End-to-End Delay Minimization for IPv6 Power Communication Networks Based on Deep Deterministic Policy Gradient

Xubin Lin[1], Feifei Hu[1*], Liu Wu[1]
[1]Power Dispatching and Control Center, CSG, Guangzhou Guangdong, 510663, China
Corresponding author: Feifei Hu (huff_1024@163.com).

ABSTRACT The IPv6 power communication network is the core infrastructure supporting smart grid dispatch and control, data transmission, and business linkage. In this environment, substations, distributed-energy devices, and edge terminals use IPv6 for unified addressing and end-to-end interconnection, while Traffic Class and Flow Label provide network-layer identifiers for differentiated services and flow-level scheduling. End to end latency is the core indicator for measuring the quality of power communication services, which directly affects the safe and stable operation of key power services such as relay protection, remote measurement and control, and fault alarm. This paper proposes an end-to-end delay minimization method for IPv6 power communication networks based on Deep Deterministic Policy Gradient (DDPG) to address the problems of poor dynamic adaptability, low efficiency in high-dimensional resource optimization, and insufficient convergence of delay in complex business scenarios in existing delay optimization methods. The protocol-aware MDP represents address prefixes, Traffic Class, Flow Label, link load, queue state, and current route information, and jointly adjusts bandwidth share, route weight, and queue service ratio. Thus, IPv6 fields participate directly in differentiated delay control rather than serving only as packet-header descriptions.

INDEX TERMS IPv6; Power communication network; End to end latency; Deep deterministic policy gradient

I. INTRODUCTION

A. RESEARCH SIGNIFICANCE

With the comprehensive promotion of the construction of new power systems, smart grids are rapidly developing towards digitization, networking, intelligence, and ubiquity, while IPv6 provides scalable addressing and unified identification for large-scale power terminals. Compared with fragmented IPv4 address management, hierarchical IPv6 prefixes support direct addressing across dispatch centers, substations, access nodes, and distributed-energy terminals, providing a common network layer for heterogeneous optical, 5G, and power-line carrier links. New energy sources such as wind power and photovoltaics are connected to the grid on a large scale, and energy storage devices and flexible loads are widely connected. The operation status of the power grid presents strong randomness and high dynamism. As the "neural network" of the smart grid, the IPv6 power communication network carries a full range of power services such as grid dispatch automation, remote power monitoring, relay protection, distribution automation, electricity information collection, and new energy grid connected data transmission. It is the core support network that ensures the safe, stable, and efficient operation of the power grid. Unlike traditional civilian communication networks, power communication networks have unique characteristics such as significant differences in business priority, high requirements for transmission reliability, obvious layering of delay sensitivity, relatively fixed network topology but significant dynamic fluctuations in business traffic. Their communication performance directly determines the control accuracy and operational safety of the power grid. IPv6 deployment is therefore not limited to address expansion. Prefix aggregation supplies the routing domain, while Traffic Class and Flow Label provide service-class and flow-level identifiers. Their value depends on a consistent mapping among ingress classification, routing, bandwidth allocation, and queue scheduling, which makes protocol-aware resource control necessary.

End to end latency is one of the core performance indicators of IPv6 power communication networks, which refers to the entire process time of power business data transmission from the sending node to the receiving node, covering data encapsulation, link transmission, node forwarding, data parsing, and other processes. In IPv6 transmission, Traffic Class expresses service priority and Flow Label maintains flow identity across forwarding nodes, allowing latency constraints to be associated with specific service classes and flows. Different power services have strict and differentiated threshold requirements for end-to-end latency: among them, core control services such as relay protection and stability devices require end-to-end latency of less than 10ms, which belongs to ultra-high real-time services; Remote measurement and control of the power grid, transmission of dispatch instructions, and other related services require a latency of less than 50ms; the latency threshold for distribution automation, new energy data collection, and other related services is at the 100ms level; However, non real time services such as electricity information collection and backend data analysis are less sensitive to latency. If the end-to-end delay of the power communication network exceeds the standard, it is highly likely to cause safety hazards such as delayed dispatch instructions, misoperation or refusal of relay protection, and untimely handling of power grid faults. In severe cases, it can lead to regional power grid fluctuations, power outages, and threaten the safe and stable operation of the new power system.

Deep Reinforcement Learning (DRL) has powerful dynamic decision-making and high-dimensional spatial optimization capabilities. It can achieve autonomous iterative optimization of strategies through continuous interaction between intelligent agents and dynamic environments, without the need for precise system analysis models. It perfectly adapts to complex optimization scenarios in IPv6 power communication networks that are dynamic, nonlinear, and multi coupled. Deep Deterministic Policy Gradient (DDPG), as a classic deep reinforcement learning algorithm, is based on the Actor Critic framework to achieve precise decision-making in continuous action spaces. Compared with traditional discrete reinforcement learning algorithms, it is more suitable for continuous variable optimization scenarios such as bandwidth allocation, routing selection, and power regulation in power communication. Based on this, this article combines the improved DDPG algorithm to conduct research on end-to-end delay minimization in IPv6 power communication networks. Through the algorithm's autonomous learning of IPv6 flow identifiers, routing states, and network dynamic characteristics, real-time dynamic scheduling of communication resources and global delay optimization are achieved. This has important theoretical value and engineering application significance for improving the transmission performance of power communication networks and ensuring the safe and stable operation of new power systems.

B. RESEARCH CONTENT AND TECHNICAL ROUTE

This article takes the smart grid IPv6 power communication network as the research object, with the global minimization of end-to-end delay as the core objective, combined with the deep deterministic policy gradient algorithm, to carry out systematic research on the delay mechanism analysis, network modeling, algorithm improvement, policy optimization, simulation verification, etc. of the power communication network. The core research content is as follows:

(1) Delay mechanism and demand analysis of power communication network. The system elaborates on the network architecture, transmission modes, and business classification of the power communication network, disassembles the core components of end-to-end delay in power communication, analyzes the impact mechanism of factors such as link load, resource allocation, routing paths, and business traffic on delay, sorts out the delay thresholds and optimization requirements of different power services, and constructs a mathematical model of delay in the power communication network.

(2) MDP modeling for dynamic optimization scenarios in power communication networks. Based on the characteristics of dynamic fluctuations in power communication network business and time-varying link loads, a Markov decision process is used to model the delay optimization problem, clarifying the network state space, resource scheduling action space, delay optimization reward function, and state transition rules. The end-to-end delay minimization problem is transformed into a continuous space dynamic decision optimization problem.

(3) Improvement and adaptation optimization of DDPG algorithm. In response to the shortcomings of the traditional DDPG algorithm, such as unstable training, low exploration efficiency, and susceptibility to local optima, combined with the characteristics of power communication optimization scenarios, the algorithm is improved from three dimensions: dual experience pool hierarchical sampling, adaptive learning rate updating, and differentiated reward and punishment function design, to enhance its dynamic adaptation ability, convergence speed, and global optimization ability.

(4) Construction of latency minimization strategy based on improved DDPG. Based on the improved DDPG algorithm, a joint optimization strategy for bandwidth resource allocation, routing path selection, and traffic scheduling in the power communication network is designed to minimize the global end-to-end latency of the network, while also considering the ultra-low latency requirements of high priority power services.

(5) Multi scenario simulation verification and performance analysis. Build a simulation experimental platform for power communication networks, and construct three typical power network scenarios: light load, full load, and sudden flow. Compare and test the improved algorithm in this paper with traditional optimization algorithms, and verify the superiority and effectiveness of the algorithm from dimensions such as average delay, delay jitter, convergence speed, resource utilization, and business support capability.

The overall technical route of this article follows the research logic of "mechanism analysis model construction algorithm improvement strategy design simulation verification summary analysis". Firstly, review the current research status of delay optimization in power communication networks, clarify the research difficulties and innovative points, and analyze the architecture, delay composition, and business requirements of power communication networks; Secondly, establish a mathematical model for power communication network delay and an MDP dynamic decision-making model; Once again, improve the shortcomings of the traditional DDPG algorithm and design optimization algorithms and delay minimization scheduling strategies that are suitable for power communication scenarios; Finally, through multi scenario simulation experiments, the algorithm performance is compared and analyzed, and the research results and shortcomings are summarized to complete the entire study.

II. DELAY MECHANISM AND THEORETICAL BASIS OF IPV6 POWER COMMUNICATION NETWORKS

A. OVERALL ARCHITECTURE OF IPV6 POWER COMMUNICATION NETWORK

The IPv6 power communication network is a dedicated communication network that serves the production, scheduling, operation, and marketing of the power system. Its address system follows hierarchical prefix allocation: regional and organizational prefixes identify administrative domains, /64 subnets identify individual links, and interface identifiers distinguish dispatch, substation, and terminal nodes. According to the network hierarchy, it can be divided into three layers: backbone communication network, aggregation communication network, and access communication network. The functional positioning, transmission rate, and business carrying type of each layer of network have significant differences, forming a fully covered, multi-level, and highly reliable power exclusive communication system, as shown in Table 1 below. At the IPv6 layer, the 8-bit Traffic Class field carries differentiated-service and congestion information, and the 20-bit Flow Label identifies packets belonging to the same flow. This study encodes an address as a domain prefix plus a node identifier rather than inputting a raw 128-bit value. Flow Label is treated as a categorical flow key, not as a priority value; service priority is determined by the mapping between business type and Traffic Class.

TABLE 1. Classification and delay threshold of power communication services

Priority Level	Typical Services	End-to-End Delay Threshold	Traffic Characteristics
Ultra-high priority	Relay protection, stability control device, emergency dispatching command	$\leq$ 10 ms	Bursty, ultra-high real-time requirement
High priority	Remote measurement & control, fault alarm, new energy grid-connected control	$\leq$ 50 ms	Periodic transmission, low delay jitter demand
Medium priority	Distribution automation, regional load monitoring	$\leq$ 100 ms	Stable flow, moderate delay tolerance
Low priority	Electricity data collection, equipment operation log, background analysis	$\leq$ 500 ms	Large data volume, low real-time demand

The power backbone communication network is the core transmission network between provincial and municipal power grid dispatch centers, mainly using fiber optic transmission. It has a large transmission bandwidth, high link stability, and complete topology redundancy. It mainly carries ultra-high priority services such as power grid core dispatch, relay protection, and stable control, and is the core skeleton of the power communication network. The number of backbone network nodes is small, the link span is large, and the network state is relatively stable. However, the requirements for latency and reliability for carrying business are the most stringent. Single link failures or delays exceeding the standard will directly affect the safety of regional power grid operation.

The power convergence communication network is a mid-level transmission network between regional substations and power supply stations, connecting the backbone network and access network. It adopts a fusion transmission mode of optical fiber and power wireless private network, mainly carrying high priority services such as distribution automation, regional load monitoring, and data aggregation of new energy stations. The network topology complexity is moderate, and the business traffic shows regional fluctuation characteristics.

The power access communication network is an access network for terminal equipment in the power grid, covering a large number of terminal nodes such as smart meters, distribution terminals, new energy inverters, and energy storage devices. It mainly uses power wireless private networks, power carriers, and short-range optical fibers as transmission methods to carry a large number of low priority services such as electricity information collection, terminal status monitoring, and equipment operation and maintenance data. It has the characteristics of a large number of nodes, dispersed topology, strong sudden business traffic, and severe dynamic fluctuations. It is the main source of delay fluctuations in the power communication network, as shown in Table 2 below.

TABLE 2. Parameter definition of MDP model for IPv6 power communication network

MDP Element	Specific Meaning	Value Range
State space	IPv6 prefix/flow ID, Traffic Class, Flow Label, link/load, route/next hop, queue, delay	Multi-dimensional continuous vector $[0, 1]$ normalized
Action space	Bandwidth share, IPv6 route weight, queue scheduling ratio	Continuous real number $[-0.3, 0.3]$
Instant reward	Comprehensive reward weighted by delay, penalty, load balance	Real number $[-100, 10]$
State transition	Post-action state determined by IPv6 traffic, link, and route changes	Stochastic distribution determined by traffic

From the perspective of transmission mode, the current power communication network has formed a heterogeneous transmission architecture of “optical fiber as the main, wireless as the auxiliary, and multi mode integration”. The transmission rate, delay characteristics, and anti-interference ability of different transmission modes are significantly different: optical fiber communication transmission has stable delay, low loss, and high reliability, which is suitable for core business transmission; The 5G wireless private network for power transmission is flexible and has low latency, suitable for dynamic access and sudden business of terminals; Power line carrier communication has wide coverage and low cost, but has large delay fluctuations and weak anti-interference ability, and is mostly used for non real time business transmission. The multi standard heterogeneous architecture leads to complex factors affecting the delay of power communication networks, and static optimization methods are difficult to achieve global optimality. IPv6 provides a unified network layer above these heterogeneous bearers, so the same address, service-class, and flow-label information can be used during cross-domain forwarding and resource scheduling.

B. MECHANISM OF END-TO-END DELAY IN IPV6 POWER COMMUNICATION NETWORKS

The end-to-end delay of an IPv6 power communication network refers to the total time required for business data to be generated, encapsulated, transmitted, forwarded from the source node to the destination node for parsing and receiving. It consists of two parts: fixed delay and time-varying delay, including four core components: propagation delay, transmission delay, queuing delay, and processing delay. IPv6 basic-header parsing, optional extension-header processing, Traffic Class identification, and Flow Label lookup are included in processing delay. Traffic Class affects queue priority, whereas Flow Label allows forwarding nodes to apply a consistent route and scheduling policy to packets of the same service flow. The generation mechanism and fluctuation characteristics of each delay component are significantly different. The specific analysis is as follows. For a payload of Lp bytes on a link with capacity C, transmission delay includes the fixed 40-byte IPv6 base header and is calculated from $8(\mathrm{Lp}+40)/\mathrm{C}$ when extension headers are disabled. Processing delay includes prefix lookup, Traffic Class classification, and Flow Label policy lookup, whereas route selection changes the number and type of links traversed.

Queuing delay refers to the time it takes for data packets to be forwarded in the node cache queue after arriving at a network node. It is the core source of latency fluctuations in power communication networks and belongs to highly dynamic time-varying delays. The queuing delay is influenced by multiple factors such as node traffic load, queue scheduling strategy, IPv6 Traffic Class, Flow Label, business priority, and sudden traffic, and there is no fixed analytical formula. When the traffic load of network nodes is low, the queue is idle and the queuing delay is approximately 0; When the network is in a fully loaded or sudden traffic scenario, queue congestion and queue delay increase sharply, which can easily lead to problems such as excessive delay and data loss. The sudden data transmission of massive terminals in the power access network is the main cause of queue delay fluctuations, and it is also the delay component that this article focuses on optimizing, as shown in Table 3.

TABLE 3. IPv6 simulation configuration and algorithm hyperparameters

Algorithm	Learning Rate	Batch Size	Discount Factor γ
Improved DDPG	Adaptive 0.0001–0.001	256	0.95
Original DDPG	Fixed 0.001	256	0.95
Q-learning	Fixed 0.002	—	0.95
GA	Crossover=0.7, Mutation=0.01	Population=60	—
PSO	Inertia weight=0.6	Particle=50	—
IPv6-only; /64/link; 40-B base header; no extension header; DSCP=EF/AF41/AF21/BE; 20-bit Flow Label; UDP; HL=64			

C. CLASSIFICATION AND LATENCY REQUIREMENTS FOR POWER COMMUNICATION SERVICES

The power communication business has clear hierarchical characteristics, and the real-time and reliability requirements of different businesses vary greatly. In the IPv6 environment, ultra-high-, high-, medium-, and low-priority services are mapped to differentiated Traffic Class values, while a stable Flow Label identifies packets belonging to the same service flow. In order to achieve precise delay optimization and ensure the priority transmission of core businesses, this article divides the power communication business into four categories based on the national power communication industry standards, combined with business functions and delay thresholds: ultra-high priority, high priority, medium priority, and low priority. The specific business characteristics and delay requirements are shown in the table below.

The hierarchical characteristics of power services require delay optimization algorithms to have differentiated scheduling capabilities. On the basis of minimizing global delay, priority should be given to ensuring low delay transmission of ultra-high and high priority services, avoiding the problem of delay exceeding the standard in core power services due to network optimization. This is also an important constraint for algorithm improvement and strategy design in this article.

D. PRINCIPLE OF DEPTH DETERMINISTIC STRATEGY GRADIENT ALGORITHM

Deep Deterministic Policy Gradient (DDPG) is a deep reinforcement learning algorithm based on the Actor Critic framework, which combines the advantages of policy gradient algorithm and value function algorithm. It can efficiently handle optimization decision problems in high-dimensional continuous action spaces and perfectly adapt to continuous variable optimization scenarios such as bandwidth allocation, IPv6 flow routing weight adjustment, and queue scheduling in power communication networks. The DDPG algorithm consists of four parts: Actor network, Critic network, Target network, and Experience replay pool. The core principles are as follows.

Actor network is a policy network responsible for outputting deterministic scheduling actions and making resource scheduling decisions based on the current environmental state. The state vector contains IPv6 source and destination identifiers, Traffic Class, Flow Label, link load, residual bandwidth, current route or next-hop information, queue length, and real-time delay. Unlike traditional stochastic gradient algorithms, DDPG outputs deterministic actions with higher decision-making efficiency, making it more suitable for power communication scheduling scenarios with high real-time requirements. Source and destination addresses are encoded by prefix domain and node identity, and the arbitrary numerical value of a Flow Label is encoded categorically to avoid a false magnitude relationship. The action vector contains bandwidth-share adjustment, weights of reachable IPv6 routes, and queue service ratios. The reward combines weighted end-to-end delay with penalties for service-threshold violations, packet loss, frequent route switching, and unfair resource occupation. State transition follows the next interval of traffic arrivals, link measurements, queue evolution, and route updates; candidate paths are filtered by prefix reachability, loop-free forwarding, and Hop Limit constraints before execution.

III. IMPROVED DDPG ALGORITHM DESIGN FOR IPV6 POWER COMMUNICATION NETWORKS

A. ANALYSIS OF DEFECTS IN TRADITIONAL DDPG ALGORITHM

Although the traditional DDPG algorithm can achieve continuous spatial decision-making, its direct application to delay optimization scenarios in IPv6 power communication networks has obvious shortcomings and cannot adapt to the optimization requirements of high dynamic, high real-time, and differentiated constraints in power networks. The specific shortcomings are as follows:

(1) Low efficiency of experience sampling: Traditional DDPG adopts a random uniform sampling strategy to extract experience pool samples, without distinguishing sample values. High quality low latency scheduling samples and invalid congestion samples are equally extracted in power communication networks, resulting in low utilization of high-quality samples, poor algorithm learning efficiency, and slow convergence speed.

(2) Lack of training stability: Traditional algorithms use a fixed learning rate to update network parameters. In the early stages of algorithm training, a large learning rate is required for rapid iteration, and in the later stages, a small learning rate is needed for precise convergence. A fixed learning rate cannot adapt to changes in the training stage, and is prone to parameter oscillation and unstable convergence, making it difficult to adapt to dynamic scenarios in power networks.

(3) Poor balance ability in exploration and utilization: Traditional algorithms have fixed exploration strategies, and insufficient exploration in the early stages of training can easily lead to local optima. Excessive exploration in the later stages of training can cause policy fluctuations, making it difficult to balance global exploration and local optimal utilization, and achieve global delay minimization in power communication networks.

(4) Non differentiated business adaptation: Traditional reward functions do not consider the priority differences of power services, and do not differentiate the delay constraints between core and ordinary services, which cannot guarantee the ultra-low delay requirements of ultra-high priority power services and is not in line with the actual situation of power communication engineering.

In response to the above deficiencies, this article systematically improves the DDPG algorithm from four dimensions: experience sampling mechanism, learning rate update strategy, reward and punishment function optimization, and exploration strategy adjustment, to enhance the algorithm's adaptability to power communication delay optimization scenarios, as shown in Table 4 below.

TABLE 4. Weighted average end-to-end delay under three simulation scenarios (unit: ms)

Algorithm	Light Load Scenario (30% load)	Full Load Scenario (75% load)	Burst Traffic Scenario
Q-learning	18.62	42.35	58.71
GA	16.35	38.62	52.36
PSO	15.89	36.18	49.85
Original DDPG	13.26	31.52	43.62
Improved DDPG	10.45	24.86	34.28

B. IMPROVED MECHANISM BASED ON DUAL EXPERIENCE POOL HIERARCHICAL SAMPLING

To solve the problems of low utilization rate of traditional random sampling samples and insufficient mining of high-quality samples, this paper designs a dual experience pool hierarchical sampling mechanism, which divides the experience pool into high-quality sample pool and ordinary sample pool, realizes sample hierarchical storage, differentiated sampling, and improves algorithm learning efficiency.

IV. IPV6 POWER COMMUNICATION NETWORK DELAY MINIMIZATION SCHEDULING STRATEGY BASED ON IMPROVED DDPG

A. MULTI RESOURCE JOINT OPTIMIZATION SCHEDULING FRAMEWORK

Traditional power communication delay optimization often adopts a single resource scheduling mode, which only optimizes routing or bandwidth resources and cannot achieve global delay optimization. This article is based on an improved DDPG algorithm to construct a three in one multi resource joint optimization and scheduling framework of "bandwidth allocation routing optimization traffic scheduling", breaking the limitations of single resource optimization and achieving multi-dimensional resource collaborative regulation of the IPv6 power communication network, minimizing end-to-end latency. This framework is based on real-time network state perception, with improved DDPG intelligent decision-making as the core and business differentiation guarantee as the constraint, to achieve real-time adaptive scheduling in dynamic scenarios. The overall framework is divided into three layers:

(1) Network state perception layer: Real-time collection of IPv6 source and destination identifiers, Traffic Class, Flow Label, link load, residual bandwidth, current route or

next-hop status, queue status, and latency data; the data are normalized to construct an IPv6 flow-oriented network state vector for intelligent decision-making;

(2) Intelligent decision optimization layer: Based on the improved DDPG algorithm, continuous actions are generated for bandwidth adjustment, IPv6 flow routing weight, and queue scheduling ratio. Flow Label binds each action to a service flow, while Traffic Class and business priority constrain differentiated delay assurance;

(3) Resource scheduling execution layer: The execution layer adjusts link bandwidth, updates IPv6 business-flow routes according to Flow Label, and schedules queues according to Traffic Class and business priority. The resulting network state is returned to the decision layer to form a closed-loop optimization mechanism, as shown in Table 5 below.

TABLE 5. Performance comparison indicators of different optimization algorithms

Algorithm	Delay Jitter (ms)	Bandwidth Utilization	Ultra-high Priority Service Compliance Rate
Q-learning	12.1	67.2%	85.3%
GA	10.8	71.5%	87.6%
PSO	10.2	73.8%	89.1%
Original DDPG	8.6	79.4%	95.8%
Improved DDPG	5.5	89.7%	100.0%

B. DYNAMIC BANDWIDTH ALLOCATION OPTIMIZATION STRATEGY

Bandwidth resources are the core variables that affect the transmission delay of power communication. In response to the problems of resource waste and congestion delay caused by traditional static bandwidth allocation, this paper designs a dynamic bandwidth allocation strategy based on an improved DDPG algorithm. According to link load, IPv6 Traffic Class, Flow Label, and business priority, the bandwidth resources of each link are adaptively adjusted. Flow Label is used only to associate packets with a learned allocation. The controller first preserves the minimum share of each Traffic Class and then distributes residual capacity according to queue pressure and delay margin, while the sum of flow allocations remains bounded by available link bandwidth.

In the scenario of light network load, the overall network traffic is small and the degree of link congestion is low. The algorithm allocates bandwidth resources moderately and evenly to ensure stable transmission of various services, and reserves redundant resources to cope with sudden traffic; In scenarios where the network is fully loaded, priority should be given to allocating sufficient bandwidth for ultra-high and high priority services, reducing transmission latency for core services, moderately compressing bandwidth for low priority services, and achieving precise resource allocation; In sudden traffic scenarios, algorithms quickly sense sudden changes in traffic, instantly adjust the bandwidth of congested links, alleviate queue congestion, and suppress sudden increases in latency.

This bandwidth allocation strategy abandons the traditional mode of fixed allocation and average allocation. By improving the continuous decision-making ability of the DDPG algorithm, it achieves fine-grained, dynamic, and differentiated allocation of bandwidth resources. While ensuring ultra-low latency for core services, it maximizes the utilization of bandwidth resources across the entire network and optimizes latency from the perspective of transmission delay dimension.

C. LOAD BALANCING ROUTING OPTIMIZATION STRATEGY

The multi-path redundant architecture of the power communication network provides a foundation for routing optimization. Traditional shortest path routing strategies can easily lead to long-term congestion of the optimal path, uneven link load, and overall increase in network latency. This article is based on the improved DDPG algorithm to design a load balancing routing optimization strategy. It no longer pursues the shortest path alone, but aims to minimize the network latency by dynamically adjusting the transmission routing weights of various services to balance the load of all network links.

The algorithm perceives the load rate, latency, and bandwidth redundancy status of each link in real-time. For high load and high latency links, it automatically reduces their routing weights and guides business traffic to transfer to low load and low latency redundant links; For low load idle links, moderately increase routing weights and fully utilize redundant resources. At the same time, for ultra-high priority services, priority should be given to selecting the optimal path with the lowest latency and strongest stability to eliminate core service routing congestion. This strategy can effectively solve the problems of local link congestion and global resource idle in the power communication network, balance the load distribution of the entire network, and significantly reduce queuing delay and overall end-to-end delay. The destination prefix defines the reachable route set, and the DDPG action adjusts weights among those candidates. A stable Flow Label keeps packets of one flow on a consistent treatment path, while a route-switching penalty prevents unnecessary oscillation. Loop-free and Hop Limit checks are applied before a new next hop is installed.

D. DIFFERENTIATED TRAFFIC SCHEDULING STRATEGY

This paper designs a queue traffic scheduling strategy based on improved DDPG to address the differentiated latency requirements of power business layers. IPv6 Traffic Class maps packets to service-priority queues, and Flow Label distinguishes concurrent flows within the same service class. The strategy optimizes the business scheduling sequence of node cache queues and achieves precise latency optimization from the perspective of queuing latency. Traditional queue scheduling often adopts a first come, first served model, which cannot distinguish business priorities. Core businesses are easily blocked by a large number of low priority businesses, resulting in excessive latency.

This article improves the strategy by dynamically adjusting the queue scheduling weights of IPv6 flows identified by Flow Label and grouped by Traffic Class through algorithms. When network resources are tight and queues are congested, the scheduling priority of ultra-high and high priority services is significantly increased, and core control services and scheduling service data are prioritized for forwarding, minimizing their latency to the greatest extent possible; Adopt delayed scheduling and batch forwarding mode for medium and low priority services to ensure normal transmission of ordinary services without affecting core services. At the same time, through the differentiated penalty mechanism of the reward function, the constraint algorithm always prioritizes ensuring that the core business latency meets the standard, achieving the dual goals of optimal network latency and high reliability transmission of core business. At each forwarding node, Traffic Class selects the service queue and Flow Label distinguishes concurrent flows within that queue. Queue weights are bounded so that urgent protection traffic receives strict delay assurance without allowing persistent high-priority traffic to starve lower service classes.

E. MULTI SCENARIO ADAPTIVE SCHEDULING MECHANISM

The power communication network has three typical operating scenarios: light load, full load, and sudden flow. The focus of delay optimization varies in different scenarios. This article designs a multi scenario adaptive scheduling mechanism based on an improved DDPG algorithm to achieve precise optimization in different scenarios.

(1) Light load scenario: The entire network has sufficient resources and overall low latency. The algorithm aims to improve resource utilization and balance load, evenly allocate bandwidth and traffic, and maintain low latency and stable operation of the entire network;

(2) Full load scenario: With tight network resources and high risk of link congestion, the algorithm aims to ensure core business latency and minimize network latency, accurately tilt resources, optimize routing, and implement differentiated scheduling to suppress the increase in latency;

(3) Sudden traffic scenario: Business traffic surges instantly, queue congestion is severe, algorithms quickly respond to sudden changes in traffic, adjust bandwidth and routing instantly, quickly alleviate sudden traffic, reduce latency jitter, and avoid sudden increases in latency and data loss.

V. SIMULATION EXPERIMENT AND RESULT ANALYSIS FOR IPV6 POWER COMMUNICATION NETWORKS

A. SIMULATION EXPERIMENT PLATFORM AND PARAMETER SETTINGS

To verify the effectiveness and superiority of the improved DDPG algorithm and delay minimization scheduling strategy in this article, a joint simulation platform of Python and MATLAB was built. The network layer was configured in IPv6-only mode without IPv4 compatibility or address translation. Based on MATLAB, a topology and business flow model of the IPv6 power communication network was constructed, and the training optimization of the improved DDPG algorithm and the comparative algorithm was implemented using Python to complete multi scenario simulation testing.

A regional IPv6 power communication network topology was constructed with 2 dispatch-center nodes, 8 substation aggregation nodes, 20 terminal access nodes, 30 network nodes in total, and 42 heterogeneous communication links covering optical fiber, power 5G, and powerline carrier transmission. Each point-to-point link used an independent /64 prefix allocated from the simulation prefix 2001:db8:100::/48. Service data were encapsulated in IPv6 packets with a 40-byte base header, Next Header set to UDP (17), initial Hop Limit set to 64, and ECN set to 00. Traffic Class DSCP values for the four service levels were mapped to EF, AF41, AF21, and BE. Each service flow was assigned a nonzero 20-bit Flow Label that remained unchanged during the flow lifetime, and packet arrivals followed a Poisson distribution. All comparison algorithms use the same IPv6 header overhead, Traffic Class mapping, and Flow Label lifetime. Extension headers and IPv4 translation are disabled. At each decision interval, the environment reports link utilization, residual bandwidth, queue length, measured delay, current next hop, and route availability, and each action is projected onto bandwidth-capacity and loop-free routing constraints before state update.

To demonstrate the superiority of the algorithm, four mainstream optimization algorithms were selected as the comparative control group: traditional Q-learning algorithm, original DDPG algorithm, genetic algorithm (GA), and particle swarm optimization algorithm (PSO). Comparative tests were conducted under identical network topology, business scenarios, and constraint conditions.

B. SIMULATION SCENE SETTINGS

To comprehensively verify the scenario adaptation capability of the algorithm, three typical power communication operation scenarios were set up for simulation experiments, covering both conventional and extreme operating states of the power grid:

Scenario 1: Light load scenario. The overall business traffic load rate is 30%, with uniform business distribution and no sudden traffic. It simulates the low load operation state of the power grid at night and mainly tests the latency optimization and resource utilization performance of the algorithm under normal conditions.

Scenario 2: Full load scenario. The overall network traffic load rate is 75%, and the traffic distribution of each link is uneven. Simulating the daytime peak operation state of the power grid, the main test algorithm is to optimize latency and load balancing ability under high load.

Scenario 3: Sudden Traffic Scenario. The basic load rate is 60%, and a large number of new energy grid connected and fault alarm emergency services are randomly connected. The extreme scenarios of power grid fault disturbance and new energy fluctuation are simulated, and the dynamic response and delay jitter suppression ability of the algorithm are mainly tested.

C. DEFINITION OF EVALUATION INDICATORS

To quantitatively compare the optimization performance of various algorithms, five core evaluation indicators are set to comprehensively evaluate the delay optimization effect, algorithm stability, and resource utilization ability:

(1) Weighted average end-to-end latency across the entire network: a core optimization indicator, the smaller the value, the better the overall network latency optimization effect;

(2) Delay jitter amplitude: reflects delay stability, the smaller the value, the smaller the network delay fluctuation and the more stable the transmission;

(3) High priority business latency compliance rate: reflects the core power business guarantee capability, the higher the compliance rate, the higher the safety of power grid operation;

(4) Network resource utilization rate: reflects the efficiency of bandwidth resource utilization, with higher values indicating less resource waste;

(5) Algorithm convergence iteration times: Reflecting the optimization efficiency of the algorithm, the fewer iterations, the faster the convergence speed.

D. SIMULATION RESULTS AND COMPARATIVE ANALYSIS

The weighted average delay optimization results of each algorithm in the three scenarios are shown in the table below. Through data comparison, it can be seen that the improved DDPG algorithm in this paper has the lowest average delay in all scenarios, and the optimization effect is significantly better than other compared algorithms.

Data analysis shows that compared to the original DDPG algorithm, the improved algorithm in this paper reduces the average latency by 21.2%, 21.1%, and 21.6% in light load, full load, and burst scenarios, respectively, with an average reduction of 21.3%; Compared to traditional heuristic algorithms and Q-learning algorithms, the advantage of delay optimization is more obvious. In light load scenarios, the difference in latency between algorithms is small, and the lower the network load, the smaller the optimization space; Under full load and sudden scenarios, traditional algorithms have significantly increased latency, while the algorithm proposed in this paper can still maintain low latency and has strong dynamic adaptation capabilities, which can effectively cope with network congestion and sudden traffic changes.

Delay jitter is a core indicator for measuring the stability of power communication transmission, which directly affects the accuracy of power grid regulation. The simulation results show that the delay jitter amplitude of the improved DDPG algorithm in this paper is significantly lower than that of traditional algorithms, and the delay stability of the entire network is greatly improved. In the scenario of sudden traffic, the traditional DDPG algorithm has a delay jitter amplitude of 8.6ms, while the Q-learning, GA, and PSO algorithms all have jitter amplitudes exceeding 10ms. However, the improved algorithm in this paper has a delay jitter amplitude of only 5.5ms, which is 35.7% lower than the original DDPG. The core reason is that the algorithm in this article improves training stability through adaptive learning rate, and effectively suppresses delay oscillations caused by traffic fluctuations by combining with load balancing scheduling strategy, which can provide stable low delay transmission guarantee for power business.

The delay compliance rate of high priority power services is the core indicator for evaluating the practicality of algorithm engineering, and the performance of core service guarantee varies significantly among different algorithms. In the most severe sudden traffic scenarios, the improved algorithm in this article achieves a 100% compliance rate for ultra-high priority business latency and a 99.2% compliance rate for high priority business latency, fully guaranteeing the ultra-low latency requirements of core businesses such as relay protection and scheduling instructions; The original DDPG algorithm has a compliance rate of 95.8% for ultra-high priority services, while traditional heuristic algorithms have compliance rates below 90%, indicating a risk of exceeding core service latency. The effectiveness of the differentiated reward and punishment function and traffic scheduling strategy in this article has been verified, which can achieve priority guarantee for core business and meet the actual needs of power communication engineering.

VI. CONCLUSION

This article proposes an improved DDPG end-to-end delay minimization scheduling method to address the issue of time delay exceeding caused by dynamic traffic in heterogeneous IPv6 power communication networks of new power systems. Build a refined delay model based on hierarchical delay constraints in power business, and complete scheduling decision modeling based on MDP. IPv6 source and destination identifiers, Traffic Class, Flow Label, link state, and routing information are incorporated into the state vector; To address the shortcomings of inefficient sampling and training oscillations in traditional DDPG, a dual experience pool hierarchical sampling, adaptive learning rate, and differentiated reward and punishment function are introduced to construct a bandwidth routing traffic multi resource joint scheduling framework. The multi scenario simulation results show that compared to Q-learning PSO 'The original DDPG and other algorithms have reduced the weighted average latency of the entire network by 21.3%, reduced latency jitter by 35.7%, maintained a 100% compliance rate for ultra-high priority business latency, significantly improved convergence speed, and can adapt to light load, full load, and sudden traffic conditions in IPv6 power communication networks. Because the method relies on standard IPv6 addressing and header fields, it is applicable to IPv6-based dispatch, substation aggregation, and terminal access networks. The study only considers normal link operation scenarios, and can integrate link fault perception mechanisms in the future to further enhance the robustness of algorithm engineering. In deployment, ingress gateways must maintain the mapping among power-service type, Traffic Class, and Flow Label, while forwarding nodes provide link, queue, and route telemetry to the controller. Future work should verify extension-header processing, multi-domain prefix policies, link failures, and controller communication delay in operational IPv6 power networks.

ACKNOWLEDGMENT

This work was supported by the Science and Technology Project of Southern Power Grid Corporation (No.ZDKJXM20230027)

REFERENCES

[1] L. Bin, L. Chao, Q. Bing, et al., "Enhancement of failure independent path protection preconfigured cycles for electric power communication networks," *Opt. Switch. Netw.*, vol. 48, 2023, doi: 10.1016/J.OSN.2023.100731.

[2] F. Bing, T. Hongtao, and L. Yaqun, "Critical link identification algorithm for power communication networks in SDN architecture," *Int. J. Crit. Infrastruct. Prot.*, vol. 40, 2023, doi: 10.1016/J.IJCIP.2022.100584.

[3] L. Wenjuan, L. Yijie, L. Bo, et al., "Low-Latency Autonomous Service Orchestration for the Power Communication Network," *J. Phys.: Conf. Ser.*, vol. 2399, no. 1, 2022, doi: 10.1088/1742-6596/2399/1/012048.

[4] M. Congcong, "ODL Centralized Control of Power Communication Network Based on Bio-Inspired Algorithms and SDN," *Math. Probl. Eng.*, vol. 2022, 2022, doi: 10.1155/2022/9925620.

[5] L. Bin, L. Chao, Q. Bing, et al., "Risk and traffic based service routing optimization for electric power communication network," *Int. J. Electr. Power Energy Syst.*, vol. 137, 2022, doi: 10.1016/J.IJEPES.2021.107782.

[6] Y. Zhu, A. Tian, Z. Lu, et al., "Research on operation and maintenance management risk assessment of power communication network based on DEMATEL," *Int. J. Ind. Syst. Eng.*, vol. 40, no. 3, pp. 325–342, 2022, doi: 10.1504/IJISE.2022.122247.

[7] L. Ma, J. Zhou, B. Xu, et al., "State scheduling method of redundant nodes in power communication network based on least square method," *Int. J. Auton. Adapt. Commun. Syst.*, vol. 15, no. 3, pp. 185–201, 2022, doi: 10.1504/IJAACS.2022.10034570.

[8] F. Xinxin, C. Xiuguo, Y. Hongmei, et al., "Design of an 8X8 Optical Matrix Protection Device based on Electric Power Communication Network," *J. Phys.: Conf. Ser.*, vol. 2083, no. 2, 2021, doi: 10.1088/1742-6596/2083/2/022062.

[9] Huang, Hao, Ma, et al., "A novel network flow prediction method based on UWB in power communication networks," *Wireless Netw.*, pp. 1–11, 2021, doi: 10.1007/S11276-021-02755-3.

[10] S. Li, Y. Xinyu, Z. Zhiqin, et al., "Reliability Analysis of Optical Path Automatic Protection Switching in Ubiquitous Power Communication Network," *J. Phys.: Conf. Ser.*, vol. 1992, no. 3, 2021, doi: 10.1088/1742-6596/1992/3/032128.

[11] Z. Yan, X. Jie, Y. Hui, et al., "Research on Informatization Construction of Electric Power Communication Network under Smart Grid," *IOP Conf. Ser.: Earth Environ. Sci.*, vol. 632, no. 4, Art. no. 042090, 2021, doi: 10.1088/1755-1315/632/4/042090.

[12] L. Lu, H. Xue, G. Lizhuo, et al., "Design and application of electric power communication network comprehensive simulation training system," *MATEC Web Conf.*, vol. 336, Art. no. 05019, 2021, doi: 10.1051/MATECCONF/202133605019.

[13] J. Zhou, Y. Jiang, B. Xu, et al., "Research on protection of vulnerable nodes in power communication network based on the model of invalid proximity of edge," *Int. J. Auton. Adapt. Commun. Syst.*, vol. 14, no. 4, pp. 354–370, 2021, doi: 10.1504/IJAACS.2021.119126.

[14] W. Miao, M. Weiwei, W. Haiyang, et al., "Intelligent Auxiliary Operation and Maintenance System of Power Communication Network based on Knowledge Graph," *J. Phys.: Conf. Ser.*, vol. 1684, no. 1, Art. no. 012105, 2020, doi: 10.1088/1742-6596/1684/1/012105.

[15] H. Liu, Y. Ma, X. Chen, et al., "Reliability Evaluation and Optimization Method in Power Communication Network Based on Environmental Factors and Controlled Hybrid Stochastic Petri Net," *한국통신학회APNOMS*, 2020.

[16] H. Liu, Y. Ma, X. Chen, et al., "Reliability Evaluation and Optimization Method in Power Communication Network Based on Environmental Factors and Controlled Hybrid Stochastic Petri Net," *IEICE Proc. Ser.*, vol. 62, pp. 349–352, 2020, doi: 10.34385/PROC.62.P3-8.

[17] "Algorithms; Findings in Algorithms Reported from H.Z. Yang and Co-Researchers (Traffic Grooming Algorithm for Power Communication Optical Networks)," *J. Math.*, p. 305, 2018.

[18] Sony Corporation, "Patent Issued for Multiple-Input and Multiple-Output Modem for Connection in a Power Line Communication Network (USPTO 9866275)," *Telecommun. Weekly*, p. 2667, 2018.

[19] Q. Han, Q. Feng, Y. Su, et al., "Services Paths Planning for Electric Power Communication Network Based on Improved Ant Colony Optimization," *IEICE Proc. Ser.*, vol. 17, Art. no. P2-3, 2013, doi: 10.34385/PROC.17.P2-3.