Integrated GNN and DRL in E2E Networking Solutions

Subjects: Computer Science, Theory & Methods
Contributor: Prohim Tam , Seyha Ros , Inseok Song , Seungwoo Kang , Seokhoon Kim

Graph neural networks (GNN) and deep reinforcement learning (DRL) are at the forefront of algorithms for advancing network automation with capabilities of extracting features and multi-aspect awareness in building controller policies. While GNN offers non-Euclidean topology awareness, feature learning on graphs, generalization, representation learning, permutation equivariance, and propagation analysis, it lacks capabilities in continuous optimization and long-term exploration/exploitation strategies. Therefore, DRL is an optimal complement to GNN, enhancing the applications towards achieving specific policies within the scope of end-to-end (E2E) network automation.

deep reinforcement learning | end-to-end networking | graph neural networks | network automation | optimization approaches

1. Introduction

Following the establishment of comprehensive advanced 5G and 6G standards, 2019 to 2023 has witnessed the pioneering commercial deployment of fast-speed wireless networks, which supports the advent of smart digital transformation. The internet evolution presents advancements in ultra-reliable low-latency, high-throughput, mobility-aware, and high-coverage connectivity that set a new benchmark compared to the previous network generations [1][2]. Forecasts by the International Telecommunication Union (ITU) anticipate exponential growth in global mobile data traffic, with projections extending from 390 exabytes to 5016 exabytes between 2024 and 2030, respectively [3]. As digital transformation and its volume expand with the benefits of widespread coverage and lightning-fast connections, it also faces significant challenges in managing the growth in data, devices, and services [4][5]. To address these evolving challenges, a shift towards network automation is essential to breaking down barriers within end-to-end (E2E) solutions, which spans three domains: radio access networks (RAN), transport networks, and core networks.

Traditional RAN requires redesigning with AI-empowered control [6], shared cloudification [7], optimized power allocation [8][9], and highly programmable handover and interoperability [10]. During the redesign process, initial challenges arise in data exposure capability and the level of network infrastructure knowledge necessary to support rich-feature input and processing for network automation. Considering the significant objectives of integrating AI, O-RAN, and software-defined networking (SDN)-enabled management, the ability to encode network conditions (signal, interference, spectrum availability, etc.) and decode hidden relationships between each timeslot remains burdensome. Furthermore, transport and core networks also require the ability to understand traffic (congestion) patterns, resource utilization, and anomaly detection in complex topology graphs [11][12][13]. Therefore, before focusing on other potential issues in E2E networking, one key research is the selection of optimization algorithms that handle complex graph-structured topologies and extract data to support self-organizing capabilities [14][15].

Previous works supported by standardization, academia, and industry experts, are coming to conduct the creation of cutting-edge testbeds and simulation tools for network intelligence [16][17][18][19]. The motivation from existing testbeds has guided researchers towards integrating three key objectives, namely zero-touch autonomy, topology-aware scalability, and long-term efficiency, into network and service management [20][21]. In terms of these goal-oriented optimizations, graph neural networks (GNN) [22][23][24] and deep reinforcement learning (DRL) [25][26][27] are at the forefront of algorithms for advancing network automation with capabilities of extracting features and multi-aspect awareness in building controller policies. While GNN offers non-Euclidean topology awareness, feature learning on graphs, generalization, representation learning, permutation equivariance, and propagation analysis [28][29][30][31], it lacks capabilities in continuous optimization and long-term exploration/exploitation strategies. Therefore, DRL is an optimal complement to GNN, enhancing the applications towards achieving specific policies within the scope of E2E network automation.

2. GNN

2.1. GNN and Its Variants

GNN represents a class of deep learning models designed to perform inference on data structured as graphs. Initially, GNN is particularly powerful for tasks where the data are inherently graph structured, such as social networks [32], chemistry [33], and communication networks [34]. The core idea behind GNN is to learn representations (embeddings) for each node/edge that capture both (1) key features and (2) the structure of local graph neighborhood. GNN iteratively updates the representation of a node by aggregating representations of its neighboring nodes and combining them with its current representation.

Several well-known variants of GNNs have been developed, where each with its own approach to modify on aggregation and updating steps, including (1) graph convolutional networks (GCN) [35] simplify the aggregation step by using a weighted average of neighbor features, where weights are typically based on the degree of the nodes; (2) graph attention networks (GAT) [36] introduce attention mechanisms to weigh the importance of each neighbor’s features during aggregation dynamically; (3) GraphSAGE [37] extend GNN by sampling a fixed-size neighborhood for each node and using various aggregation functions, such as mean, LSTM, or pooling; (4) message passing neural networks (MPNN) [38] generalize several GNN models by defining a message passing framework, where messages (aggregated features) are passed between nodes; (5) edge-node GNN [39] target on edge updates alongside node updates for radio resource management, which demonstrated superior performance in beamforming and power allocation to achieve higher rates with less computation time.

2.2. Applied GNN in E2E Networking

Beyond traditional networking approaches, GNN offers a paradigm shift for network intelligence through the capability to model and analyze the hidden relationships and dynamic attributes in graph-structured massive network topologies. Furthermore, GNN with permutation equivariance offers a significant advantage in communication networks by treating equivalent network configurations, even if nodes swap positions, as the same from a network function perspective. This key factor translates to reduced training effort, making GNN particularly well suited for analyzing and optimizing complex network structures [39][40].

3. DRL

3.1. DRL and Its Variants

DRL combines the principles of reinforcement learning with the representation learning capabilities of deep neural networks (DNN) by (1) enabling agents to learn optimal policies for decision making, (2) interacting with the environment through observing states and applying actions, (3) receiving feedback by proposing specific reward functions, and (4) targeting to maximize cumulative long-term rewards [41]. The foundations of DRL involve the Bellman equation used to update the value, as Equations (3) and (4), where (1) $V(s)$ is the value of state s, (2) $Q(s,a)$ is the value of taking action a in state s, (3) Rt is the reward at time t, and (4) γ is the discount factor.

3.2. Applied DRL in E2E Networking

DRL marks a significant evolution in networking intelligence, diverging from conventional strategies by its adaptability and learning-driven approach to optimize network functions [42][43][44][45]. **Table 1** outlines DRL notable studies in E2E networking contexts, including the networking domains, key remarks, state observation, action implementation, and reward targets.

Table 1. Selected comprehensive works on applied DRL.

Network Domains	Key Remarks	State	Action	Reward	Ref.	Year
Access networks: (1) maximizing the sum rate (2) adhering low latency requirements in smart transportation services	Utilization of an attention mechanism to focus on relevant state information among agents	Partial CSI, including received interference information, remaining payload, and remaining time for V2V agents	Sub-band selection and power allocation for V2V agents	Maximization of the total throughput on V2I links while ensuring low latency and high reliability for V2V links	[46]	2022
Access networks: (1) optimizing total weighted	Model-free DRL framework employing Q-learning with	Global network state including task requests from ground	(1) task offloading decisions (local processing or	The negative weighted sum of task processing	[47]	2021

Network Domains	Key Remarks	State	Action	Reward	Ref.	Year
costs for task offloading and resource allocation in an SDN-enabled Multi-UAV-MEC network	enhancements to handle the mixed-integer conditions of task offloading and resource allocation	equipment, available UAV resources, and current network configurations	offloading to a UAV) and (2) resource allocation strategies (assigning computation resources to tasks)	delay and energy consumption		
Transport networks: (1) maximizing overall system throughput for real-time traffic demand across autonomous systems	Utilization of policy gradients and handling partial observability while adopting actor-critic algorithms for stability	Source and destination of flows, current traffic loads on links to neighbors, and observed throughputs	Selection of next-hops for routing traffic flows	Average throughput of all concurrent flows traversing an agent	[48]	2020
Transport networks: (1) optimizing the routing decisions by minimizing delay and loss while maximizing throughput	The proposed model used DQN for SDN to proactively compute optimal routes (leveraging path-state metrics for dynamic traffic adaptation)	Source-destination pairs	Selection of specific E2E routing paths	Path-state metrics including path bandwidth, path delay, and path packet loss ratio	[49]	2021
Core networks: (1) optimizing the allocation of VNF forwarding graphs to maximize the number of accepted requests	Enhanced DDPG with heuristic fitting algorithm to translate actions into allocation strategies	VNF forwarding graphs, including computing resources for VNFs and QoS requirements for VLs	Allocation decisions for VNFs on substrate nodes and paths for VLs	Acceptance ratio based on successful deployment of VNFs and VLs while meeting resources and QoS requirements	[50]	2019
Core networks: (1) optimizing adaptive online orchestration of NFV while focusing on maximizing E2E QoE of all arriving service requests	Utilization of a policy gradient-based approach with Q-learning enhancements to handle the state transitions and real-time network state changes	CPU, memory bandwidth, delay, orchestration results of executing SFC, and the arrival requests with different QoS requirements	The allocation of network resources and VNFs to fulfill the request	Maximizing QoE while satisfying QoS constraints	[51]	2021

1. Salh, A.; Audah, L.; Shah, N.S.M.; Alhammadi, A.; Abdullah, Q.; Kim, Y.H.; Al-Gailani, S.A.; Hamzah, S.A.; Esmail, B.A.F.; Almohammedi, A.A. A Survey on Deep Learning for Ultra-Reliable

4. Integrated GNN and DRL in E2E Networking Solutions

The synergy of GNN and DRL capitalizes on (1) **GNN**: the capability to encode complex graph environments, approximate actions/rewards, and compute q-values, along with (2) **DRL**: the ability to explore GNN architectures and evaluate the accuracy of readout predictions. **Figure 1** presents the overview of fusing both algorithms and key features that complement each other. Together, GNN+DRL extract auxiliary network states, advance generalization/adaptability, and adopt data-driven learning for multi-aspect awareness reward functions towards pioneering network automation.

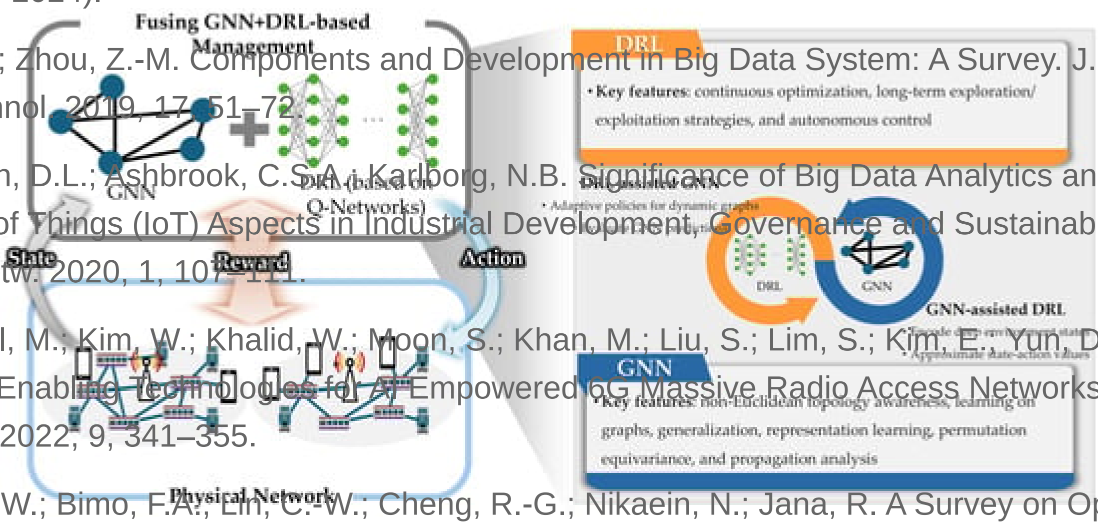

Figure 1. Overview of GNN+DRL and the key features.

4.1. Access Networks

Figure 2 illustrates the schematic representation of the wireless network input in relation to the policy objectives that emphasize the strategic applications of integrating GNN and DRL. The key to understanding how GNN works is focusing on how graph information is input to subsequent hidden layers, which primarily involves the concepts of message passing, aggregation, feature transformation, and update mechanisms that enable the network to learn from the graph structures and node features. After the initial round, the updated node features can serve as input to the next hidden layers. Each hidden layer can perform its own steps, which allows the network to capture more complex patterns and relationships at higher levels of abstraction. The depth of the network (number of hidden layers) typically correlates with the reach of a node (e.g., how many hops away in the graph the node information can propagate from).

and Low-Latency Communications: Challenges in 6G Wireless Systems. IEEE Access 2021, 9, 55098–55131.

2. Zhou, W.; Islam, A.; Chang, K. Real-Time RL-Based 5G Network Slicing Design and Traffic Model Distribution: Implementation for V2X and EMBB Services. KSII Trans. Internet Inf. Syst. 2023, 17, 2573–2589.
3. IMT Traffic Estimates for the Years 2020 to 2030, Document ITU-R SG05, July 2015. Available online: https://www.itu.int/dms_pub/itu-r/opb/rep/R-REP-M.2370-2015-PDF-E.pdf (accessed on 2 February 2024).
4. Yu, J.-H.; Zhou, Z.-M. Components and Development in Big Data System: A Survey. J. Electron. Sci. Technol. 2019, 17, 51–72.
5. Andersen, D.L.; Ashbrook, C.S.A.; Karlborg, N.B. Significance of Big Data Analytics and the Internet of Things (IoT) Aspects in Industrial Development, Governance and Sustainability. Int. J. Intell. Netw. 2020, 1, 107–111.
6. Shahjalal, M.; Kim, W.; Khalid, W.; Moon, S.; Khan, M.; Liu, S.; Lim, S.; Kim, E.; Yun, D.-W.; Lee, J.; et al. Enabling Technologies for AI Empowered 6G Massive Radio Access Networks. ICT Express 2022, 9, 341–355.
7. Azariah, W.; Bimo, F.A.; Lin, C.-W.; Cheng, R.-G.; Nikaein, N.; Jana, R. A Survey on Open Radio Access Networks: Challenges, Research Directions, and Open Source Approaches. Sensors 2024, 24, 1038.
8. Li, C. Optimal Power Allocation for NOMA-Based Cellular Two-Way Relaying. KSII Trans. Internet Inf. Syst. 2023, 17, 202–215.
9. Xu, Y.; Liu, F.; Zhang, Z.; Sun, Z. Uplink Achievable Rate Analysis of Massive MIMO Systems in Transmit-Correlated Ricean Fading Environments. KSII Trans. Internet Inf. Syst. 2023, 17, 261–279.
10. Mangipudi, P.K.; McNair, J. SDN Enabled Mobility Management in Multi Radio Access Technology 5G Networks: A Survey. arXiv 2023, arXiv:2304.03346.
11. Wang, N.; Wang, H.; Wang, X. Service Deployment Strategy for Customer Experience and Cost Optimization under Hybrid Network Computing Environment. KSII Trans. Internet Inf. Syst. 2023, 17, 3030–3049.
12. Tian, Z.; Patil, R.; Gurusamy, M.; McCloud, J. ADSeq-5GCN: Anomaly Detection from Network Traffic Sequences in 5G Core Network Control Plane. In Proceedings of the 2023 IEEE 24th International Conference on High Performance Switching and Routing (HPSR), Albuquerque, NM, USA, 5–7 June 2023.

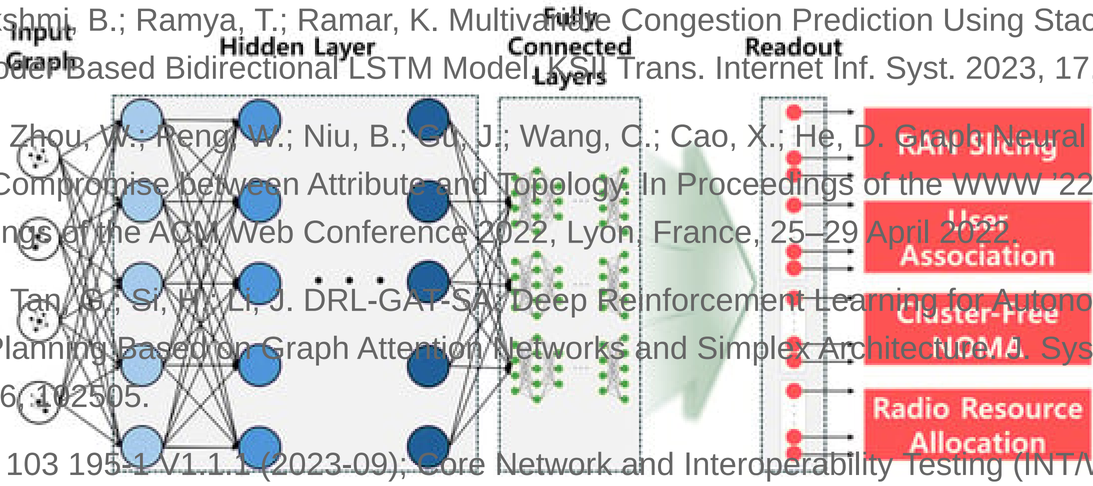

Figure 2. Schematic graph processing from input network graphs towards access network policies.

4.1.1. RAN Slicing

Arash et al. [52] proposed a GNN-based multi-agent DRL framework for RAN/mobile edge computing (MEC) slicing and admission control in 5G metropolitan networks. The authors leveraged GAT and GATv2 for topology-independent feature extraction, which enabled scalability and generalizability across different networks. The approach used multi-agent DRL, combining a GNN-based slicing agent with a topology-independent multi-layer perceptron (MLP) for admission control for optimizing long-term revenue under E2E service delay and resource constraints. The framework demonstrated significant improvements in infrastructure provider's revenue, achieving up to 35.2% and 25.5% overall gain over other DRL-based and heuristic baselines. The proposed scheme maintained good performances without re-training or re-tuning, even when applied to unseen network topologies, which showcased its generalizability and robustness.

4.1.2. Radio Resource Allocation

In E2E solutions, efficient radio resource allocation is crucial for optimal service delivery while ensuring fairness, quality of service (QoS), efficiency, and cost-effectiveness in operational expenses. Zhao et al. [53] introduced graph reinforcement learning by first transforming the traditional state and action representations from matrices to graphs, which enabled the functionality of GNN in capturing graph-structured network topologies and node-level relationships efficiently. The graph-based representation was then utilized within a DDPG framework, where the actor and critic networks were adapted to handle graph inputs to allow the model to learn optimal policies for resource allocation. The proposed approach not only reduced the dimensionality of the input data but also captured the relational dynamics between network elements more effectively than traditional methods. The results showcased significant improvements in training efficiency and performance for radio resource allocation tasks. The graph-based DDPG algorithm demonstrated faster convergence, lesser computing resource consumption, and lower space complexity compared to traditional DDPG algorithms.

4.1.3. User Association

13. Vijayalakshmi, B.; Ramya, T.; Ramar, K. Multivariate Congestion Prediction Using Stacked LSTM Autoencoder Based Bidirectional LSTM Model. KSII Trans. Internet Inf. Syst. 2023, 17, 216–238.

14. Yang, L.; Zhou, W.; Peng, W.; Niu, B.; Gu, J.; Wang, C.; Cao, X.; He, D. Graph Neural Networks beyond Compromise between Attribute and Topology. In Proceedings of the WWW '22: Proceedings of the ACM Web Conference 2022, Lyon, France, 25–29 April 2022.

15. Peng, Y.; Tan, G.; Si, H.; Li, J. DRL-GAT-SA: Deep Reinforcement Learning for Autonomous Driving Planning Based on Graph Attention Networks and Simplex Architecture. J. Syst. Archit. 2022, 126, 102505.

16. ETSI TR 103 195-1 V1.1.1 (2023-09); Core Network and Interoperability Testing (INT/WG AFI) Generic Autonomic Network Architecture. Part 1: Business Drivers for Autonomic Networking. ETSI: Sophia Antipolis, France, 2023.

17. GENI Testbed. Available online: https://github.com/GENI-NSF (accessed on 2 February 2024).

18. Ros, S.; Tam, P.; Kang, S.; Song, I.; Kim, S. A survey on state-of-the-art experimental simulations for privacy-preserving federated learning in intelligent networking. Electron. Res. Arch. 2024, 32, 1333–1364.

19. Rajab, M.E.; Yang, L.; Shami, A. Zero-Touch Networks: Towards Next-Generation Network Automation. arXiv 2024, arXiv:2312.04159.

20. Mehmood, K.; Kralevska, K.; Palma, D. Intent-Driven Autonomous Network and Service Management in Future CellularR2 Networks: A Structured Literature Review. Comput. Netw. 2022, 220, 109477.

21. Bringhenti, D.; Marchetto, G.; Sisto, R.; Valenza, F. Automation for Network Security Configuration: State of the Art and Research Trends. ACM Comput. Surv. 2023, 56, 1–37.

22. He, S.; Xiong, S.; Ou, Y.; Zhang, J.; Wang, J.; Huang, Y.; Zhang, Y. An Overview on the Application of Graph Neural Networks in Wireless Networks. IEEE Open J. Commun. Soc. 2021, 2, 2547–2565.

23. Jiang, W. Graph-Based Deep Learning for Communication Networks: A Survey. Comput. Commun. 2022, 185, 40–54.

24. Tam, P.; Song, I.; Kang, S.; Ros, S.; Kim, S. Graph Neural Networks for Intelligent Modelling in Network Management and Orchestration: A Survey on Communications. Electronics 2022, 11, 3371.

25. Munikoti, S.; Agarwal, D.; Das, L.; Halappanavar, M.; Natarajan, B. Challenges and Opportunities in Deep Reinforcement Learning with Graph Neural Networks: A Comprehensive Review of Algorithms and Applications. IEEE Trans. Neural Netw. Learn. Syst. 2023, 1–21.

Ibtihal et al. [54] proposed DQN-GNN processing flow for optimizing user association in wireless networks involves a sequence of steps. Initially, the system represents the user association problem as a graph, where nodes correspond to users or base stations (BS), and edges represent wireless connections. A GNN is then used to encode this graph structure by learning a representation for each node to understand the importance and connectivity within the network. Following these steps, a DQN agent is trained to decide the best base station for user connection based on the network state, which includes user–BS associations and other network parameters. The integration of GNN with DQN leverages the encoded graph structure to inform the DQN agent decisions, which aims to optimize network performance by selecting the optimal user–BS associations to maximize the reward evaluation.

4.1.4. Cluster-Free NOMA

A NOMA framework is designed to enhance the flexibility of successive interference cancellation operations, which eliminates the need for user clustering. The cluster-free objective aims to efficiently mitigate interference and improve system performance by enabling more adaptable and scenario-responsive NOMA communications. Xu et al. [55] proposed a comprehensive framework that significantly increases the flexibility of successive interference cancellation operations, which is supported by advanced DRL with GNN paradigms (automated learning GNN termed as AutoGNN) to achieve scenario-adaptive and efficient communications in next-generation multiple access environments. The proposed algorithm leveraged the GNN+DRL integration to minimize interference and optimize beamforming in a flexible flow for cluster-free NOMA setting. The results highlighted that the proposed AutoGNN approach for cluster-free NOMA can outperform conventional cluster-based NOMA across various channel correlations.

4.2. Transport Networks

4.2.1. Routing Optimization

Swaminathan et al. [56] proposed GraphNET approach, by integrating GNN with DRL frameworks to optimize routing decisions in SDN. There are two primary phases, namely inference and training. Initially, a network state matrix synchronized with the proposed GNN, which then predicts the most optimal path with minimal delay. The GNN, acting as a DQN within the DRL framework, is trained using experiencing routing episodes, which employs a custom reward function focused on packet delivery and minimizing delays. The GNN+DRL algorithm significantly reduced packet drops and achieved lower average delays compared to traditional Q-routing and shortest path algorithms.

4.2.2. Flow Migration

Sun et al. [57] proposed an optimization approach on flow migration, which referred to the dynamic relocation of traffic among different network function instances to adapt the loading statuses and balancing between network service quality and resource utilization efficiency. The proposed framework was termed DeepMigration, which utilized (1) GNN to handle graph-structured topology and flow distribution and (2) DRL for generating flow migration policies, while maximizing QoS satisfactions and minimizing resource consumption. DeepMigration demonstrated

26. Luong, N.C.; Hoang, D.T.; Gong, S.; Niyato, D.; Wang, P.; Liang, Y.-C.; Kim, D.I. Applications of Deep Reinforcement Learning in Communications and Networking: A Survey. IEEE Commun. Surv. Tutor. 2019, 21, 3133–3174.
27. Nie, M.; Chen, D.; Wang, D. Reinforcement Learning on Graphs: A Survey. IEEE Trans. Emerg. Top. Comput. Intell. 2023, 7, 1065–1082.
28. Tang, H.; Liu, Y. Towards Understanding the Generalization of Graph Neural Networks. arXiv 2023, arXiv:2305.08048.
29. Liu, S.; Wu, C.; Zhu, H. Topology-Aware Graph Neural Networks for Learning Feasible and Adaptive AC-OPF Solutions. IEEE Trans. Power Syst. 2022, 38, 5660–5670.
30. Luo, D.; Cheng, W.; Yu, W.; Zong, B.; Ni, J.; Chen, H.; Zhang, X. Learning to Drop: Robust Graph Neural Network via Topological Denoising. In Proceedings of the 14th ACM International Conference on Web Search and Data Mining, Virtual Event, 8–12 March 2021.
31. Almasan, P.; Suárez-Varela, J.; Rusek, K.; Barlet-Ros, P.; Cabellos-Aparicio, A. Deep Reinforcement Learning Meets Graph Neural Networks: Exploring a Routing Optimization Use Case. Comput. Commun. 2022, 196, 184–194.
32. Fan, W.; Ma, Y.; Li, Q.; He, Y.; Zhao, E.; Tang, J.; Yin, D. Graph Neural Networks for Social Recommendation. In Proceedings of the World Wide Web Conference on—WWW '19, San Francisco, CA, USA, 13–17 May 2019.
33. Reiser, P.; Neubert, M.; Eberhard, A.; Torresi, L.; Zhou, C.; Shao, C.; Metni, H.; van Hoesel, C.; Schopmans, H.; Sommer, T.; et al. Graph Neural Networks for Materials Science and Chemistry. Commun. Mater. 2022, 3, 93.
34. Suárez-Varela, J.; Almasan, P.; Ferriol-Galmés, M.; Rusek, K.; Geyer, F.; Cheng, X.; Xiang, S.; Xiao, S.; Scarselli, F.; Cabellos-Aparicio, A.; et al. Graph Neural Networks for Communication Networks: Context, Use Cases and Opportunities. IEEE Netw. 2021, 37, 146–153.
35. Zhang, S.; Tong, H.; Xu, J.; Maciejewski, R. Graph Convolutional Networks: A Comprehensive Review. Comput. Soc. Netw. 2019, 6, 11.
36. Wang, X.; Ji, H.; Shi, C.; Wang, B.; Ye, Y.; Cui, P.; Yu, P.S. Heterogeneous Graph Attention Network. In Proceedings of the World Wide Web Conference 2019, San Francisco, CA, USA, 13–17 May 2019.
37. Liu, T.; Jiang, A.; Zhou, J.; Li, M.; Kwan, H.K. GraphSAGE-Based Dynamic Spatial–Temporal Graph Convolutional Network for Traffic Prediction. IEEE Trans. Intell. Transp. Syst. 2023, 24, 11210–11224.
38. Gilmer, J.; Schoenholz, S.S.; Riley, P.F.; Vinyals, O.; Dahl, G.E. Message Passing Neural Networks. Machine Learning Meets Quantum Physics. Lect. Notes Phys. 2020, 968, 199–214.

significant performance improvements in network functions virtualization (NFV)-enabled migration by reducing the cost and saving up to 71.6% of computation time compared to selected baselines.

4.2.3. Traffic Steering

Rafiq et al. [58] integrated RouteNet model [59] with a delay-aware traffic flow steering module for optimal SFC deployment and traffic steering in SDN controller. The proposed scheme predicted optimal paths considering delays through GNN. The system autonomously selected paths with minimal delay for traffic steering and SFC deployment by leveraging the knowledge plane for decision making. As a result, the system demonstrated efficient resource utilization and optimal SFC deployment across different scenarios. For instance, deploying 5 VNFs across separate compute nodes showed the model's capability in the experiment to efficiently allocate resources, while achieving significant improvements in latency and resource management.

4.2.4. Dynamic Path Reconfiguration

Liu et al. [60] introduced a novel GNN-based dynamic resource prediction model and deep dyna-Q-based reconfiguration algorithm for optimizing SFC paths in IoT networks. The proposed GNN model was used for forecasting VNF instance resource requirements for facilitating proactive reconfiguration decisions. The system dynamically adapted SFCs based on predicted and real-time data that aim to balance between resources and service performances. The authors addressed the SFC reconfiguration problem by proposing a trade-off optimization between maximizing revenue and minimizing reconfiguration costs, including both migration and bandwidth expenses. Utilizing deep dyna-Q-based method, the study overcome the NP-hard nature of the problem, while integrating with GNN for graph-structured scalability. The effectiveness of the proposed model was validated against exact solutions for small networks. The experimental evaluation demonstrated the model's effectiveness with an average CPU root-mean-square error (RMSE) of 0.17 on improved GNN, which was significantly lower than 0.75 achieved by original GNN.

4.3. Core Networks

4.3.1. VNF Optimization

By leveraging the virtualization and softwarization from SDN/NFV-based infrastructure, GNN+DRL can obtain efficient computing capabilities with replay buffer for multi-epoch training towards the optimization of VNF placements, as shown in **Figure 3**. Sun et al. [61] proposed a combination of DRL framework with graph network-based neural network for optimal VNF placement, which addresses the challenges of resource constraints in different VNF identifiers and QoS requirements in massive network traffic. The authors proposed DeepOpt architecture to operate within an SDN-enabled environment, where graph network's is utilized to generalize network topology (resource, storage, bandwidth, and tolerable delays).

39. Wang, Y.; Li, Y.; Shi, Q.; Wu, Y.-C. ENGNN: A General Edge-Update Empowered GNN Architecture for Radio Resource Management in Wireless Networks. Available online. arXiv 2022, arXiv:2301.00757.

40. Shen, Y.; Shi, Y.; Zhang, J.; Letaief, K.B. Graph Neural Networks for Scalable Radio Resource Management: Architecture Design and Theoretical Analysis. IEEE J. Sel. Areas Commun. 2021, 39, 101–115.

41. Li, Y. Deep Reinforcement Learning: An Overview. arXiv 2017, arXiv:1701.07274.

42. Xiang, H.; Zhang, M.; Jia, C. Federated Deep Reinforcement Learning-Based Online Task Offloading and Resource Allocation in Harsh Mobile Edge Computing Environment. Clust. Comput. 2023.

43. Song, I.; Tam, P.; Kang, S.; Ros, S.; Kim, S. DRL-Based Backbone SDN Control Methods in UAV-Assisted Networks for Computational Resource Efficiency. Electronics 2023, 12, 2984.

44. Chen, M.; Liu, W.; Wang, T.; Liu, A.; Zeng, Z. Edge Intelligence Computing for Mobile Augmented Reality with Deep Reinforcement Learning Approach. Comput. Netw. 2021, 195, 108186.

45. Tam, P.; Math, S.; Lee, A.; Kim, S. Multi-Agent Deep Q-Networks for Efficient Edge Federated Learning Communications in Software-Defined IoT. Comput. Mater. Contin. 2022, 71, 3319–3335.

46. Ding, Y.; Huang, Y.; Tang, L.; Qin, X.; Jia, Z. Resource Allocation in V2X Communications Based on Multi-Agent Reinforcement Learning with Attention Mechanism. Mathematics 2022, 10, 3415.

47. Sha, D.; Zhao, R. DRL-Based Task Offloading and Resource Allocation in Multi-UAV-MEC Network with SDN. In Proceedings of the 2021 IEEE/CIC International Conference on Communications in China (ICCC) 2021, Xiamen, China, 28–30 July 2021.

48. Zhao, X.; Wu, C.; Le, F. Improving Inter-domain Routing through Multi-agent Reinforcement Learning. In Proceedings of the IEEE INFOCOM 2020—IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS), Toronto, ON, Canada, 6–9 July 2020; pp. 1129–1134.

49. Casas-Velasco, D.M.; Rendon, O.M.C.; da Fonseca, N.L.S. DRSIR: A Deep Reinforcement Learning Approach for Routing in Software-Defined Networking. IEEE Trans. Netw. Serv. Manag. 2021, 19, 4807–4820.

50. Quang, P.T.A.; Hadjadj-Aoul, Y.; Outtagarts, A. A Deep Reinforcement Learning Approach for VNF Forwarding Graph Embedding. IEEE Trans. Netw. Serv. Manag. 2019, 16, 1318–1331.

51. Chen, J.; Chen, J.; Zhang, H. DRL-QOR: Deep Reinforcement Learning-Based QoS/QoE-Aware Adaptive Online Orchestration in NFV-Enabled Networks. IEEE Trans. Netw. Serv. Manag. 2021, 18, 1758–1774.

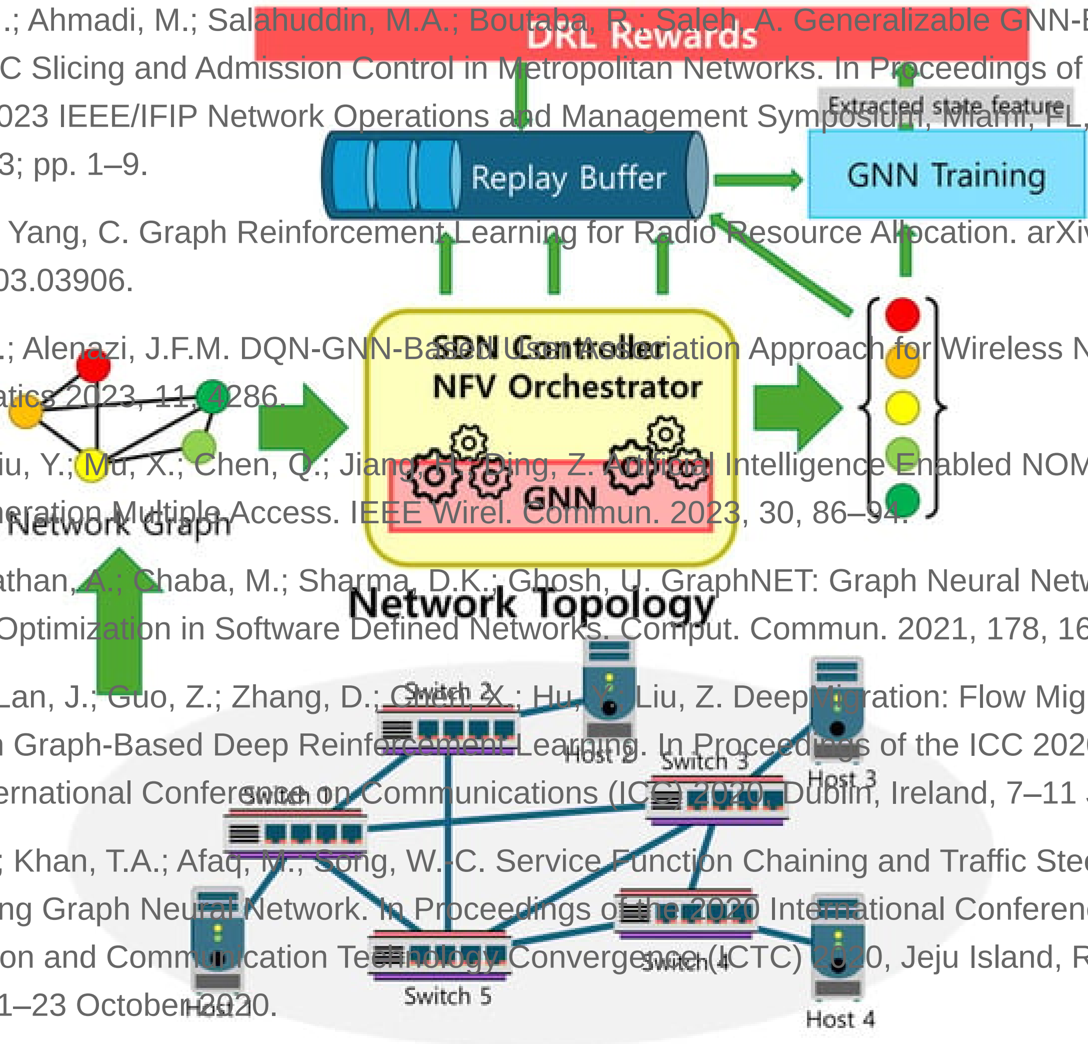

Figure 3. GNN enhances DRL with replay buffer-assisted training in SDN/NFV.

4.3.2. Adaptive SFC

Hara et al. [62] critically considered the high-dimensional changes in graph-structured network topology and service demands that handles future massive service chain requests. In SDN/NFV-enabled environment, authors adopted GNN for approximating the q-values within double DQN framework. The model transformed the network by reinterpreting links as nodes. In this transformed network, nodes are connected if their corresponding links in the original network share a common node, which allows the original network's link features to be viewed as node features in the transformed network and leveraging the adjacency matrix for analysis. The authors obtained the enhancing key performance indicators on packet drop reduction, average delay reduction, robustness against network topology changes, and optimal response to various hyperparameter settings. **Figure 4** presents the overview of logical adaptive SFC for slicing applications between high to low-mission-critical.

52. Arash, M.; Ahmadi, M.; Salahuddin, M.A.; Boutaba, R.; Saleh, A. Generalizable GNN-Based 5G RAN/MEC Slicing and Admission Control in Metropolitan Networks. In Proceedings of the NOMS 2023—2023 IEEE/IFIP Network Operations and Management Symposium, Miami, FL, USA, 8–12 May 2023; pp. 1–9.

53. Zhao, J.; Yang, C. Graph Reinforcement Learning for Radio Resource Allocation. arXiv 2022, arXiv:2203.03906.

54. Ibtihal, A.; Alenazi, J.F.M. DQN-GNN-Based User Association Approach for Wireless Networks. Mathematics 2023, 11, 4286.

55. Xu, X.; Liu, Y.; Mu, X.; Chen, Q.; Jiang, H.; Ding, Z. Artificial Intelligence Enabled NOMA toward next Generation Multiple Access. IEEE Wirel. Commun. 2023, 30, 86–94.

56. Swaminathan, A.; Chaba, M.; Sharma, D.K.; Ghosh, U. GraphNET: Graph Neural Networks for Routing Optimization in Software Defined Networks. Comput. Commun. 2021, 178, 169–182.

57. Sun, P.; Lan, J.; Guo, Z.; Zhang, D.; Chen, X.; Hu, Y.; Liu, Z. DeepMigration: Flow Migration for NFV with Graph-Based Deep Reinforcement Learning. In Proceedings of the ICC 2020—2020 IEEE International Conference on Communications (ICC), 2020, Dublin, Ireland, 7–11 June 2020.

58. Rafiq, A.; Khan, T.A.; Afaq, M.; Song, W.-C. Service Function Chaining and Traffic Steering in SDN Using Graph Neural Network. In Proceedings of the 2020 International Conference on Information and Communication Technology Convergence (ICTC) 2020, Jeju Island, Republic of Korea, 21–23 October 2020.

59. Rusek, K.; Suarez-Varela, J.; Almasan, P.; Barlet-Ros, P.; Cabellos-Aparicio, A. RouteNet: Leveraging Graph Neural Networks for Network Modeling and Optimization in SDN. IEEE J. Sel. Areas Commun. 2020, 38, 2260–2270.

60. Liu, Y.; Lu, Y.; Li, X.; Yao, Z.; Zhao, D. On Dynamic Service Function Chain Reconfiguration in IoT Networks. IEEE Internet Things J. 2020, 7, 10969–10984.

61. Sun, P.; Lan, J.; Li, J.; Guo, Z.; Hu, Y. Combining Deep Reinforcement Learning with Graph Neural Networks for Optimal VNF Placement. IEEE Commun. Lett. 2020, 25, 176–180.

62. Hara, T.; Masahiro, S. Capacitated Shortest Path Tour Based Service Chaining Adaptive to Changes of Service Demand and Network Topology. IEEE Trans. Netw. Serv. Manag. 2024, 25, 176–180.

63. Tan, Y.; Liu, J.; Wang, J. 5G End-To-End Slice Embedding Based on Heterogeneous Graph Neural Network and Reinforcement Learning. IEEE Trans. Cogn. Commun. Netw. 2024.

64. Jalodia, N.; Taneja, M.; Davy, A. A Graph Neural Networks Based Framework for Topology-Aware Proactive SLA Management in a Latency Critical NFV Application Use-Case. arXiv 2022, arXiv:2212.00714.

65. Long, D.; Wu, Q.; Fan, Q.; Fan, P.; Li, Z.; Fan, J. A Power Allocation Scheme for MIMO-NOMA and D2D Vehicular Edge Computing Based on Decentralized DRL. Sensors 2023, 23, 3449.

66. He, Z.; Wang, L.; Hao, Y.; Li, G.Y.; Juang, B. Resource Allocation Based on Graph Neural Networks in Vehicular Communications. In Proceedings of the GLOBECOM 2020—2020 IEEE Global Communications Conference, Taipei, Taiwan, 7–11 December 2020.

67. Lu, S.; Liu, S.; Zhu, Y.; Liang, W.; Li, K.; Lu, Y. A DRL-Based Decentralized Computation Offloading Method: An Example of an Intelligent Manufacturing Scenario. IEEE Trans. Ind. Inform. 2023, 19, 9631–9641.

68. Islam, A.; Ismail, M.; Atat, R.; Boyaci, O.; Shannigrahi, S. Software-Defined Network-Based Proactive Routing Strategy in Smart Power Grids Using Graph Neural Network and Reinforcement Learning. e-Prime 2023, 5, 100187.

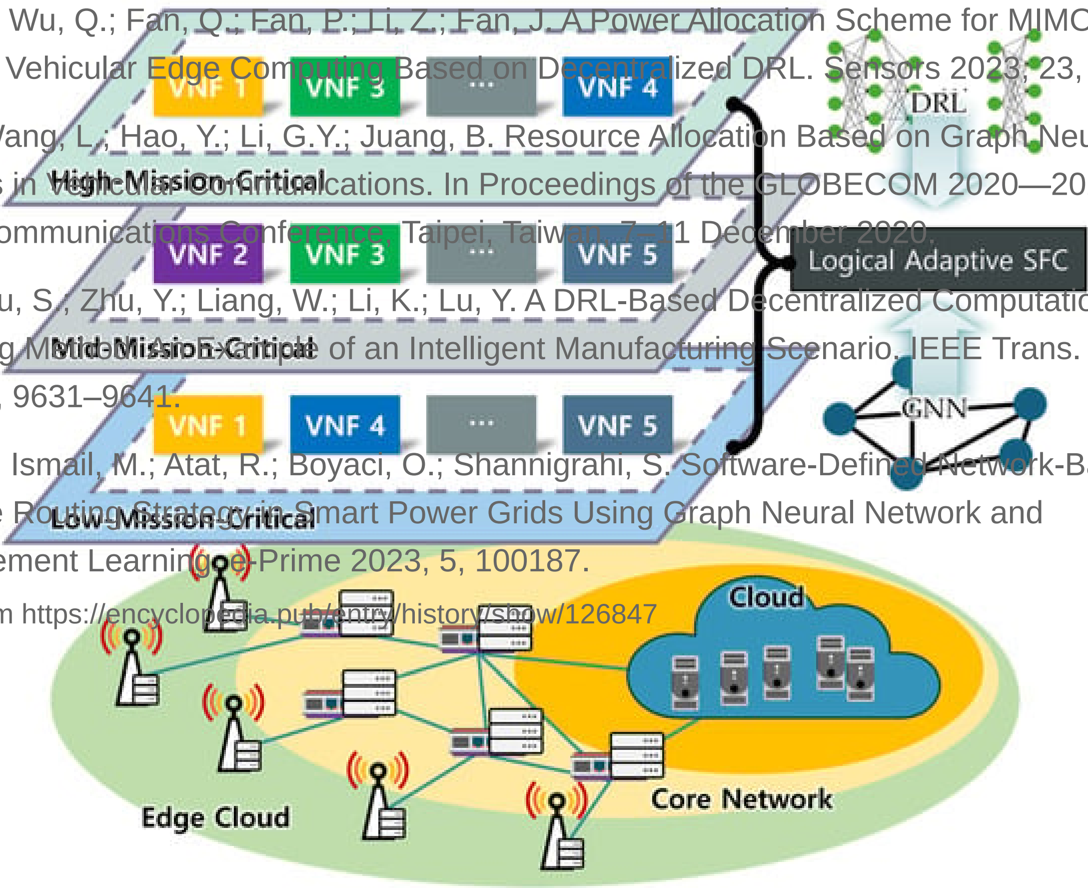

Figure 4. GNN+DRL for orchestrating service chains.

4.3.3. Core Slicing

Tan et al. [63] proposed a novel E2E 5G slice embedding framework that integrates GNN+DRL, primarily in core, to dynamically embed network slices. Utilizing a heterogeneous GNN-based encoder, the scheme captured the complex multidimensional embedding environment, including the substrate and slice networks' topologies and their relationships. A dueling network-based decoder with variable output sizes was employed to generate optimal embedding decisions. The system was trained using the dueling double DQN algorithm, namely D3QN, for enhancing the flexibility and efficiency of slice embedding decisions under various traffic conditions and future service requirements. The proposed GNN+DRL integration achieved higher accumulated revenues for mobile network operators (MNOs) with moderate embedding costs. Specifically, authors obtained significant improvements in embedding efficiency and cost-effectiveness, which showcased its potential for practical deployment in 5G and beyond networks.

4.3.4. SLA Management

Jalodia et al. [64] combined graph convolutional recurrent networks for accurate spatio-temporal forecasting of system SLA metrics and deep Q-learning for enforcing dynamic SLA-aware scaling policies. By capturing both spatial and temporal dependencies within the network, the graph convolutional recurrent networks model forecasted potential SLA violations. The deep Q-learning component utilized these forecasts to train on scaling actions, which aimed to optimize for long-term SLA compliance. The proposed approach allowed for proactive management of network resources, while reducing the risk of SLA breaches and enhancing overall network efficiency. The proposed framework achieved a 74.62% improvement in forecasting performance over the baseline approaches, which demonstrated better prediction accuracy for preventing SLA violations.

5. Application Deployment Scenarios

5.1. Smart Transportation

In [65], the authors address the complexity of V2X communications from the perspective of task allocation, which can be processed either locally or by an MEC server. The authors identified communication scenarios as a significant aspect of channel conditions in MIMO-NOMA-based V2I communications. The paper proposed a decentralized DRL approach for power allocation in the vehicular edge computing (VEC) model that enhanced optimal policy of DDPG in terms of power consumption and reward improvement. Furthermore, [66] employed DQN to learn the optimal value for the V2X pair, which considered the agent within the RL framework in terms of action and resource allocation observation.

5.2. Smart Factory

In [67], authors presented a DRL-based decentralized computation offloading method tailored for intelligent manufacturing scenarios. The paper introduced the dual-critic DDPG algorithm that uses two-critic networks to accelerate the convergence process and minimize computational costs in edge computing systems. By implementing a multi-user system model with a single-edge server, the dual-critic DDPG algorithm efficiently addresses computation offloading and resource allocation challenges while demonstrating good performance in reducing system computational costs for intensive tasks in smart factory.

5.3. Smart Grids

GNN+DRL offers significant opportunities to enhance smart grid reliability, efficiency, and sustainability, moving towards more intelligent and resilient energy systems. By pointing out potential challenges (e.g., various QoS levels including periodic fixed scheduling and emergency-driven packets), traditional smart grids struggle with adaptability to massive/congested network conditions and adhere QoS requirements. In [68], the authors discussed an SDN proactive routing solution using GNN for improved traffic prediction. The paper targeted on improving QoS by (1) predicting future network congestion using GNN and (2) dynamically adjusting routing paths and queue service rates through DRL. The proposed method enhanced the smart grid proactivity in handling of regular and emergency data traffic, which showcased an innovative approach to managing network resources and ensuring service delivery under peak and off-peak conditions.

6. Potential Challenges and Future Directions

6.1. Explainable GNN+DRL

While integration offers remarkable potential, granularity and complexity present a significant challenge, particularly, when these models deploy in critical infrastructure, the decision-making hypothesis becomes increasingly concerned and requires deep inspection. The interpretable GNN architectures require further explorations that inherently reveal the reasons behind each flow-level, node-level, and graph-level predictions (including attention mechanisms or layer-wise explanations). Beyond architecture interpretation, future studies should enable or guide users to understand how altering inputs would affect model outputs, which fosters trust and debugging capabilities. Moreover, researchers can extend by developing methods to extract insights from pre-trained models. Addressing explainability is not only ethically necessary but also crucial for regulatory compliance and gaining wider adoption in safety-critical domains. **Figure 5** describes how explainable modelling interacts to stakeholders with understanding interfaces and outputs.

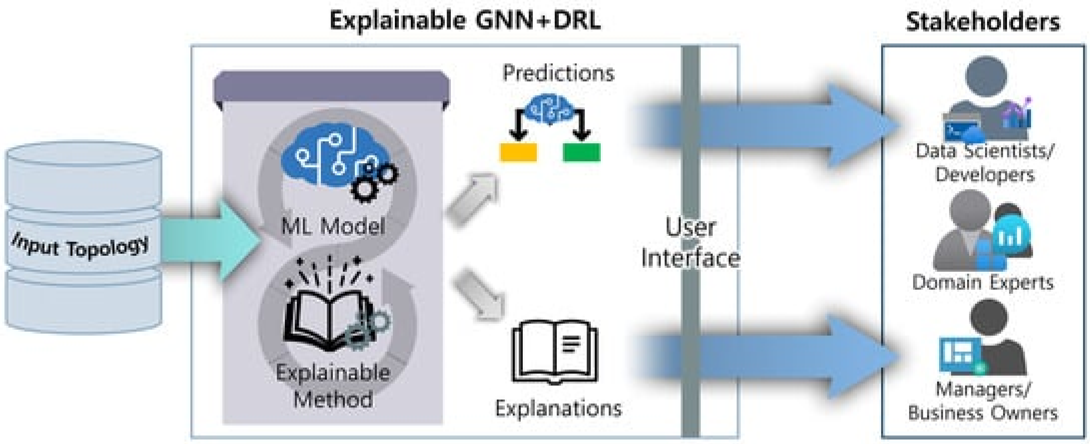

Figure 5. Explainable methods for explaining stakeholders with proper dashboard interfaces.

6.2. Overhead Consumption: Latency, Energy and Computing

The computational demands of GNN+DRL raise concerns about its real-world applicability. Beyond formulating reward functions that jointly consider latency, energy, and computing resources, future research should focus on:

- **Lightweight GNN architectures**, which designs efficient GNNs with reduced parameter counts and computational complexity, potentially leveraging knowledge distillation or pruning techniques.
- **Hardware acceleration**, which explores specialized hardware (e.g., GPUs, TPUs) or hardware-software co-design to accelerate GNN computations and enable (near) real-time capability.
- **Model compression and quantization**, which reduces model size and memory footprint while maintaining accuracy.

6.3. Interoperability with Existing Schemes

Integrating GNN+DRL with existing network infrastructure presents a significant challenge. The key research directions include (1) hybrid approaches, which combines with traditional network protocols and architectures (e.g., SDN, NFV, MEC) for enabling a gradual transition and leveraging existing operations, (2) standardized interfaces, which defines open and adaptable interfaces that allow GNN+DRL models to seamlessly interact with diverse network components and protocols, and (3) backward compatibility, which ensures that new models can work with older systems (minimizing disruption and facilitating wider adoption). **Figure 6** illustrates the overview of interoperating GNN+DRL in existing software-defined and virtualized infrastructures.

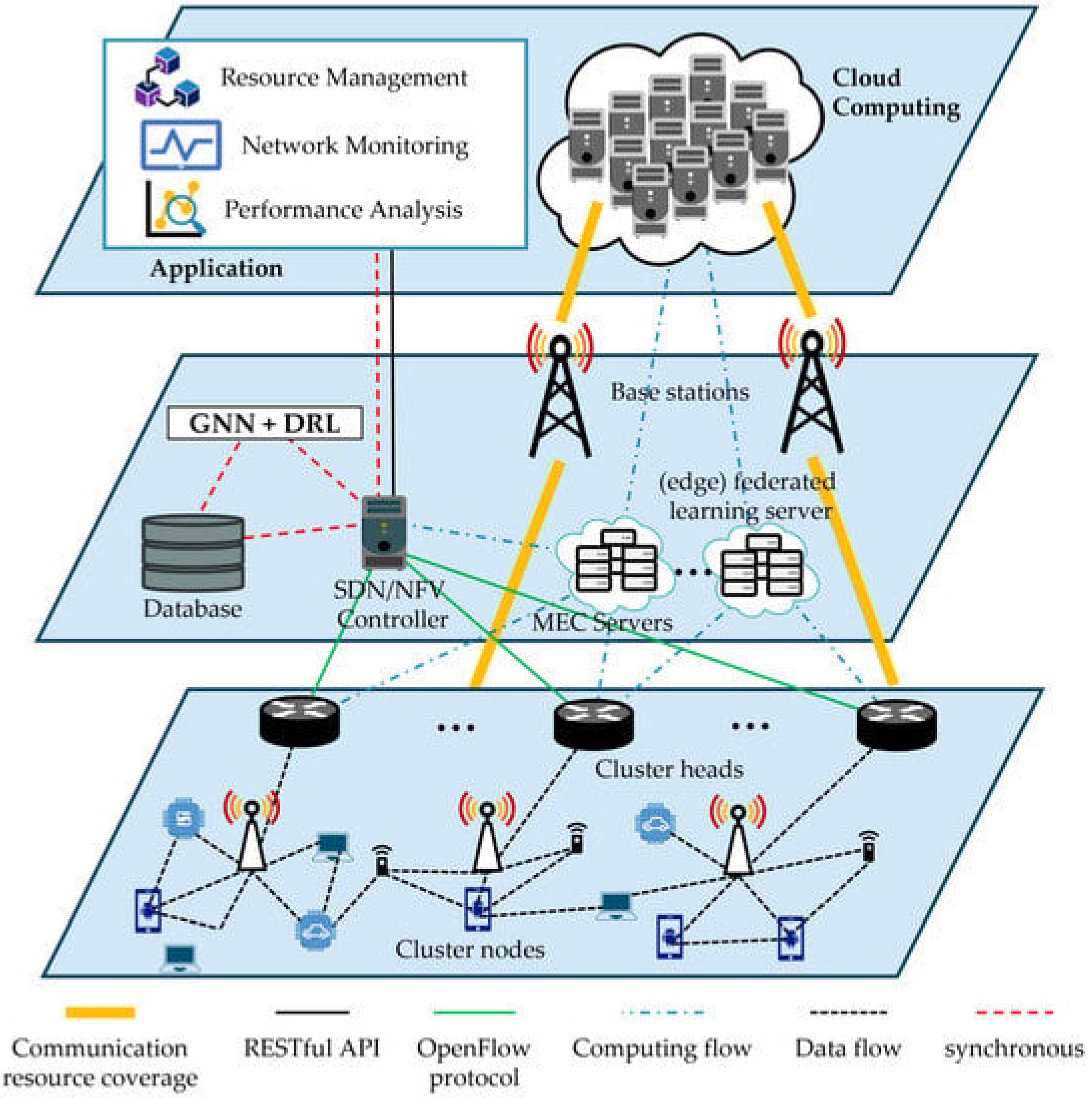

Figure 6. Interoperability of GNN+DRL with SDN, NFV, MEC, and federated learning.

6.4. Reproducibility Awareness

The diverse and complex requirements of future digital networks necessitate robust reproducibility practices in GNN+DRL research. Building a strong foundation of reproducibility is essential for fostering research growth in GNN+DRL and ensuring its practical impact. The key research areas include:

- **Building standardized benchmarks and datasets**, which develop publicly available, well-documented datasets and benchmarks that represent real-world network scenarios; therefore, enabling consistent evaluation and comparison across different studies. Due to a lack of comprehensive studies or data across all domains (access, transport, and core networks), researchers face several issues to conduct the comparison and identify the key metrics to target during experimentation. Different studies may use varied metrics, which making direct comparisons challenging.

- **Code and model sharing**, which encourage open-source code and model sharing to facilitate collaboration, reproducibility, and accelerate research progress.

- **Experimental design guidelines**, which establish best practices for experimental design, data collection, and model evaluation to ensure the validity and generalizability of the research findings.