system	0.736
		wide-area power grid measurement	0.894
	zone II	relay protection information management	0.408
		telemetry & energy measurement	0.490
		electricity market trading & operation	0.523
transactional management services	zone III	lightning monitoring & positioning	0.243
		substation integrated monitoring	0.176
	zone IV	power quality management	0.428
		video conferencing	0.315
		information disaster recovery	0.165
		administrative telephone system	0.152

2) 业务的服务质量要求：对于归一化处理过的业务，允许的最大传输时延为 $D_{\max}$ ，最大丢包率为 $L_{\max}$ ，最小带宽需求为 $B_{\min}$ ，则设计业务的综合服务质量指标(Service Index, SI, 以 I_s 表示)为：

$$I_s = (1 - D_{\max} - L_{\max} + B_{\min}) \tag{1}$$

3) 设备接收到的信号强度: 根据接收信号强度指标(Received Signal Strength Indicator, RSSI)评估设备所处地理位置的信号遮蔽效应与电磁干扰强度。在电力通信网中, 设备接收到的信号强度指标(RSSI)低于-90 dBm时, 信号质量较差, 可能会导致丢包、数据传输错误等问题, 通常不适合使用无线传输, 将此类设备划分为二级地理等级; 接收信号强度高于-90 dBm的设备则划分为一级地理等级。

4) 设备间距离(ED): 根据拓扑结构中各节点坐标计算得到。

5) 设备移动性(Mod): 移动性设备 Mod 设为 1; 否则为 0。

因此, 特征集 $X=\{Imp, SI, RSSI, ED, Mod\}$ 。根据提取的特征, 使用随机森林算法训练多个决策树模型, 如图 1 所示。通过 Bootstrap 方法进行有放回抽样, 生成 w 个样本集, 并用样本集构成 w 棵决策树, 形成随机森林。决策树的训练流程如下:

1) 输入包含电力业务重要度与综合服务质量指标的业务训练集 D , 基尼指数阈值 G_θ , 特征集 X 。

2) 随机选择特征集 X 的子集作为子特征集 X' , 根据训练数据集, 从根节点开始, 递归地对每个节点进行以下操作, 构建决策树。

a) 设当前节点的训练数据集为 D' , 针对特征集 X' 中的特征, 计算现有特征对各特征值对该数据集的基尼指数。假设第 k 个类别的数量为 C_k , 数据集 D' 的基尼系数为:

$$\text{Gini}(D') = 1 - \sum_{k=1}^s \left(\frac{|C_k|}{|D'|} \right)^2 \quad (2)$$

根据某一特征 x 是否取特征值 a , 计算数据集 D' 的条件基尼指数为:

$$\text{Gini_index}(D', a) = \sum_{v=1}^V \frac{|D_v|}{|D'|} \text{Gini}(D_v) \quad (3)$$

式中: $|D'|$ 为数据集 D' 中的总样本数; $|D_v|$ 为划分后第 v 个分支所包含的样本数。

b) 计算数据集 D' 上所有特征的基尼指数, 选择基尼指数最小的特征 x 作为最优特征。若该特征为离散特征, 则对应的特征值 a 为最优特征值。根据最优特征和最优特征值, 将数据集划分成 D_1 和 D_2 两个子集, 建立当前节点的左右节点, 并从 X 中移除已使用的特征。

c) 若当前特征为连续特征, 则在左右子节点上递归执行特征选择与分裂操作(即步骤 a)和步骤 b)), 直至满足以下任一终止条件: 节点的基尼指数低于预设阈值(表示样本纯度达标); 子数据集 X' 为空(表示无剩余样本可分配); 此时停止递归, 将当前节点标记为叶节点, 最终生成完整的 CART 决策树。

决策树构建过程中不会使用子集中所有的特征属性, 而是随机选择部分特征, 以增加各决策树之间的差异性。

3) 各决策树独立进行训练与预测, 最后通过随机森林中所有决策树组合投票的方式确定最优的通信传输方式, 从而实现电力业务路由通信方式的分类。统计 w 棵决策树的决策结果, 选择票数最多的输出变量 Y 作为最终的分类结果。

$$H(x) = \arg \max_Y \sum_{i=1}^w I(h_i(x) = Y) \quad (4)$$

式中: $H(x)$ 为组合分类模型; $h_i(x)$ 为第 i 棵决策树分类模型; Y 为输出变量, 即最终选择的通信方式; $I(\cdot)$ 为示性函数。

2 基于 MAPPO 的多约束路由优化算法

MAPPO 算法基于集中式训练和分散执行(Centralized Training and Decentralized Execution, CTDE)框架^[10], 在训练阶段允许价值网络访问环境的全局状态信息; 而在执行阶段, 每个代理通过分散的策略网络(基于局部观测

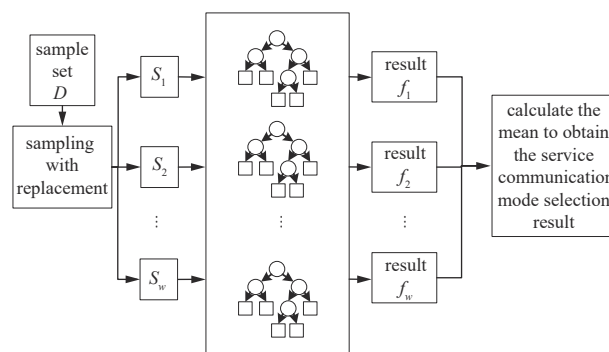


Fig.1 RF-CMS algorithm model

图 1 RF-CMS 模型

信息)执行动作。本文根据 RF-CMS 算法选择通信方式,综合考虑业务时延、带宽、丢包率等 QoS 需求,采用 MAPPO 算法为业务动态分配路由,实现网络负载均衡。

将网络路由优化问题建模为马尔可夫决策过程,使用多元组 $MDP=(S_i, A_i, R_i)$ 描述路由决策过程,并使用智能体代表决策实体,则 S_i 表示第 i 个智能体的状态空间集, A_i 表示第 i 个智能体的动作空间集, R_i 表示第 i 个智能体的奖励函数。本文设置 A、B 两类智能体,分别针对有线资源和无线资源进行路由决策,并为每个设备配置合适的智能体:铺设光缆的设备配置 A 类智能体;铺设无线路径的设备配置 B 类智能体。当电力业务流 $f_k=(s_k, d_k)$ (其中 s_k 、 d_k 分别为源节点和目标节点)到达网络时,智能体开始进行路由决策。由于有线资源和无线资源传输特性的不同, A、B 类智能体的动作空间相同,但状态空间不同。根据以上分析建立决策过程:

1) 状态空间 S_i : 当前网络的节点和链路状态决定了路由的最优决策, A、B 类智能体的状态空间分别为: $S_A=\{D, B_w, L, NIR, W\}$, $S_B=\{D, B_w, L, NIR, W_1\}$ 。其中 D 表示网络时延, B_w 表示链路剩余可用带宽, L 表示网络丢包率, NIR 表示网络的节点空闲率(Node Idle Rate, NIR), W 表示有线带宽资源的剩余量, W_1 表示无线频谱资源的剩余量。

2) 动作空间 A_i : 定义动作为网络的链路选择权重,根据权重使用 Dijkstra 算法求解得到最优路径。

3) 奖励函数 R_i :

本文路由优化目标主要有以下 3 点: a) 网络的时延、丢包率满足业务 QoS 需求; b) 实现网络的流量负载均衡,避免某些节点因为负载过高而导致网络拥塞和时延增加,影响业务的稳定性和可靠性; c) 为不同的电力业务选择合适的链路类型,提升网络资源的利用效率。因此,在各类链路容量不同的网络中,应尽可能使各链路负载流量与链路总容量的比例趋于一致,从而保证网络的负载均衡。本文采用负载均衡度(Load Balance, LB, 以 B_L 表示)衡量网络拓扑 $G(V, E)$ (其中 V 为网络中的路由器等通信设备集合, E 为光纤、无线等通信链路集合)的流量均衡效果。负载均衡度越大,说明流量分布越均匀。

$$B_L = -\eta^*(x, y) \sum_{(x, y) \in E} \log_2 \eta^*(x, y) \quad (5)$$

$$\eta(x, y) = \frac{C(x, y) - l(x, y)}{C(x, y)} \quad (6)$$

式中: $\eta(x, y)$ 为链路 (x, y) 已负载流量占链路总容量的百分比; $\eta^*(x, y)$ 为 $\eta(x, y)$ 的归一化值; $C(x, y)$ 为链路 (x, y) 的总容量; $l(x, y)$ 为链路的剩余容量。

网络的 NIR 对于网络的拥塞程度也有着重要的影响。在网络路由传输数据过程中,从网络结构整体上看,相同网络负载下空闲节点越多,拥塞程度越低。因此,将 NIR 纳入网络均衡的评价指标。当业务到达节点时,若节点繁忙,业务进入队列排队,队列为空时业务立即转发,因此采用节点队列为空时的概率描述节点空闲率。

为检验 RF-CMS 算法的有效性,引入业务资源匹配度(Matching Ratio, MR)描述业务与通信资源的匹配程度,进而衡量各链路类型的负载均衡情况和网络资源利用率。

$$R_M = CE(B_{L, W}, B_{L, W_1}) \quad (7)$$

式中: $CE(B_{L, W}, B_{L, W_1})$ 为 LB_W 和 LB_{W_1} 的交叉熵; $B_{L, W}$ 为有线通信资源的负载均衡度, B_{L, W_1} 为无线通信资源的负载均衡程度。

综上所述,网络的性能主要与网络时延、丢包率、负载均衡情况、节点空闲率和业务资源匹配度 5 个指标有关。

奖励函数 R_i 为:

$$R_i = -\alpha D + \beta R_M - \theta L + \varepsilon B_L + \delta R_{IN}, (\alpha, \beta, \theta, \varepsilon, \delta) \in (0, 1) \quad (8)$$

式中 α 、 β 、 θ 、 ε 、 δ 分别为 5 个指标的权重参数, $\alpha + \beta + \theta + \varepsilon + \delta = 1$ 。

训练结束后, Actor 网络使用目标函数裁剪的更新方法,损失函数 $L_t^{CLIP}(\theta)$ 为:

$$L_t^{CLIP}(\theta) = \hat{E}_t[r_t(\theta) \hat{A}_t^i, \text{CLIP}(r_t(\theta), 1 - \varepsilon, 1 + \varepsilon) \hat{A}_t^i] \quad (9)$$

式中: $r_t(\theta)$ 为新旧策略的概率比值; $\hat{A}_t^i$ 为第 i 个代理的广义优势估计优势值; $\hat{E}_t$ 为对时间 t 的经验轨迹数据求统计期望。

Critic 网络的更新则是利用目标值 y 与当前 Q 值的均方误差构造损失函数，进行梯度更新。

$$L(\theta_i)=E_{s,a,r,s'}\left[\left(Q_i^\mu(s_i,a_i)-y\right)^2\right] \tag{10}$$

式中 $E_{s,a,r,s'}$ 为期望值； $Q_i^\mu(s_i,a_i)$ 为 Critic 网络估计的 Q 值函数。

目标值 y 利用式(11)计算得到：

$$y=r+\gamma Q_i^\mu(s'_i,a'_i)|_{a'_i=\mu'_i(o_i)} \tag{11}$$

式中： r 为 Critic 网络通过访问所有智能体的经验回放池得到的平均奖励值； γ 为折扣因子； u' 为目标策略集合。

3 算法性能分析

为验证所提算法的有效性，采用 OMNET++6.0.1 搭建具有 14 个节点、33 条光链路和 28 条无线链路的通信网络拓扑。每条光链路容量在 500~800 Gbps 间随机分布，无线网络链路数据传输速率设置为 54 Mbps。利用 Pytorch 深度学习框架进行仿真实验和结果分析。仿真中主要的环境参数设置见表 2。

表 2 模拟参数设置
Table2 Simulation parameter setting

parameter	learning rate	discount factor γ	experience replay	number of hidden	gradient clipping	sampling steps per	batch size for
			buffer size	layer neurons	range	update	sampling
value	1×10^{-6}	0.8	10 000	512	0.2	2 048	128

在上述网络拓扑上随机生成 400 条业务，对于随机生成的每一条业务信息，业务的源节点、宿节点和业务类型是随机的。根据随机生成的业务类型，可确定业务其他属性信息(业务重要度和业务带宽等)。针对生成的电力业务，将本文所提的 RF-CMS+MAPPO 算法、MAPPO 算法以及文献[6]的 DDPG 路由算法进行对比实验。

图 2 为本文所提算法与对比算法在训练过程中网络的平均端到端时延变化。由于训练过程中时延数据波动较大，本文将每 100 步的时延数据取均值，绘制了简化的下降曲线。从图 2 可以看出，训练过程中，与 MAPPO 算法相比，本文所提算法能够在终端接入网络中更快地收敛，收敛后的平均端到端时延降低了 15.62%。

图 3 为本文所提算法与对比算法在训练过程中有线链路的平均负载均衡度的变化情况。仿真结果表明，随着训练步数的增加，有线链路组成网络的负载均衡度在波动中上升，并且所提算法的负载均衡度总是高于其他算法，且能够快速收敛到稳定状态。由此可见，本文所提算法通过为电力业务匹配合适的通信方式，提供了更优质的服务质量，有效提高了资源利用效率。

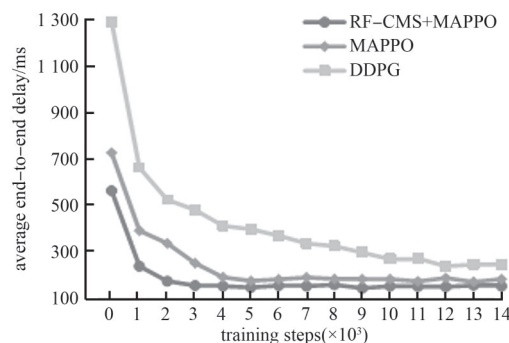


Fig.2 Training process of network average end-to-end delay
图2 网络平均端到端时延训练过程

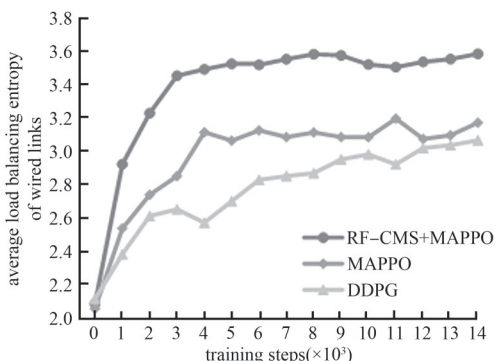


Fig.3 Average load balance entropy training
图3 有线链路的平均负载均衡度训练过程

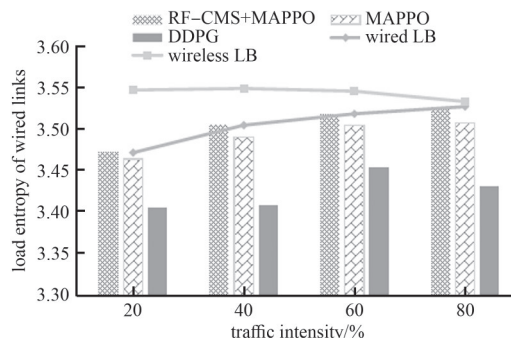


Fig.4 Entropy of network load balancing under different traffic intensities
图4 不同流量强度下的网络负载均衡度

为深入评估所提算法在业务流量强度变化下的性能，进行了不同流量强度下的对比实验。图 4 为所提算法与

对比算法在不同流量强度下的网络负载均衡程度对比结果。从图4中可以看出,随着流量强度的增加,所提算法下接入网中无线链路部分的网络负载均衡程度呈现小幅波动下降趋势,有线网络负载均衡度则显著增加。这是因为在实际电力终端接入网络中,无线网络资源充裕,负载均衡度变化幅度较小。无线资源的补充缓解了有线网络的业务承载压力,能够帮助有线网络达到更好的网络负载均衡状态。值得注意的是,随着负载的加重,有线和无线网络的负载均衡度逐渐趋于一致。说明本文所提算法能够为业务分配合适的通信方式,使终端接入网能够在高负载状态下有效实现网络负载均衡。

4 结论

在能源互联网迅猛发展的背景下,车网互动等新型电力业务对电力终端接入网的服务质量提出了更高要求。针对当前终端接入网中通信技术重叠和通信方式选择不当带来的挑战,本文提出了基于随机森林的通信方式选择算法(RF-CMS)。RF-CMS算法通过智能分类机制,有效识别并适配各类新型电力业务所需的通信方式,显著减少了资源的浪费。通过结合MAPPO算法,实现了电力业务路由的动态优化,确保了数据的高效传输。实验结果表明,相较于单独使用MAPPO算法,RF-CMS与MAPPO的协同应用将平均端到端时延降低15.62%,同时提升了光纤链路的负载均衡性能。所提算法为电力终端接入网提供了一种高效、智能的通信方式选择和路由分配方案,对于提升能源互联网的服务质量和实现网络的全时全域、可控可控具有重要的实际应用价值。

参考文献:

- [1] 刘玮男. 电力系统多能源互联网协同优化控制研究[J]. 今日制造与升级, 2023(10):18-21. (LIU Weinan. Research on collaborative optimization control of multi-energy energy internet in power system[J]. Manufacture & Upgrading Today, 2023(10): 18-21.)
- [2] 朱子旭,刘宣,李丹,等. 面向能源互联网的电力线通信路径搜索研究[J]. 电网与清洁能源, 2023,39(10):63-69. (ZHU Zixu, LIU Xuan, LI Dan, et al. A study on path search of power line communication applied in energy Internet[J]. Advances of Power System & Hydroelectric Engineering, 2023,39(10):63-69.) DOI:10.3969/j.issn.1674-3814.2023.10.008.
- [3] 吴麒,左琳立,丁建,等. 基于Q-learning的自适应链路状态路由协议[J]. 重庆邮电大学学报(自然科学版), 2024,36(5):945-953. (WU Qi, ZUO Linli, DING Jian, et al. Q-learning based adaptive link state routing protocol[J]. Journal of Chongqing University of Posts and Telecommunications(Natural Science Edition), 2024,36(5):945-953.) DOI: 10.3979/j.issn.1673-825X. 202311080358.
- [4] 张华南,李石君,金红. 无线传感器网络强化学习增强路由研究[J]. 应用科学学报, 2024,42(1):83-93. (ZHANG Huanan, LI Shijun, JIN Hong. Research on enhanced routing for reinforcement learning in wireless sensor networks[J]. Journal of Applied Sciences, 2024,42(1):83-93.) DOI:10.3969/j.issn.0255-8297.2024.01.007.
- [5] LIN S C, AKYILDIZ I F, WANG P, et al. QoS-aware adaptive routing in multi-layer hierarchical software defined networks: a reinforcement learning approach[C]// 2016 IEEE International Conference on Services Computing(SCC). San Francisco, CA, USA: IEEE, 2016:25-33. DOI:10.1109/SCC.2016.12.
- [6] LIU Waixi, CAI Jun, CHEN Qinchun, et al. DRL-R: deep reinforcement learning approach for intelligent routing in software-defined data-center networks[J]. Journal of Network and Computer Applications, 2021(177):102865.
- [7] 张庚,王亚男,邢祥栋,等. 电力光通信网中风险均衡路由算法研究[J]. 光通信研究, 2021(1):15-18,35. (ZHANG Geng, WANG Yanan, XING Xiangdong, et al. Research on risk balancing routing algorithm in power optical communication network[J]. Study on Optical Communications, 2021(1):15-18,35.) DOI:10.13756/j.gtxy.2021.01.004.
- [8] 庄保乾,韩路,李晓虎,等. 基于集成深度随机森林算法的智能电厂设备健康评估方法[J]. 计算机测量与控制, 2024,32(8):322-328. (ZHUANG Baoqian, HAN Lu, LI Xiaohu, et al. Health status assessment method of intelligent power plant equipment based on integrated deep random forest algorithm[J]. Computer Measurement & Control, 2024,32(8):322-328.) DOI:10.16526/j.cnki.11-4762/tp.2024.08.046.
- [9] 张冉,蒋慧超,赵子岩,等. 基于多指标最小环路的电力通信网重要节点判别[J]. 电力信息与通信技术, 2023,21(9):75-81. (ZHANG Ran, JIANG Huichao, ZHAO Ziyang, et al. Identification of important nodes in power communication network based on multi-index minimum loop[J]. Electric Power Information and Communication Technology, 2023,21(9):75-81.) DOI:10.16543/j.2095-641x.electric.power.ict.2023.09.11.
- [10] YU C, VELU A, VINITSKY E, et al. The surprising effectiveness of PPO in cooperative multi-agent games[J]. Advances in Neural Information Processing Systems, 2022(35):24611-24624.

(下转第501页)