EAI Endorsed Transactions
on Energy Web
Research Article EAI.EU

Dynamic Path Selection in SDN Based on Reinforcement Learning and Link Utilization

Ying Zeng[1], Xingnan Li[1], Yuben Bao[1], Guangyu Hu[2], Jikai Hong[2] and Fan Zha[2]

[1] Guangdong Power Grid Power Dispatching Control Center, Guangzhou, China
[2] NARI Information & Communication Technology Co., Ltd., Nanjing, China

Abstract

INTRODUCTION: The development of modern power systems imposes stringent requirements on communication networks, including highly dynamic loads, low latency, and high reliability. Although recent Software-Defined Networking routing, link-utilization-aware scheduling, and reinforcement learning-based methods improve path optimization, challenges remain in bottleneck-link perception, congestion feedback, and stable decision-making under dynamic traffic conditions.
OBJECTIVES: To counteract problems like delayed response times and inadequate congestion identification in conventional path selection algorithms owing to dynamic link modifications, this research paper presents a path selection model that combines bottleneck link usage and reinforcement learning.
METHODS: Under the Software-Defined Networking control architecture, the proposed model incorporates link utilization, Graph Convolutional Network structures, and Gated Recurrent Units, while introducing a deep reinforcement learning algorithm to optimize routing strategies.
RESULTS: Experimental results demonstrate that under a 60 Mbit/s load, the proposed method achieves a throughput of 57 Mbit/s, maintains the minimum transmission delay within 0.043 s, and yields a link utilization rate of 89%. In experiments carried out to evaluate dynamic decision-making, the approach for choosing paths adopted by the model averages convergence at the 15th round, resulting in an error rate of 4.2%, minimal load balancing at 0.23, and median latency time in real-time decisions of only 31 ms, which is better than other models.
CONCLUSION: These results demonstrate that the model achieves superior state awareness and adaptive routing performance in multi-source heterogeneous networks. The model also exhibits good performance in terms of congestion control and path optimization and hence can be used effectively for intelligent routing in next-generation power communication networks.

Keywords: Reinforcement learning; Link utilization; Path selection optimization; Software-defined networking; Graph convolutional network

Received on 12 May 2026, accepted on 06 July 2026, published on 10 August 2026

doi: 10.4108/ew.12999

*Corresponding author. Email: newzengying_zyxing@163.com

1. Introduction

Due to fast development of modern power systems and changes in energy structure, there emerge many difficulties associated with power communication network including highly variable loads, low delay transmission, and high requirements to reliability [1]. Traditional routing path selection relies on static topologies and fixed policies, which fail to perceive real-time link loads, thereby leading to network congestion and degraded scheduling efficiency [2]. In some studies, Software Defined Networking (SDN) was implemented in the power communication system because of its advantages including centralization and overall network visibility for network scheduling purposes. However, current SDN routing schemes rely largely on static information and do not have any real-time network knowledge [3-4]. Reinforcement Learning (RL) has unique advantages in adaptive strategy optimization through interaction with the environment, enabling dynamic path adjustment based on changes in network states [5]. Incorporating information such as Link Utilization (LU) into reinforcement learning models helps achieve accurate load balancing and resource optimization, thereby avoiding link overload [6]. To achieve this, this study leverages cutting-edge technologies, including SDN-based centralized control, GCN-based topology representation, GRU-based temporal traffic modeling, and DQN-based adaptive decision-making, to jointly support intelligent path optimization under dynamic network conditions. Therefore, to address the complex and variable scenarios in power communication networks, this study proposes an SDN dynamic path selection model that integrates LU information with a Graph Convolutional Network (GCN), a Gated Recurrent Unit (GRU), and Deep Q-network (DQN) reinforcement learning. The proposed model gathers the topology information and LU information via the SDN controller, which serves as the state input for the strategy optimization based on reinforcement learning. Its goal is to enhance the real-time performance, load balancing, and robustness in path selection, ensuring the operational security of the power communication system. Its key novelty includes incorporating the scheduling strategy that integrates bottleneck LU and Explicit Congestion Notification (ECN), allowing for the coordination of congestion recognition and traffic balancing. Moreover, an intelligent routing framework integrating GCN-GRU and DQN is developed, which helps advance the learning and adaptability towards network dynamics, facilitating the intelligent development of power communication network. The main contributions of this study are summarized as follows: (1) a BLU-dual ECN collaborative scheduling mechanism is proposed to enhance congestion awareness and adaptive traffic balancing; (2) a graph-based spatiotemporal feature extraction framework combining GCN and GRU is developed to model topology-dependent traffic evolution; and (3) a DQN-driven intelligent routing strategy is designed to achieve adaptive path optimization through continuous interaction with dynamic network environments. Unlike existing GAT-DQN, GRU-PG, and GCN-based routing methods that mainly focus on the individual aspect of topology learning, temporal prediction, or policy optimization, respectively, LUGDQ-SDN integrates bottleneck link congestion notification, spatiotemporal state modeling, and RL-based routing in one closed-loop decision-making process. What makes the proposed algorithm novel is the fact that BLU-dual ECN, GCN-GRU, and DQN are used as the congestion aware state notifier, topology-temporal state encoder, and adaptive policy optimizer. In practical power communication systems, LUGDQ-SDN can support delay-sensitive services such as wide-area protection, dispatching automation, and distributed energy coordination. Compared with conventional SDN routing, it provides a unified congestion-aware and adaptive decision framework, improving routing reliability under dynamic traffic and large-scale network conditions while remaining compatible with existing SDN controllers.

2. Related works

RL keeps optimizing its action selection strategy by means of a reward feedback process, which is particularly ideal for networks with unpredictable states and various objectives. Extensive research has been conducted globally on this topic. For example, He et al. proposed the idea of a deep reinforcement learning algorithm with the incorporation of the GCN model to solve the challenges faced in terms of information exploitation and load balancing in the case of conventional algorithms. This method dynamically extracted transferable structural knowledge to achieve fine-grained routing decisions [7]. Li et al. proposed a path optimization model based on reinforcement learning to improve the traditional emergency evacuation routing system, which often failed to respond promptly to changes such as road closures caused by floods or other disasters. In this model, the agent learned the optimal path strategy through continuous interaction and the reward mechanism [8]. As an important indicator reflecting the network load status, LU has also been widely studied. He et al. proposed a reconfigurable architecture combining link rate adaptation and network function virtualization, along with a method for constructing service function chains. This framework was designed to resolve the coordination challenges of virtual network function deployment and wireless resource allocation in integrated space-terrestrial networks, which are caused by resource heterogeneity and service diversity [9]. Yang et al. proposed a methodological framework for analyzing and assessing the effectiveness of transport protocols based on the effect of link interruption on the efficiency of transmission. The authors also introduced an interval threshold based on modeling which is used to determine if multiple link interruptions can be modeled as one [10].

Path selection algorithms for SDN are fairly well developed both theoretically and practically, and many researchers from different countries have studied this topic in their research papers. For example, Taha offered an effective intelligent algorithm for selecting routes to avoid routing congestion and packet delay in SDN networks. Experiments proved that the algorithm could cut down routing delay by up to 96.3% [11]. In addition, Ritzkal et al. designed a system for evaluating and recommending paths in SDN based on k-nearest neighbors algorithm to increase the level of intelligence in path selection. It helped to analyze potential routes using performance criteria like throughput and packet loss rate [12]. Zhen et al. provided a dynamic load-balancing technique in SDN for large-scale data centers to cope with problems associated with unbalanced load and congestion control. The algorithm utilized an advanced version of the ant colony optimization algorithm with the use of chaotic disturbance in evaluating the routes depending on link and server utilization in order to find the optimum path

[13]. Wu et al. proposed an intelligent routing technique for video conferencing in SDN networks in order to handle problems concerning delays and bandwidth guarantee. The algorithm minimized request conflicts by using weighted key links and included a delay component that dynamically changed the priority of requests [14].

In summary, although existing studies have made progress in reinforcement learning, link modeling, and SDN path selection, most existing methods still have several limitations. First, some reinforcement learning-based routing methods mainly depend on reward feedback and lack explicit perception of bottleneck-link congestion. Second, existing link-aware studies often focus on isolated indicators such as bandwidth, delay, or utilization, failing to capture the joint effects of network topology and temporal traffic evolution. Third, many SDN path selection methods rely on static rules, heuristic optimization, or local path evaluation, making it difficult to maintain stable routing decisions under rapid traffic fluctuations. Therefore, this study proposes an SDN path selection method that integrates LU information, GCN, GRU, and DQN, aiming to construct a coordinated control mechanism that balances congestion awareness and resource optimization. This method is designed to improve the intelligence of scheduling and the robustness of the network in complex scenarios, providing both theoretical and technical support for the intelligent evolution of new-type power communication networks.

3. Dynamic path selection method integrating reinforcement learning and link utilization

3.1. SDN path selection algorithm based on bottleneck link utilization

With increasing network complexity and traffic dynamism, traditional SDN path selection algorithms often fail to adjust routing in time under traffic bursts or drastic link changes, resulting in frequent congestion and rising transmission delay [15–16]. To address the issues of static evaluation delay and poor congestion reaction, an adaptive path selection scheme known as BLUECN, which combines Bottleneck Link Utilization (BLU) and dual ECN, is introduced in this research work. While transmitting data packets, the proposed algorithm analyzes link states dynamically, recognizes the bottleneck links, and carries out ECN marking for identifying possible congestion situations.

The marked packets are forwarded using minimum queuing. The minimum queue mechanism is mathematically formulated as shown in Equation (1).

$$\begin{cases} q_{\min} = k + n - \dfrac{1}{2}\sqrt{2n\left(c \times rtt + k\right)} \\ k > \dfrac{c \times rtt}{7} \end{cases} \tag{1}$$

In Equation (1), k and n represent the ECN threshold and the number of synchronized flows, while c and rtt refer to the link capacity and round-trip delay. Then, BLU λ is introduced to calculate the ECN threshold. The calculation of λ is shown in Equation (2).

$$\lambda(t) = \frac{\sum_{i=1}^{n} W_i(t) - Q(t)}{C \times RTT_b} \tag{2}$$

In Equation (2), $W_i(t)$ indicates the traffic of the i-th data flow at time t, and $Q(t)$ is the queue length in the network. C and RTT_b represent the bottleneck link capacity and the basic round-trip delay, respectively. With the ECN marking mechanism and minimum queue strategy, the BLUECN algorithm effectively relieves congestion and improves transmission efficiency. However, in large-scale networks and multi-source heterogeneous traffic scheduling scenarios, the traditional architecture still lacks flexibility and scalability in path control [17]. To address this, an SDN path selection algorithm named BLUE-SDN is proposed, which integrates the BLUECN and leverages the centralized control and global view of SDN to optimize path decisions. The overall framework is shown in Figure 1.

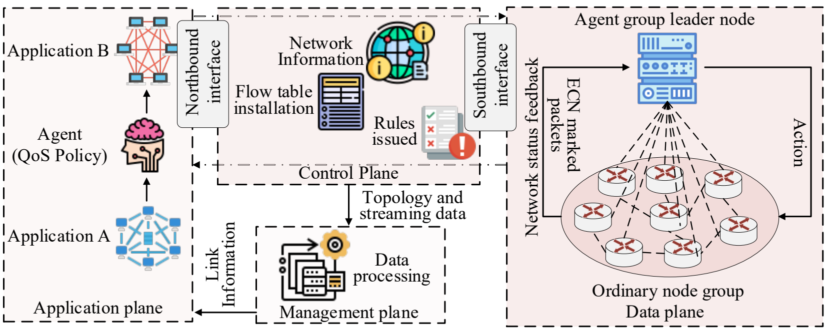

Figure 1. Schematic diagram of BLUE-SDN's optimization process

As shown in Figure 1, the application layer delivers Quality of Service (QoS) policies and service requirements to the control plane. The control plane evaluates network states, generates flow-table rules, and installs them on the data plane. The data plane comprises proxy-group leader nodes and regular nodes, where leader nodes collect state information and provide feedback to support real-time traffic adjustment and path optimization. ECN marking from proxy nodes indicates congestion levels and guides path scheduling decisions. Meanwhile, the management plane continuously monitors network operations and provides decision support to enhance overall system adaptability. Link transmission capability is also affected by available bandwidth, which is calculated in Equation (3).

$$\vartheta_{u,v} = \frac{B_{u,v}}{C_{u,v}} = 1 - \frac{Load_{u,v}}{C_{u,v}} \tag{3}$$

In Equation (3), $\vartheta_{u,v}$ and $B_{u,v}$ represent the available bandwidth rate and bandwidth of the link from u to v, while $C_{u,v}$ and $Load_{u,v}$ denote the maximum bandwidth and load. To achieve load balancing, traffic should be distributed evenly, as described in Equation (4).

$$\mu_{u,v} = \frac{\rho_{u,v} \times C_{u,v}}{C\left(D_u + D_v\right)} \tag{4}$$

In Equation (4), $\mu_{u,v}$ is the load parameter, $\rho_{u,v}$ is the interference factor, and D_u and D_v refer to the congestion certainty of links u and v.

3.2. SDN dynamic decision model integrating link utilization and reinforcement learning

Although the BLUE-SDN algorithm can optimize network traffic and overcome limitations of traditional static path selection methods, its flexibility and adaptability to traffic fluctuations remain insufficient under highly dynamic and variable network conditions. Therefore, this study introduces GCN and GRU to construct a traffic prediction model named GCN-GRU. The GCN component captures the complex relationships between network topologies and link states, while the GRU component processes temporal data, thereby improving prediction accuracy. The overall process is shown in Figure 2.

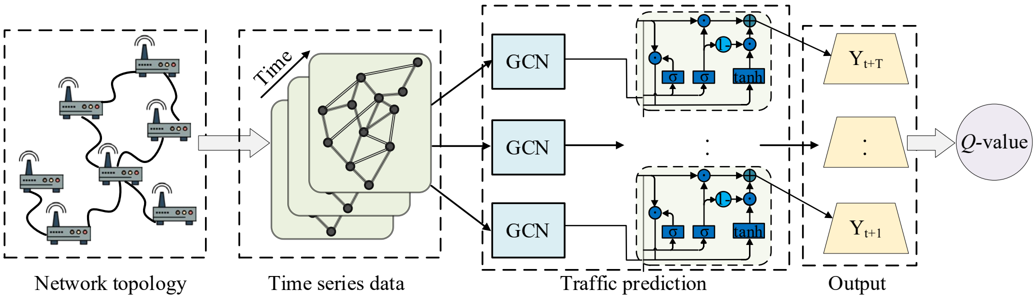

Figure 2. GCN-GRU framework diagram

As shown in Figure 2, the model first uses GCN to process link information in the network topology, learning and aggregating the relationship between network nodes and link state features to extract global network state information. Then, GRU uses these time-series features to predict future changes in network states, providing dynamic decision support for path selection. The node representation update within the GCN is governed by Equation (5).

$$h_p^{(l)} = g\left[\theta^{(l)} h_p^{(l-1)} A^*\right] = \left[\theta^{(l)} h_p^{(l-1)} D^{-\frac{1}{2}} A D^{\frac{1}{2}}\right] \tag{5}$$

In Equation (5), A and D are the adjacency and degree matrices, A^* represents normalization, $\theta^{(l)}$ is the learning weight of the l-th layer, $g[\cdot]$ is the activation function, and $h_p^{(l)}$ is the feature of node p in layer l. The calculations of different gates in GRU are shown in Equation (6) [18].

$$\begin{cases} z_t = \sigma\left(W_z x_t + U_z h_{t-1} + b_z\right) \\ r_t = \sigma\left(W_r x_t + U_r h_{t-1} + b_r\right) \end{cases} \tag{6}$$

In Equation (6), z_t and r_t denote the update gate and reset gate vectors at time t, $\sigma(\cdot)$ is the sigmoid activation function, x_t represents the input feature vector at time t, W_z and W_r are the input weight matrices for the update gate and reset gate, while U_z and U_r are the recurrent weight matrices corresponding to the previous hidden state h_{t-1}, b_z and b_r represent the bias terms of the update gate and reset gate, respectively. However, a standalone GCN-GRU structure still has limitations in highly dynamic and complex networks, especially in terms of load variation and real-time routing. Its adaptability and optimization ability need improvement [19]. Therefore, this study further introduces DQN to build an intelligent routing model named GURU-DQN, which combines GCN-GRU and DQN. This model uses deep reinforcement learning to enable the agent to learn and optimize decisions in the network environment, further enhancing the intelligence and adaptability of path selection. The target network is updated and calculates the target Q-values to adjust the Q-values of the online network. After each training round, the model uses sampled experience to update itself and optimizes the Q-values to improve the path selection strategy. After multiple training iterations, the agent learns the optimal path strategy, improving both performance and adaptability. The calculation of the target Q-value is shown in Equation (7) [20].

$$Q^{+}\left(S_t, A_t\right) = R_t + \gamma \cdot \max_{A} Q\left(S_{t+1}, A_{t+1}\right) \tag{7}$$

In Equation (7), $Q^{+}\left(S_t, A_t\right)$ represents the expected Q-value of the target network after taking action A_t at time t in state S_t. R_t is the reward received by the agent after executing A_t at time t, and γ is the discount factor. The reward function was designed by jointly considering delay, packet loss, link utilization, and load balancing. Specifically, the reward increased when the selected path achieved lower delay and packet loss with balanced link utilization, and decreased when congestion or excessive path switching occurred. The primary hyperparameters for DQN training are configured as follows: a learning rate of 0.001, a discount factor of 0.95, a batch size of 64, a replay buffer size of 5,000, an exploration rate decaying from 0.9 to 0.05, and a target network update interval of 20 episodes. In summary, this study combines the BLUE-SDN architecture based on link utilization with the GURU-DQN intelligent routing model using reinforcement learning, and proposes a new SDN dynamic path decision model named LUGDQ-SDN. The overall process is shown in Figure 3.

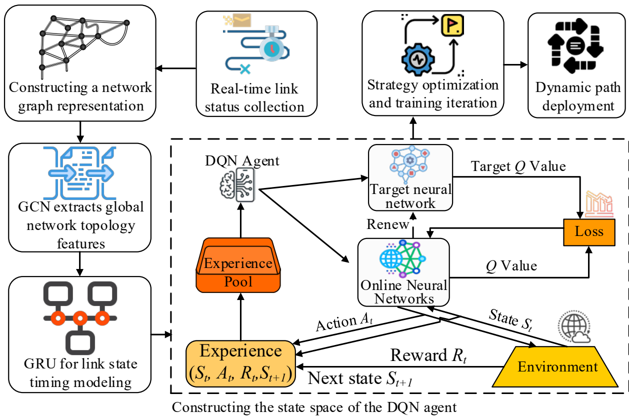

Figure 3. LUGDQ-SDN path dynamic decision model process

As shown in Figure 3, the model gathers some important parameters such as link utilizations and ECN markings in real-time via the SDN controller for creating an overview of the entire network. Afterward, the GCN-GRU layer is applied for extracting some temporal features of the network structure and links. The DQN agent takes the state features as input, generates transition samples (S_t, A_t, R_t, S_{t+1}), and stores them in the experience pool to support later strategy training. The online and target networks optimize the path strategy by updating Q-values. Finally, relying on SDN's centralized control, the model enables intelligent path delivery and dynamic global traffic scheduling, ensuring efficient and adaptive path decisions in highly dynamic network environments. In practical applications, each component of LUGDQ-SDN performs a specific function within the routing framework. The SDN controller collects network state information and manages path deployment, the BLU-ECN module provides congestion awareness and link load perception, the GCN-GRU module extracts spatial-temporal traffic features and predicts network state evolution, and the DQN module generates adaptive routing decisions based on the learned state representation. Through the coordinated interaction of these components, the application achieves intelligent path selection and dynamic traffic scheduling.

4. Performance of the LUGDQ-SDN model for dynamic path selection

4.1. Performance verification of the BLUE-SDN path selection algorithm

To verify the performance advantages of the proposed BLUE-SDN path selection algorithm in dynamic scheduling scenarios, this study compared it with SDN, Centralized-Shortest Path First (C-SPF), and Distributed-Bandwidth-Aware Routing (D-BAR). The selected baselines represented three typical routing strategies: traditional SDN routing based on centralized control, C-SPF based on static shortest-path selection, and D-BAR based on distributed bandwidth awareness. These methods covered static routing, bandwidth-aware routing, and distributed path selection, making them suitable for evaluating the advantages of BLUE-SDN in congestion awareness and dynamic scheduling. The experiments were performed using the Mininet-WiFi 2.3.1b simulation tool and the common topology of a wireless network. As the real-time monitor of network traffic, sFlow-RT was chosen. In addition, the GEANT academic backbone topology with 23 nodes and 36 links was chosen as the testing platform, and samples of 16-week link usage and traffic matrixes taken every 15 minutes were used. All experiments were executed on a 64-bit Windows 10 system equipped with an Intel Core i9-12900K CPU, 64 GB RAM, and an NVIDIA GeForce RTX 3090 GPU. Prior to training, the GEANT traffic matrices were normalized to the range of [0, 1], and the link utilization, delay, packet loss, and bandwidth were extracted as state features. The dataset was divided into training and testing sets at a ratio of 8:2. All comparison models were trained and tested under the same topology, traffic generation rules, controller configuration, and hardware environment to ensure fair comparison. Each experiment was repeated five times, and the average value was used for performance analysis. The virtual intelligent laboratory in this study refers to the Mininet-WiFi simulation environment integrated with the Ryu controller, sFlow-RT, and GEANT topology dataset. It can be accessed through the local experimental server by starting the controller, loading the network topology, and running the traffic monitoring and model training programs, thereby supporting reproducible SDN routing simulation and intelligent path decision verification. This study first compared the throughput and average transmission delay of the four algorithms. The results are shown in Figure 4.

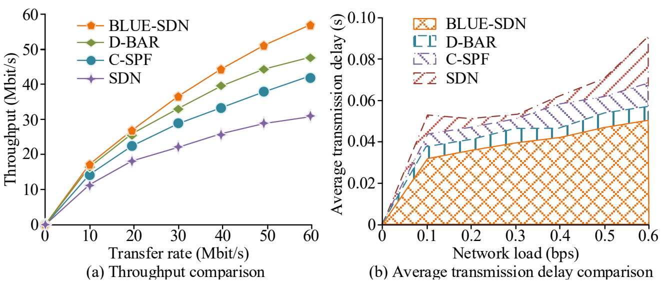

Figure 4. Comparison of throughput and average transmission delay

Figure 4(a) indicates that as the transmission rate grew, the throughput rate for the BLUE-SDN protocol stayed highest at 57 Mbit/s when the load was 60 Mbit/s. From Figure 4(b), it can be seen that the average transmission time for the BLUE-SDN protocol has been kept below 0.043 s as the load kept increasing. This demonstrated excellent real-time responsiveness and congestion mitigation. These results showed that BLUE-SDN performed well in both throughput assurance and delay control, effectively enhancing the intelligence level of path scheduling. This improvement is mainly because BLUE-SDN uses BLU and ECN feedback to identify congested links in advance and redirect traffic before severe queue accumulation occurs, thereby reducing waiting time and improving available path utilization. To further evaluate the overall performance of the proposed method in path selection, this study also compared the average link packet loss rate and link utilization of the four algorithms. The results are shown in Figure 5.

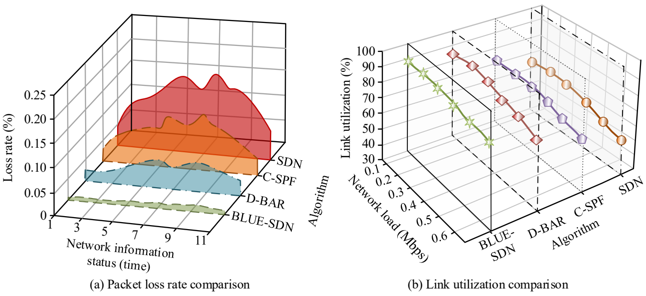

Figure 5. Comparison of average link packet loss rate and link utilization

As can be seen from Figure 5(a) in the context of dynamic networks, the BLUE-SDN algorithm was associated with the least mean packet loss ratio on all paths over the entire testing period, and the maximum value of packet loss ratios of this algorithm was 0.02%. The results clearly show the effectiveness of this algorithm with regard to the problem of detection and dynamic path routing due to the employment of BLU and ECN. It effectively reduced packet loss caused by congestion, ensuring network stability and reliability. As shown in Figure 5(b), under different network load conditions, BLUE-SDN consistently achieved the highest link utilization. When the load reached 0.6 Mbps, its utilization peaked at 89%, significantly higher than 77% for C-SPF, 68% for D-BAR, and 56% for traditional SDN. These results confirmed that BLUE-SDN could better utilize link bandwidth resources in dynamic scheduling scenarios, achieving collaborative optimization across congestion avoidance, load balancing, and throughput enhancement. The higher utilization does not indicate link overloading, but reflects that traffic is more evenly allocated to available paths while bottleneck links are dynamically avoided. This validated the proposed method's performance advantages in path selection.

4.2. Practical effect of LUGDQ-SDN model in dynamic path decision-making

Following confirmation of successful load balancing and congestion sensing using the BLUE-SDN algorithm for dynamic link scheduling, this research also investigated the adaptability and robustness of the LUGDQ-SDN framework in state sensing and optimization under varying conditions. The model was compared with Graph Attention Network with Deep Q-Network (GAT-DQN), Gated Recurrent Unit with Policy Gradient (GRU-PG), and Graph Convolutional Reinforcement Routing (GCRR). The experiments continued to use the GEANT dataset as the standard platform. The control and data planes were implemented using the Ryu controller and the Mininet-WiFi 2.3.1b environment, respectively. First, this study compared the convergence speed and the Root Mean Square Error (RMSE) of predicted traffic for the four models. The results are shown in Figure 6.

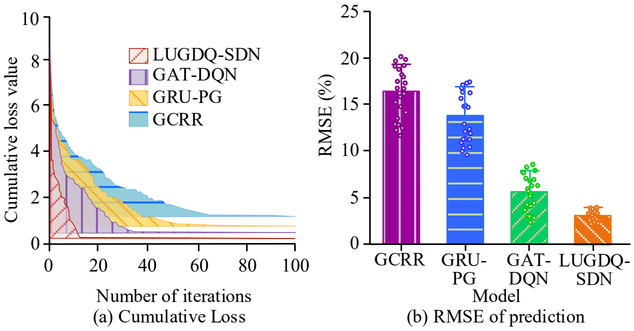

Figure 6. Convergence speed and RMSE comparison of four models

In Figure 6(a), the cumulative loss of the LUGDQ-SDN model rapidly converged after the 15th iteration and stabilized at 0.03, much faster than the 31, 45, and 64 iterations required by GAT-DQN, GRU-PG, and GCRR, respectively. This verified the model's fast convergence and training stability during strategy optimization. Furthermore, According to Figure 6(b), LUGDQ-SDN performed the best in RMSE traffic prediction with just 4.2% of RMSE, which was much smaller compared to all other models. It meant that it possessed more accurate modeling of the state space and traffic modeling. This is because the GCN-GRU module jointly captures topology-dependent spatial correlations and temporal traffic evolution, providing more informative state representations for the DQN agent and reducing unnecessary exploration during training. The findings suggested that the new model performed very well in terms of convergence speed and decision stability in complex networks, thus demonstrating great adaptability. In order to evaluate the decision-making performance of the paths, this study made comparisons in regard to load balancing and decision latency between the four models. The results are shown in Figure 7.

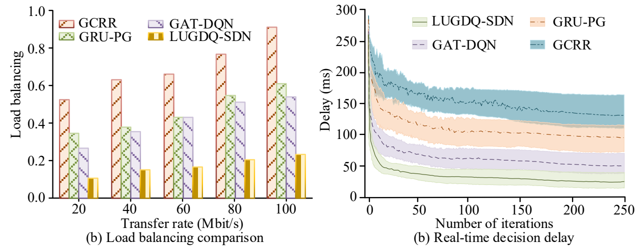

Figure 7. Comparison of the load balancing and real-time decision latency

As shown in Figure 7(a), the LUGDQ-SDN model consistently achieved the lowest load-balancing values across all transmission rates. At 100 Mbit/s, the value was only 0.23, significantly lower than 0.49 for GAT-DQN, 0.57 for GRU-PG, and 0.91 for GCRR. This indicated that the proposed model effectively avoided traffic congestion on local links and achieved distributed traffic scheduling and balanced link pressure. As can be seen from Figure 7(b), the curve of decision time lag in real time for LUGDQ-SDN stayed consistently the lowest one when the number of training steps increased, where the minimum delay was even only 31 ms on average, which was much better than other methods. This reflected a faster response to network states and more stable decision-making. The reduced decision latency is mainly attributed to the trained DQN policy, which directly maps the learned network state representation to routing actions without repeatedly recalculating paths from scratch. After evaluating routing efficiency, load balancing, and real-time decision performance, this study further investigated the robustness of the proposed model under network failure conditions. Considering that link interruption and node failure are common challenges in practical power communication systems, a failure-recovery experiment was conducted on the GEANT topology. During testing, 10% of backbone links were randomly disconnected, and the routing models were required to re-establish available paths dynamically. The results are summarized in Table 1.

Table 1. Robustness comparison under random link failure scenarios

Model	Path Recovery Success Rate (%)	Recovery Time (ms)	Throughput Retention Rate (%)
GCRR	85.7	87	72.4
GRU-PG	89.3	75	78.6
GAT-DQN	92.1	61	83.7
LUGDQ-SDN	96.8	42	89.5

As illustrated in Table 1, the highest path recovery success ratio, which is 96.8%, can be obtained using

LUGDQ-SDN while at the same time maintaining a throughput retention ratio of 89.5%. This happened after a link failure, and it has the least recovery time of 42 ms. This is mainly because the BLU-ECN mechanism can rapidly identify abnormal link states, while the GCN-GRU-DQN framework provides accurate state perception and adaptive rerouting decisions. These results indicate that the proposed model not only improves routing efficiency during normal operation but also exhibits strong robustness and service continuity under failure conditions.

To further verify the contribution of each component, ablation experiments were conducted by removing the BLU-ECN mechanism, the GCN-GRU module, and the DQN module from LUGDQ-SDN, respectively. The results are shown in Table 2.

Table 2. Ablation study of LUGDQ-SDN components

Model	RMSE (%)	Load Balancing Value	Decision Latency (ms)	Packet Loss Rate (%)
Without BLU-ECN	6.8	0.41	46	0.07
Without GCN-GRU	8.5	0.52	58	0.09
Without DQN	7.3	0.47	64	0.08
LUGDQ-SDN	4.2	0.23	31	0.02

The ablation results indicated that removing any component weakened the overall model performance. Without BLU-ECN, congestion perception became less accurate, without GCN-GRU, traffic prediction accuracy decreased, and without DQN, the adaptive decision-making capability was significantly reduced. This confirms that the three modules jointly contribute to the overall performance improvement of LUGDQ-SDN. Overall, the model has proved to be more effective in both load balancing and real-time capability, thereby fulfilling the primary requirements of intelligent routing in highly dynamic and resilient conditions. The results suggest that LUGDQ-SDN model has shown quicker convergence rate, accurate state identification, and reduced decision time. Consequently, it is an efficient tool for solving path scheduling problems in dynamic environments. These improvements are meaningful for power communication services because lower decision latency and packet loss can support time-sensitive dispatching and protection communication, while higher link utilization and better load balancing help maintain stable service transmission during traffic fluctuation or partial link degradation. Therefore, the proposed model has practical relevance for improving the communication resilience of smart grid infrastructures.

5. Discussion and Conclusion

5.1. Discussion

The proposed LUGDQ-SDN architecture may be implemented on an SDN controller as an intelligent routing application. Link-state data is collected by the controller using monitoring systems, processed through the GCN-GRU-DQN component, and then routing rules are established using the southbound interface. In actual implementation, the most important system constraints include real-time link monitoring capability, enough computational capabilities of the controller to run model inference, and support of OpenFlow switches. Since the online routing decision latency is 31 ms, the model is suitable for delay-sensitive routing services, but large-scale deployment may require lightweight model compression or edge-assisted inference. For migration from simulation to real SDN environments, the Mininet-WiFi topology scripts can be replaced by real switch and controller interfaces, while the same state collection, decision, and flow-rule installation logic can be retained. Future implementation should further consider multi-controller coordination and controller-failure recovery to enhance cross-domain scalability and operational reliability.

5.2. Conclusion

In response to response lag, uneven load distribution, and weak congestion control in traditional SDN path scheduling for power communication networks, this study proposes the LUGDQ-SDN model by integrating link utilization with reinforcement learning. First, this study designs a BLUE-SDN routing mechanism that combines BLU with explicit congestion notification to rapidly identify and dynamically bypass congested links. Experimental results show that under a 60 Mbit/s load, BLUE-SDN achieves 57 Mbit/s throughput, 0.043 s transmission delay, 0.02% packet loss rate, and 89% link utilization, significantly outperforming the comparison algorithms. Then, a GCN-GRU structure is introduced to capture topology-temporal correlation features and is further integrated with a deep Q-network for adaptive policy optimization, enabling LUGDQ-SDN to converge to a minimum loss of 0.03 within 15 iterations, reduce traffic prediction RMSE to 4.2%, and achieve a median real-time decision latency of only 31 ms. The proposed framework can be applied to smart grid communication networks, substation communication systems, distributed energy management networks, and other latency-sensitive SDN-based infrastructures requiring adaptive routing and reliable data transmission. Compared with conventional SDN routing methods, the proposed framework offers advantages in congestion awareness, adaptive path optimization, load balancing, and real-time decision-making, making it more suitable for highly dynamic power communication environments. However, multi-controller coordination in cross-domain deployments and device heterogeneity have not yet been considered. In addition, the current model still depends on simulation-based validation, and the

GCN-GRU-DQN structure may introduce additional computational overhead during large-scale deployment. Its performance under controller failures, extreme traffic bursts, and real power communication devices also requires further verification. Future work will extend validation in complex real-world networks to further enhance practicality and generalizability. In particular, future studies will focus on multi-controller collaborative routing, heterogeneous device adaptation, lightweight model deployment, failure-resilient path recovery, and real-time testing in practical power communication infrastructures.

Acknowledgments

The research work in this paper was supported by the Science and Technology Project of China Southern Power Grid Company [Project No.: 030000KC23120100 (GDKJXM20231515)], for which we would like to express our gratitude.

References

[1] Logeshwaran J, Shanmugasundaram RN, Lloret J. Load based dynamic channel allocation model to enhance the performance of device-to-device communication in WPAN. Wireless Netw. 2024;30(4):2477-2509. DOI:10.1007/s11276-024-03680-x.

[2] Wei W, Fu L, Gu H, Zhang Y, Zou T, Wang C, Wang N. GRL-PS: Graph embedding-based DRL approach for adaptive path selection. IEEE Trans Netw Serv Manag. 2023;20(3):2639-2651. DOI:10.1109/TNSM.2023.3240396.

[3] Veisi F, Montavont J, Theoleyre F. Enabling centralized scheduling using software defined networking in industrial wireless sensor networks. IEEE Internet Things J. 2023;10(23):20675-20685. DOI:10.1109/JIOT.2023.3302994.

[4] Mehraban S, Yadav RK. Traffic engineering and quality of service in hybrid software defined networks. China Commun. 2024;21(2):96-121. DOI:10.23919/JCC.fa.2022-0860.202402.

[5] Liu CL, Huang TH. Dynamic job-shop scheduling problems using graph neural network and deep reinforcement learning. IEEE Trans Syst Man Cybern Syst. 2023;53(11):6836-6848. DOI:10.1109/TSMC.2023.3287655.

[6] Ye X, Yu Y, Fu L. Deep reinforcement learning based link adaptation technique for LTE/NR systems. IEEE Trans Veh Technol. 2023;72(6):7364-7379. DOI:10.1109/TVT.2023.3236791.

[7] He Q, Wang Y, Wang X, Xu W, Li F, Yang K, Ma L. Routing optimization with deep reinforcement learning in knowledge defined networking. IEEE Trans Mob Comput. 2024;23(2):1444-1455. DOI:10.1109/TMC.2023.3235446.

[8] Li D, Zhang Z, Alizadeh B, Zhang Z, Duffield N, Meyer MA, Behzadan AH. A reinforcement learning-based routing algorithm for large street networks. Int J Geogr Inf Sci. 2024;38(2):183-215. DOI:10.1080/13658816.2023.2279975.

[9] He J, Cheng N, Yin Z, Zhou C, Zhou H, Quan W, Lin XH. Service-oriented network resource orchestration in space-air-ground integrated network. IEEE Trans Veh Technol. 2024;73(1):1162-1174. DOI:10.1109/TVT.2023.3301676.

[10] Yang L, Liang J, Wang R, Liu X, De Sanctis M, Burleigh SC, Zhao K. A study of Licklider transmission protocol in deep-space communications in presence of link disruptions. IEEE Trans Aerosp Electron Syst. 2023;59(5):6179-6191. DOI:10.1109/TAES.2023.3273523.

[11] Taha M. An efficient software defined network controller based routing adaptation for enhancing QoE of multimedia streaming service. Multimed Tools Appl. 2023;82(22):33865-33888. DOI:10.1007/s11042-023-14938-5.

[12] Ritzkal R, Suhadi S, Amalia R, Afrianto Y, Triawan A, Syafrial S, Fatimah F. K-nearest neighbor algorithm analysis for path determination in network simulation using software defined network. Bull Electr Eng Inform. 2023;12(4):2388-2400. DOI:10.11591/eei.v12i4.4868.

[13] Zheng H, Guo J, Zhou Q, Peng Y, Chen Y. Application of improved ant colony algorithm in load balancing of software-defined networks. J Supercomput. 2023;79(7):7438-7460. DOI:10.1007/s11227-022-04957-8.

[14] Wu C, Zhang Y, Li N, Rezaeipanah A. An intelligent fuzzy-based routing algorithm for video conferencing service provisioning in software defined networking. Telecommun Syst. 2024;87(4):887-898. DOI:10.1007/s11235-023-01044-y.

[15] Ye Z, Sun G, Guizani M. ILBPS: An integrated optimization approach based on adaptive load-balancing and heuristic path selection in SDN. IEEE Internet Things J. 2024;11(4):6144-6157. DOI:10.1109/JIOT.2023.3309832.

[16] Rafid M, Ghafoor H, Koo I. An SVM-based optimal-controller selection and path selection protocol for heterogeneous communications in SDVNs. IEEE Trans Intell Transp Syst. 2024;25(5):3828-3842. DOI:10.1109/TITS.2023.3326651.

[17] Liu Y, Huo L, Zhang X, Wu J. A multi-objective resource pre-allocation scheme using SDN for intelligent transportation system. IEEE Trans Intell Transp Syst. 2024;25(1):571-586. DOI:10.1109/TITS.2023.3310978.

[18] Sheng Z, Wen S, Feng ZK, Shi K, Huang T. A novel residual gated recurrent unit framework for runoff forecasting. IEEE Internet Things J. 2023;10(14):12736-12748. DOI:10.1109/JIOT.2023.3254051.

[19] Prabha R, Senthil GA, Bharathi GP, Sridevi S. Hybrid multipath routing cluster head prediction based on SDN-enabled IoT and heterogeneous context-aware graph convolution network. Peer-to-Peer Netw Appl. 2024;17(4):2016-2030. DOI:10.1007/s12083-024-01685-z.

[20] Huang L, Ye M, Xue X, Wang Y, Qiu H, Deng X. Intelligent routing method based on Dueling DQN reinforcement learning and network traffic state prediction in SDN. Wireless Netw. 2024;30(5):4507-4525. DOI:10.1007/s11276-022-03066-x.c