

Review

A Survey of Intelligent End-to-End Networking Solutions: Integrating Graph Neural Networks and Deep Reinforcement Learning Approaches

Prohim Tam ¹, Seyha Ros ¹, Inseok Song ¹, Seungwoo Kang ¹ and Seokhoon Kim ^{1,2,*}

¹ Department of Software Convergence, Soonchunhyang University, Asan 31538, Republic of Korea; prohimtam@gmail.com (P.T.); rosseyha003@gmail.com (S.R.); sis5041@sch.ac.kr (I.S.); oooksw12@sch.ac.kr (S.K.)

² Department of Computer Software Engineering, Soonchunhyang University, Asan 31538, Republic of Korea

* Correspondence: seokhoon@sch.ac.kr

Abstract: This paper provides a comprehensive survey of the integration of graph neural networks (GNN) and deep reinforcement learning (DRL) in end-to-end (E2E) networking solutions. We delve into the fundamentals of GNN, its variants, and the state-of-the-art applications in communication networking, which reveal the potential to revolutionize access, transport, and core network management policies. This paper further explores DRL capabilities, its variants, and the trending applications in E2E networking, particularly in enhancing dynamic network (re)configurations and resource management. By fusing GNN with DRL, we spotlight novel approaches, ranging from radio access networks to core management and orchestration, across E2E network layers. Deployment scenarios in smart transportation, smart factory, and smart grids demonstrate the practical implications of our survey topic. Lastly, we point out potential challenges and future research directions, including the critical aspects for modelling explainability, the reduction in overhead consumption, interoperability with existing schemes, and the importance of reproducibility. Our survey aims to serve as a roadmap for future developments in E2E networking, guiding through the current landscape, challenges, and prospective breakthroughs in the algorithm modelling toward network automation using GNN and DRL.



Citation: Tam, P.; Ros, S.; Song, I.; Kang, S.; Kim, S. A Survey of Intelligent End-to-End Networking Solutions: Integrating Graph Neural Networks and Deep Reinforcement Learning Approaches. *Electronics* **2024**, *13*, 994. <https://doi.org/10.3390/electronics13050994>

Academic Editor: Gemma Piella

Received: 8 February 2024

Revised: 2 March 2024

Accepted: 4 March 2024

Published: 6 March 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: 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]. Figure 1 presents the ITU prediction outputs. 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. There are several ways of categorizing the domains; however, we prioritize the functionality and specifications from research

perspectives, whether dealing with radio aspects, transmission of data, or critical network functions, in order to support independent technological advancements.

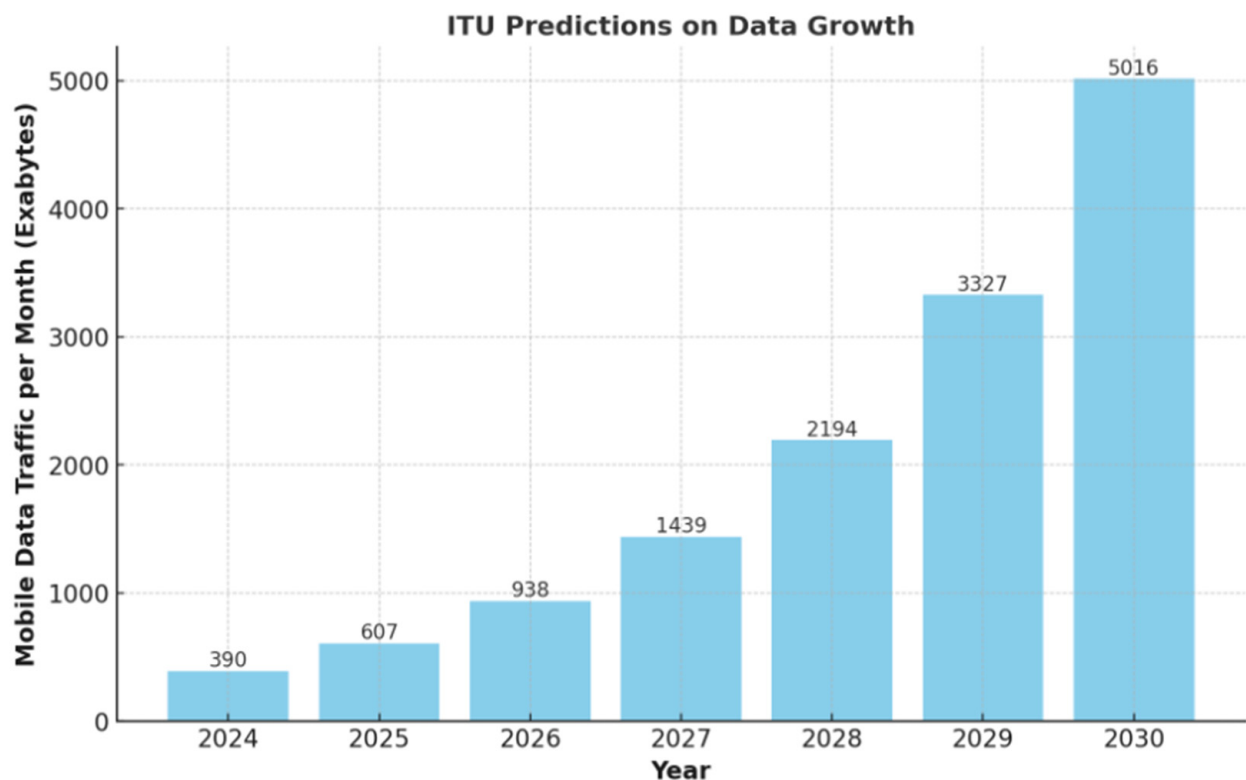


Figure 1. Predictions by ITU on mobile data traffic growth.

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 [11–13]. 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 [16–19]. 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) [22–24] and deep reinforcement learning (DRL) [25–27] 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 [28–31], 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.

Building upon the backgrounds, challenges, and motivations mentioned above, we have compiled a comprehensive review of existing works on GNN+DRL from a communications perspective. Table 1 presents our targeted comparison and the contributions of previous notable literature reviews, as well as how our work adds novelty to this research domain by emphasizing the integration modelling of GNN+DRL algorithms.

Table 1. Comprehensive existing works and our target contributions.

Domain	GNN	DRL	E2E Networking			Ref.	Year
			Access	Core	Transport		
Application of GNNs in wireless networks (resource allocation and emerging fields)	✓	o	✓	o	✓	[22]	2021
Graph-based deep learning for communication networks	✓	o	✓	✓	✓	[23]	2022
GNN in network management and orchestration	✓	o	o	✓	o	[24]	2022
Challenges and opportunities in DRL with GNN	✓	✓	x	x	x	[25]	2023
Applications of DRL in communications and networking	x	✓	✓	✓	✓	[26]	2019
Graph reinforcement learning with its applications and future directions	✓	✓	o	x	x	[27]	2023

✓: primarily included. o: partially included. x: not included.

In our review, we gathered papers from search engines, primarily using Google Scholar, with keywords combining of DRL or/and GNN for searching in networking key terms such as routing optimization, resource management, energy efficiency, access networks, core networks, and transport networks. We found more than 60 papers were published between January 2017 and January 2024. Then, we filtered the key articles (integrating GNN+DRL) to conduct the review in the main section. The remaining articles are analyzed for preliminary sections. Figure 2 presents the paper structure.

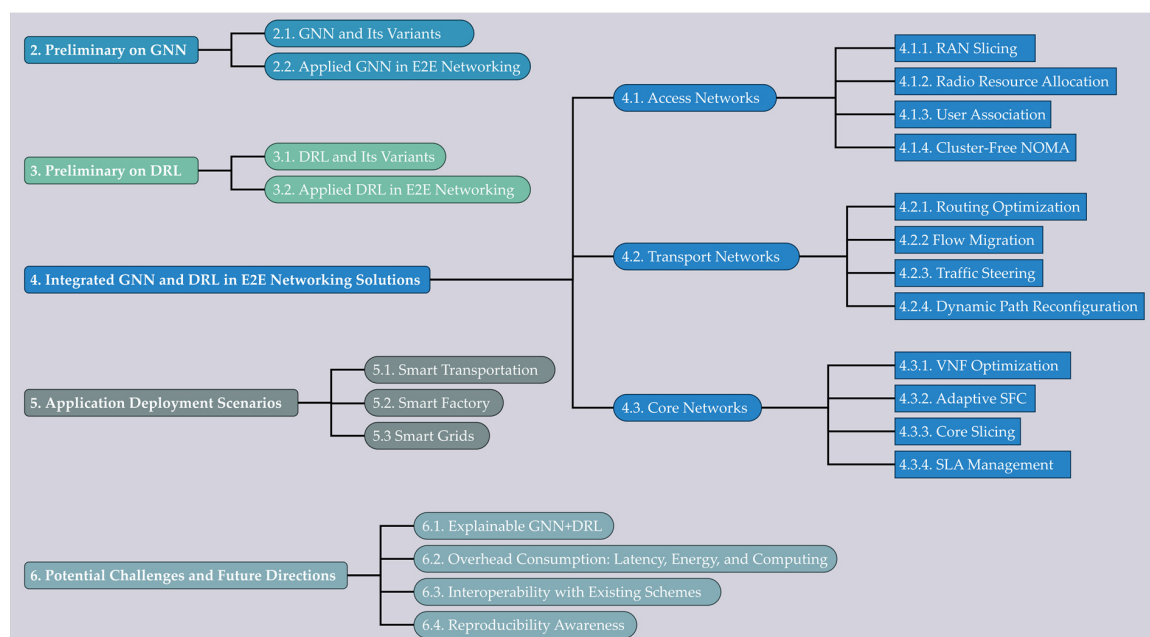


Figure 2. Paper structure.

Our contributions primarily stand with the applications of GNN+DRL in three network domains across access, transport, and core layers. Furthermore, we review the deployment strategies in three use cases, namely smart transportation, smart factory, and smart grids. Given that both GNN+DRL are in the early stages of development and are primarily explored through theoretical research perspectives, we aim to point out the potential challenges and future directions. Our review targets understanding the current limitations and envisioning the roadmap for advancing these paradigm integrations in practical applications and innovative solutions for (future) massive data traffic.

Our review is structured as follows. Sections 2 and 3 present the background studies of both GNN and DRL (its variants and applications). Section 4 provides the integrated GNN+DRL approaches for E2E networking solutions. Section 5 showcases the deployment scenarios in smart applications. Section 6 highlights the research challenges and future research directions. Finally, Section 7 concludes our literature summary. Table 2 presents the key acronyms used in this paper.

Table 2. List of acronyms.

Acronym	Description	Acronym	Description
CNN	Convolutional neural networks	MPTCP	Multi-path TCP
CSI	Channel state information	NFV	Network functions virtualization
DDPG	Deep deterministic policy gradient	NOMA	Non-orthogonal multiple access
DNN	Deep neural networks	PPO	Proximal policy optimization
DQN	Deep Q-networks	QoS	Quality of service
DRL	Deep reinforcement learning	QoE	Quality of experience
GAT	Graph attention networks	RAN	Radio access networks
GCN	Graph convolutional networks	SDN	Software-defined networking
GNN	Graph neural networks	SFC	Service function chaining
ITU	International Telecommunication Union	SLA	Service level agreement
MEC	Mobile edge computing	UAV	Unmanned aerial vehicles
MLP	Multi-layer perceptron	VEC	Vehicular edge computing
MPNN	Message passing neural networks	VL	Virtual link

2. Preliminary on 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. This message-passing process involves two main steps:

- **Aggregation:** Each node aggregates its own features with those of its neighbors, creating a unified vector that represents the local network structure. Equation (1) presents an overview of aggregating information from neighbor nodes of i , denoted as $a_i^{(l+1)}$, where (1) $h_j^{(l)}$ is the feature vector of node j at layer l , and (2) $j \in N(i)$ denotes the set of connected neighbors.
- **Update:** The representation of each node is updated by combining its current representation with the aggregated neighboring node representation (often using a neural

network). Equation (2) presents a general GNN layer for combining $h_i^{(l)}$ and $a_i^{(l+1)}$ towards next-layer features $h_i^{(l+1)}$.

$$a_i^{(l+1)} = \text{AGGREGATE}^{(l)}\left(\left\{h_j^{(l)} : j \in N(i)\right\}\right) \quad (1)$$

$$h_i^{(l+1)} = \text{UPDATE}^{(l)}\left(h_i^{(l)}, a_i^{(l+1)}\right) \quad (2)$$

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]. Table 3 presents an overview of the selected GNN implementation in E2E networking. We focus on (1) specifying the networking domains addressed by the authors, which span from access to core network optimization policies, (2) pointing out the primary input graphs and selected features, (3) describing the processing methodology of the proposed (variant) GNN models, and (4) identifying the types of readout (either flow level, node-level or graph-level prediction) with the target output.

Table 3. Selected comprehensive works on applied GNN.

Networking Domains	Input	Process	Readout	Ref.	Year
Access networks: (1) optimizing large-scale radio resource management problems (beamforming design, power allocation, and cooperative beamforming in interference channels) (2) achieving higher sum rates with much shorter computation times	- Directed graph: the wireless networks - Nodes: agents (mobile users or base stations) - Edges: communication channels or interference link - Features: agent- and channel-related parameters	- Using message passing and extending CNN principles to the graphs, while exploiting permutation equivariance to ensure consistent outputs regardless of node orders - Using “wireless channel GCN” for radio resource management tasks	- Graph-level prediction - Output: optimized configurations for power control and beamforming	[40]	2021

Table 3. Cont.

Networking Domains	Input	Process	Readout	Ref.	Year
Access networks: (1) optimizing large-scale radio resource management problems (beamforming design, power allocation, and cooperative beamforming in interference channels) (2) achieving higher sum rates with much shorter computation times	- Directed graph: the wireless networks - Nodes: agents (mobile users or base stations) - Edges: communication channels or interference link - Features: agent- and channel-related parameters	- Using message passing and extending CNN principles to the graphs, while exploiting permutation equivariance to ensure consistent outputs regardless of node orders - Using “wireless channel GCN” for radio resource management tasks	- Graph-level prediction - Output: optimized configurations for power control and beamforming	[40]	2021
Access networks: (1) optimizing resource allocation in wireless IoT (link scheduling and joint channel and power allocation) (2) maximizing the system performances (network capacity and overall power consumption)	- Directed graph: the wireless networks - Nodes: desirable communication links - Edges: harmful interference links - Features: distance, channel information for nodes and edges, and small-scale/large-scale fading effects	- Using CNN to aggregate local graph-structured neighborhood feature information - Using DNN for updating node representations and incorporating both node and edge features	- Node-level prediction - Output: class probabilities indicating the resource allocation decisions for the network, such as which D2D links to activate and the power and channel allocation for those links	[41]	2021
Access networks: (1) optimizing spectrum allocation in vehicular networks (2) maximizing the sum capacity of V2V and V2I links	- Directed graph: the vehicular communication networks - Nodes: V2V pairs - Edges: interference links - Features: channel gain and corresponding transmit power, while edge weights represent the interference channel gain	- Using GNN to extract low-dimensional features of each node (V2V pair) - Updating functions to aggregate information from neighbors and edges	- Node-level prediction - Output: the extracted features of each V2V pair are then used by a distributed DQN to select spectrum allocation for each pair	[42]	2020
Transport networks: (1) enhancing MPTCP connection throughput by optimizing multipath routing (2) leveraging SDN for better path management and cross-layer optimization	- Undirected graph: network topology and subflow information - Nodes: hosts, routers, and switches - Edges: physical transmission links - Features: bandwidth, delay, and packet loss rate, as well as MPTCP connection configurations	- Using MPNN architecture and update mechanisms to aggregate and update the features of links and subflows - Combining RNN for iterative learning and MLP for final output generation	- Graph-level prediction - Output: the expected throughput of specific MPTCP connections, which is used to optimize routing decisions	[43]	2020
Transport networks: (1) providing accurate and scalable predictions of network performance metric (2) addressing limitations of traditional modelling by supporting non-Markovian traffic, multi-queue QoS scheduling policies, and complex traffic models	- Directed graph: network flows - Nodes: network devices, queues, and traffic flows - Edges: physical links connecting devices - Features: traffic volume, link load, and queuing policy	- Employing a custom three-stage message-passing algorithm which represents the nodes and supports features such as multi-queue QoS scheduling policies	- Flow-level prediction - Output: individual flow metrics such as delay, jitter, and packet loss	[44]	2023

Table 3. Cont.

Networking Domains	Input	Process	Readout	Ref.	Year
Core networks: (1) optimizing network slicing management under various topologies and unseen environments (2) facilitating efficient and dynamic resource allocation and performance monitoring in 5G networks	(1) Undirected graph: substrate network, (2) directed graph: each network slice - Nodes: (1) multiple commodity servers, (2) VNFs - Edges: (1) physical links connecting the nodes, (2) VL connecting VNFs - Features: resources (e.g., CPU, memory, disks, GPU), available bandwidth, traffic volume and requirements	- Using GraphSAGE framework for efficient and scalable processing, including sampling, aggregation, and normalization techniques to handle dynamic network slicing scenarios	- Node-level prediction (the latency of individual slices) - Output: predicting E2E latency for each network slice	[45]	2022
Core networks: (1) optimizing the deployment of VNFs across servers (2) minimizing operational expenses while ensuring service requirements (service constraints and the physical network's limitations)	- Undirected graph: physical networks - Nodes: servers (deploying VNFs) and switches (cannot deploy VNFs) - Edges: physical links between nodes with attributes such as maximum bandwidth, available bandwidth, and delay - Features: (1) CPU cores and deployed VNFs, (2) bandwidth and delay	- Utilizing the aggregated information to update the node representations while capturing both the current state of the network and the potential for VNF deployment on each server	- Node-level prediction - Output: the optimal number of VNF instances per server and VNF types	[46]	2020

3. Preliminary on 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 [47]. 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) R_t is the reward at time t , and (4) γ is the discount factor.

$$V(s) = E[R_t + \gamma V(s_{t+1}) | s_t = s] \quad (3)$$

$$Q(s, a) = E \left[R_t + \gamma \max_{a'} Q(s_{t+1}, a') \mid s_t = s, a_t = a \right] \quad (4)$$

Several algorithms and architectures have been developed in DRL to address different challenges, including (1) deep Q-networks (DQN) [48] learns the action-value function with DNN and stabilizes Q-learning by using experience replay with fixed Q-targets, (2) policy gradient methods (e.g., REINFORCE, actor-critic) [49] directly learns the policy function while potentially also learning a value function to assist in the learning process, (3) proximal policy optimization (PPO) [50] improves the stability and efficiency of policy gradient methods with techniques that limit the updates to the policy, and (4) deep deterministic policy gradient (DDPG) [51] combines the ideas of Q-learning with policy gradients to handle continuous action spaces.

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 [52–55]. Table 4 outlines DRL notable studies in E2E networking contexts, including the networking domains, key remarks, state observation, action implementation, and reward targets.

Table 4. 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	[56]	2022
Access networks: (1) optimizing total weighted costs for task offloading and resource allocation in an SDN-enabled Multi-UAV-MEC network	Model-free DRL framework employing Q-learning with enhancements to handle the mixed-integer conditions of task offloading and resource allocation	Global network state including task requests from ground equipment, available UAV resources, and current network configurations	(1) task offloading decisions (local processing or offloading to a UAV) and (2) resource allocation strategies (assigning computation resources to tasks)	The negative weighted sum of task processing delay and energy consumption	[57]	2021
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	[58]	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	[59]	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	[60]	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	[61]	2021

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 3 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.

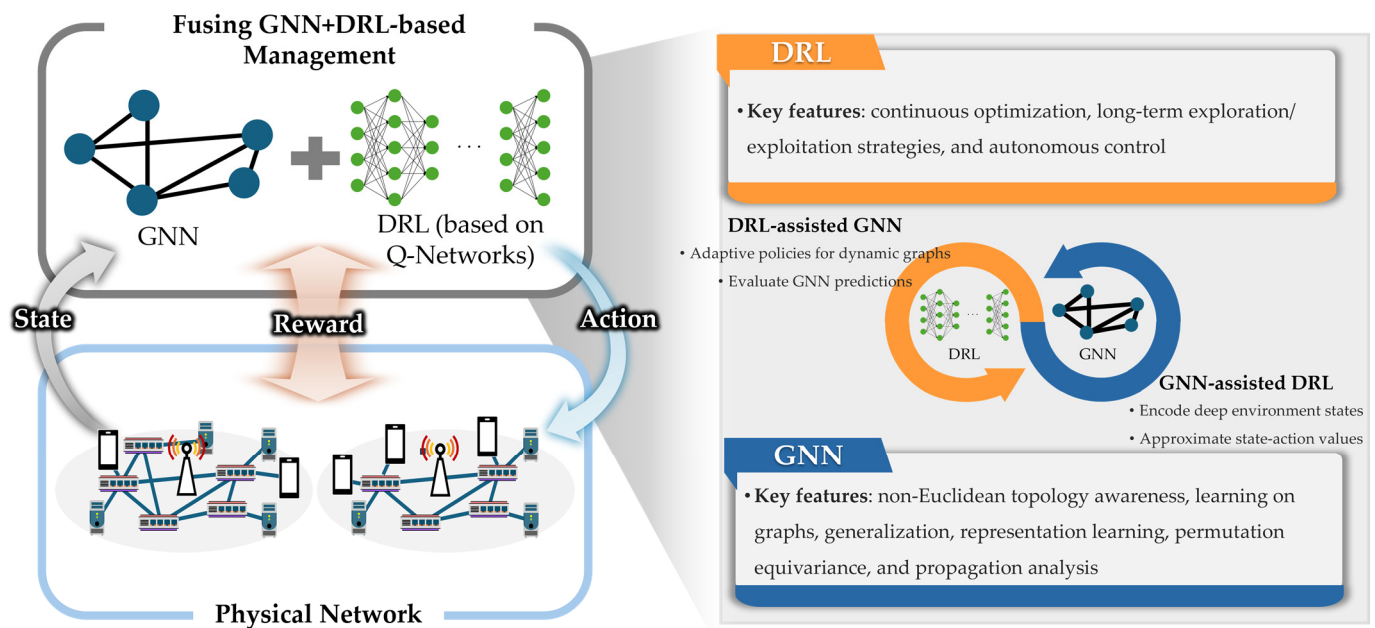


Figure 3. Overview of GNN+DRL and the key features.

Tables 5–7 provide our key literature reviews on each integrated GNN and DRL approaches ranging from access to core networks, as follows:

Table 5. Access networks.

Ref.	State	Action	Reward	DRL	GNN	Key Remarks
[62]	Slice requests, current VNFs and its forwarding graph	Slice request rejection and admission	The successful embedding of all RAN functions	Multi-agent DRL	GAT and GATv2	(1) A slicing agent evaluates service requests, (2) an Admission Control agent subsequently assesses the request based on the slicing decisions, and (3) integrating GNN for network state analysis and DRL for ensuring efficient request processing and revenue optimization
[63]	The topology of the network and attributes related to resource allocation tasks	Decision variables on resource allocations, such as power levels for video streaming or link activations for D2D communications	Minimizing energy consumption while ensuring QoS or to maximize the sum-rate for D2D communications	DDPG	Graph network	(1) Transforming states from traditional matrix-based representations to graph-based ones to utilize GNN, (2) designing GNN to meet permutation invariance and equivalence criteria, and (3) aiming for training complexity reduction and enhancing performances in radio resource allocation tasks
[64]	Service requirements, spectrum availability, transmission power, interference levels, and channel selection (each CU is equipped as an agent)	Power level adjustments, user association, and channel selection	Optimizing power efficiency while considering the interference temperature constraint	DQN	GAT, GCN	(1) Employing a cognitive radio for 6G and beyond networks (coupled network slicing with hybrid spectrum access), (2) developing multi-objective framework for resource allocation, and (3) DQN+GAT for high stability in temporal relation regularization
[65]	User distance distribution and occupied resource information	channel selection and power adaptation (aiming for efficient spectrum sharing and interference mitigation)	Maximizing data rates for secondary users while ensuring the QoS for primary users	DQN	GCN	(1) Integrating of GCN with DRL to form an E2E learning model for resource allocation in cognitive radio environment, and (2) capturing the dynamic topology through a graph-based model (GCN) and directly maps it to resource allocation decisions (DRL)
[66]	User–BS associations and network conditions (signal strength, loading metric, and level of interference)	Specific BS selection for user association	System utility optimization (throughput sums between users)	DQN	Graph network	(1) Representing the user association problem as a graph, (2) encoding it using GNN, and (3) training a DQN agent for optimal BS selection (leveraging both GNN’s spatial modelling and DQN’s decision-making capabilities)
[67]	Network conditions, user demands, channel state information, and traffic patterns	Selection on the optimal successive interference cancellation operations and beamforming vectors	Maximizing system performance metrics (e.g., throughput, spectral efficiency, or energy efficiency) and minimizing interference among users	DQN, DDPG	Graph network	(1) Employing a cluster-free NOMA framework that increases flexibility of operations, (2) executing AutoGNN for reward optimization, and (3) enabling more granular and efficient interference management

Table 6. Transport networks.

Ref.	State	Action	Reward	DRL	GNN	Key Remarks
[68]	Initial hidden states for all links	Next link and node selections for packet routing requests (received by SDN controller)	Packet delivery and delays	DQN	GNN represents DQN (calculating reward)	(1) GNN training on previous DRL episodes, (2) action approximation and reward capturing within GNN, and (3) DQN for loss optimization
[69]	Link features within the network topology and traffic demand (source, destination, required bandwidth)	The weights of each network link that influence routing decisions	Negative value of the load balancing factor that aims for optimal network load distribution	DDPG	MPNN	(1) Leveraging the dynamic network environment and topology information (using a message-passing mechanism to learn and update link weights for optimal routing policies) and (2) enabling effective load balancing and performance improvement in network traffic management
[70]	Network functions feature on processing capabilities, current traffic load instances, migration latency, and queueing length (mapped into a graph representation)	List of migrated traffic loads	The migration cost and resource cost/benefit (with penalties for SLA violations, if not meeting load thresholds)	DQN	Double GNN (current and modified graphs)	(1) Scalable GNN-based function approximator, (2) deep state feature extractions, (3) DRL for policy generation, (4) policy translator module for applying flow migration actions, and (5) less migration cost than [71]
[72]	Network topology, traffic matrix, and routing schemes	Optimal path selection for traffic routing based on predicted delays	Minimizing latency and enhancing resource efficiency	DQN	MPNN	(1) Building SDN/NFV-based system architecture, (2) implementing RouteNet [73] for optimal path prediction, (3) using SDN controller for traffic steering module, and (4) deploying SFC on separate cloud
[74]	Datacenter network topology, resource capacities (e.g., CPU, bandwidth), and requested resources	Selecting optimal data centers for SFC embedding	The success of SFC embeddings, considering both the acceptance ratio of requests and the minimization of embedding costs	PPO (policy gradient DRL)	GCN	Employing a customized GCN model tailored for encoding the network topology and resource states in multi-datacenter networks
[75]	Resource utilization ratios and service demand dynamics to model the SFC dynamic reconfiguration problem	VNF instances migration and flow re-routing	Trade-off between maximizing revenue from served requests and minimizing the total reconfiguration cost	Deep dyna-Q	Graph network	(1) Integrating reinforcement learning with dynamic resource prediction (using state and output computations), (2) building GNN-based prediction of VNF instance resource requirement, and (3) balancing resource utilization and end-service quality

Table 7. Core networks.

Ref.	State	Action	Reward	DRL	GNN	Key Remarks
[76]	The input of the (1) node attributes (processing resources, storage space utilization ratios) and (2) edge attributes (bandwidth utilization ratios, delays)	Defining a list of network nodes for VNF placement (through GNN-based neural networks)	Negative function of the policy's costs, delays, and penalties for not satisfying certain constraints (resource and E2E delays)	REINFORCE	Graph network	(1) SDN controller for state collections and optimal action translation, (2) GNN-based neural network for generalization and action specification, and (3) DRL for iterative process and reward evaluation
[77]	SFC structures, current resource utilization, and VNF components/links	The allocation of resources to VNFs	Minimizing the difference between predicted and actual resource usage	Deep Q-learning	Graph network	(1) Encoding the NFV environment's state into graph-structured processing by GNN, (2) input deep states for DRL training and evaluation, (3) building intelligent dynamic resource prediction module
[78]	Features of physical/virtual link (network and computing resources), SFC- and network-related features (utilization of physical network)	Service path candidate	Network resource utilization (usage degree of physical network links)	Double DQN	GCN	(1) Integrating SDN/NFV-enabled systems for executing massive service chain requests, (2) GNN for computing q-values, and (3) double DQN for adaptability and iterative optimization
[79]	SFC graph topology, VNF deployment status, and network resources	Server selection for VNF deployment	Reducing new server activation and penalizing SLA violations, while minimizing energy use and delay	DDQN	GCN	(1) model the VNF placement as MDP, (2) GCN as DDQN for handling SFC input, and (3) evaluating the resource constraints in each deployed SFC
[80]	Real-time embedding (current substrate network with resources, the slice needed embedding, and existing relationships)	Mapping of one slice node to a selected substrate node (following constraints on locations, security, node capacity, latency, link capacity, etc.)	The embedding success of slices (focusing on the revenue/cost ratio for evaluating embedding efficiency)	Dueling double DQN	GNN-based encoder	(1) Heterogeneous GNN to integrate diverse node and link attributes, (2) enhancing the depth of state information available for DRL agent training, and (3) decoder for handling variable output sizes (align with action space)
[81]	Spatio-temporal data (system and application metrics to forecast potential SLA violations)	Decisions on scaling policies to prevent SLA violations	Optimization of the scaling policy by preventing SLA violations and penalizing unnecessary scaling	Deep Q-learning	Graph convolutional recurrent networks	The integration of graph convolutional recurrent networks for spatio-temporal forecasting and deep Q-learning for dynamic policy enforcement, which significantly outperforming existing models (e.g., residual LSTM) in proactive SLA management

4.1. Access Networks

In this sub-section, we outline the primary contributions to access network policies, focusing on the integration of network topologies as comprehensive graphs for early processing. This approach targets readout objectives through continuous learning capabilities and non-Euclidean feature extraction. Figure 4 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).

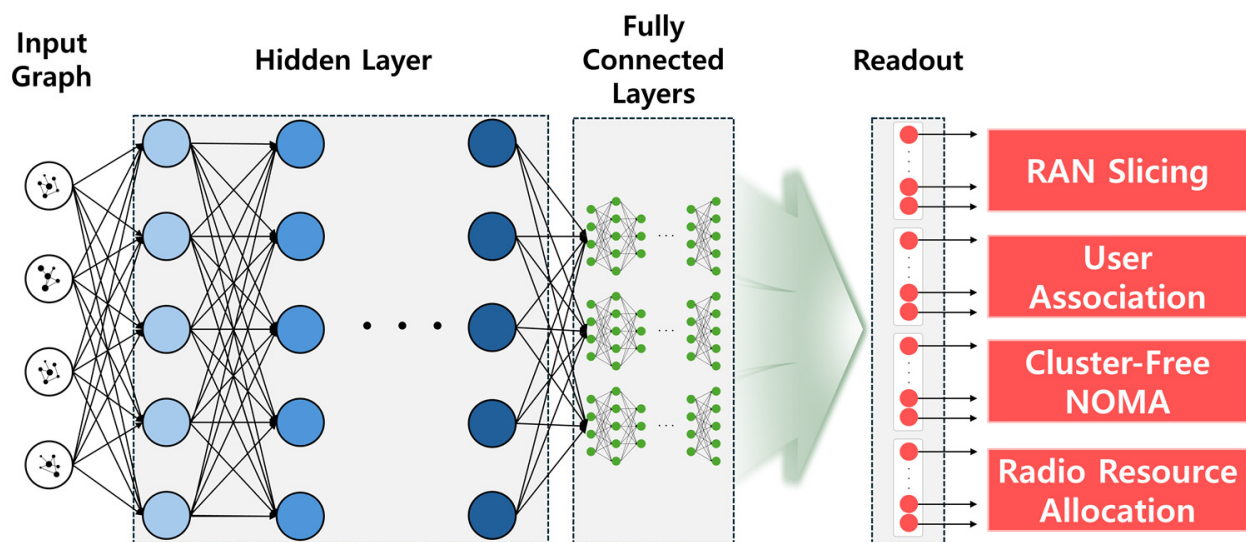


Figure 4. Schematic graph processing from input network graphs towards access network policies.

4.1.1. RAN Slicing

Arash et al. [62] 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 in ensuring fairness, quality of service (QoS), efficiency, and cost-effectiveness in operational expenses. Zhao et al. [63] 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.

Furthermore, Yuan et al. [64] focused on the dynamic assignment of spectrum and power resources in a cognitive radio network that uses both overlay (where secondary users utilize spectrum not used by primary users) and underlay (where secondary users share spectrum with primary users under certain interference constraints) access methods, enhanced by network slicing to support different QoS service requirements. The integrated scheme, termed graph convolution reinforcement learning algorithm, leveraged GCN for agents to efficiently gather and utilize both personal and neighboring information that enhances local agent collaboration. The proposed method enhanced the cognitive network's overall power efficiency.

Beyond this, Zhao et al. [65] utilized GCN for extracting interference features from the graph and combining the features (user distance distribution and resource states) with a DRL approach for decision making in channel state information estimation. The E2E model integrated feature extraction and policy generation. The learning process was guided by a policy gradient method for optimizing channel selection and power adaptation actions to improve spectrum sharing and mitigate interference.

4.1.3. User Association

Ibtihal et al. [66] 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. Finally, the effectiveness of the proposed GNN-DQN approach is tested and evaluated against other user association methods to showcase its ability in adapting and providing efficient solutions in dynamic wireless network environments. Figure 5 shows the layer structure between cognitive radio networks to agent controller in building the graph, mapping, and finalizing the policy control for user association.

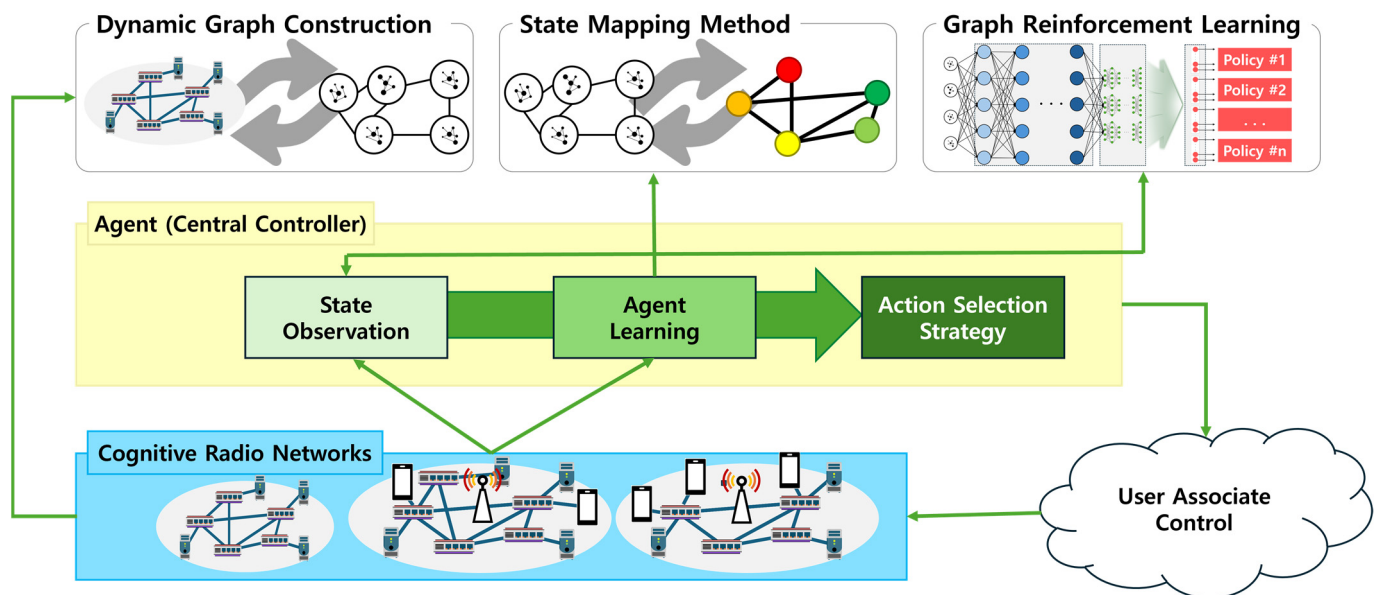


Figure 5. Agent (central controller) for user association control.

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. [67] 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. Notably, while unsupervised centralized convolutional neural networks yield the lowest performance due to the non-structural nature and scalability issues, the structural GNNs achieve system sum rates comparable to centralized optimization methods. AutoGNN achieved higher system sum rates than both the cluster-free NOMA with (1) centralized/distributed alternating direction method of multipliers and (2) the centralized convolutional neural networks approach. Specifically, the system sum rate (bps/Hz) for cluster-free NOMA with AutoGNN is highly efficient, which indicated a better performance enhancement.

4.2. Transport Networks

4.2.1. Routing Optimization

Swaminatha et al. [68] 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. Figure 6 illustrates the interactions between SDN architecture and GNN+DRL by offering (1) state observability by SDN interfaces, (2) computability on SDN databases/controllers, and (3) action configurations translate to SDN forwarding rules.

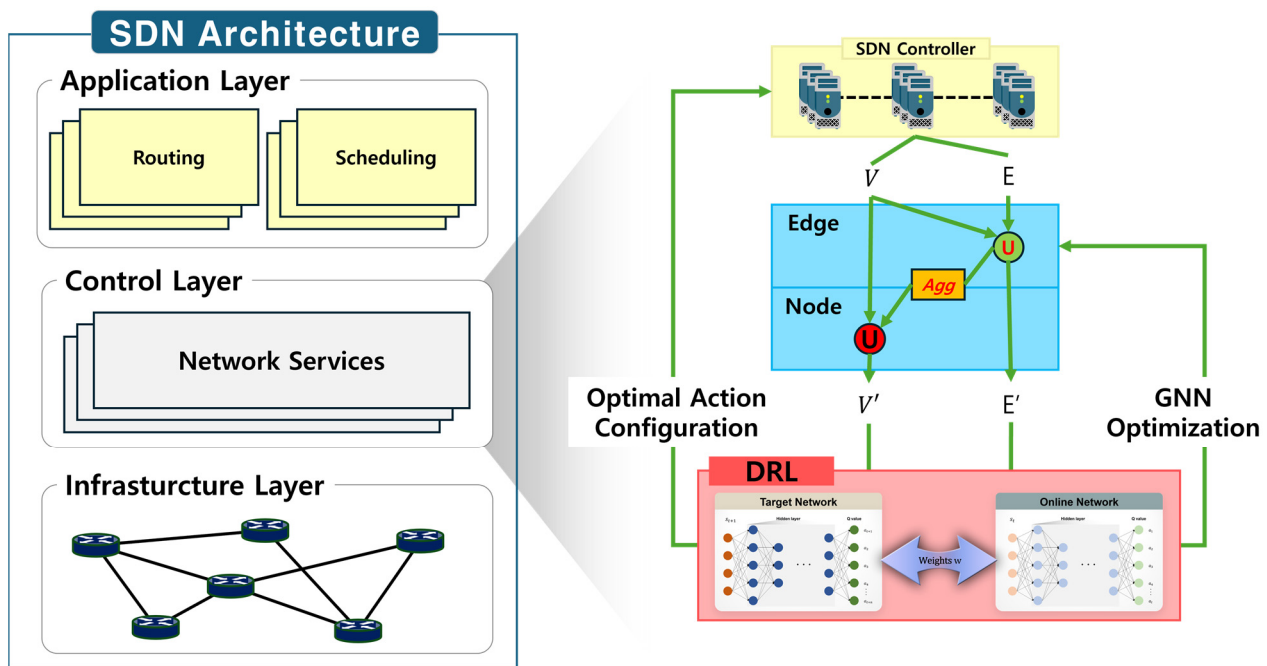


Figure 6. GNN and DRL on SDN architecture for routing optimization.

Furthermore, He et al. [69], leveraged knowledge-defined networking architecture for machine learning programmability to transform network data into actionable knowledge that intelligently guides routing policies for improved network management. The proposed integration, message passing DRL (MPDRL) architecture, utilized GNN and DDPG to address routing optimization by (1) GNN interacts with the network topology to extract knowledge through message passing, while (2) DDPG utilizes the hidden knowledge to generate optimal routing policies. In this study, authors optimized the combination for enhancing network traffic load balancing and overall performance by dynamically adapting to network changes. Through extensive experiments across three real-world internet service provider (ISP) network topologies, the proposed routing optimization method consistently achieved better network performance metrics, such as reducing maximum link utilization, decreasing E2E network delay, and enhancing overall network utility compared to (1) open shortest path first, (2) equal-cost multi-path, and (3) the original DDPG methods.

4.2.2. Flow Migration

Sun et al. [70] 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 significant performance improvements in network functions virtualization (NFV)-enabled flow migration by reducing the costs and saving up to 71.6% of computation times compared to selected baselines. Specifically, in scale-out scenarios, the proposed framework achieved 63.3% less migration cost than OFM [71].

4.2.3. Traffic Steering

In this sub-section, we focus on the intelligent management and routing of network traffic to optimize the deployment of SFC in SDN/NFV-enabled environments. Rafiq et al. [72] integrated RouteNet model [73] 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.

Furthermore, in Xiao et al. [74], SFC mapping was studied by a two-stage DRL framework, which integrated on GCN-based proximal policy optimization to address the embedding problem in multi-datacenter networks. The first stage provided a macro perspective solution by treating all SFCs in a data center's queue as a single entity, which aimed for load balancing policies across multi-datacenters. The early stage involved modelling the load transfer as a MDP to optimize SFC placement for maximizing request acceptance while minimizing costs. The second stage refined the early stage by embedding SFCs within each datacenter's local observation scope, while using a multi-agent reinforcement learning approach to achieve efficient SFC embedding with minimized costs. On average, the framework outperformed Kolin and DQN methods by 13% and 18%, respectively.

4.2.4. Dynamic Path Reconfiguration

Liu et al. [75] 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 7. Sun et al. [76] 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 is utilized to generalize network topology (resource, storage, bandwidth, and tolerable delays).

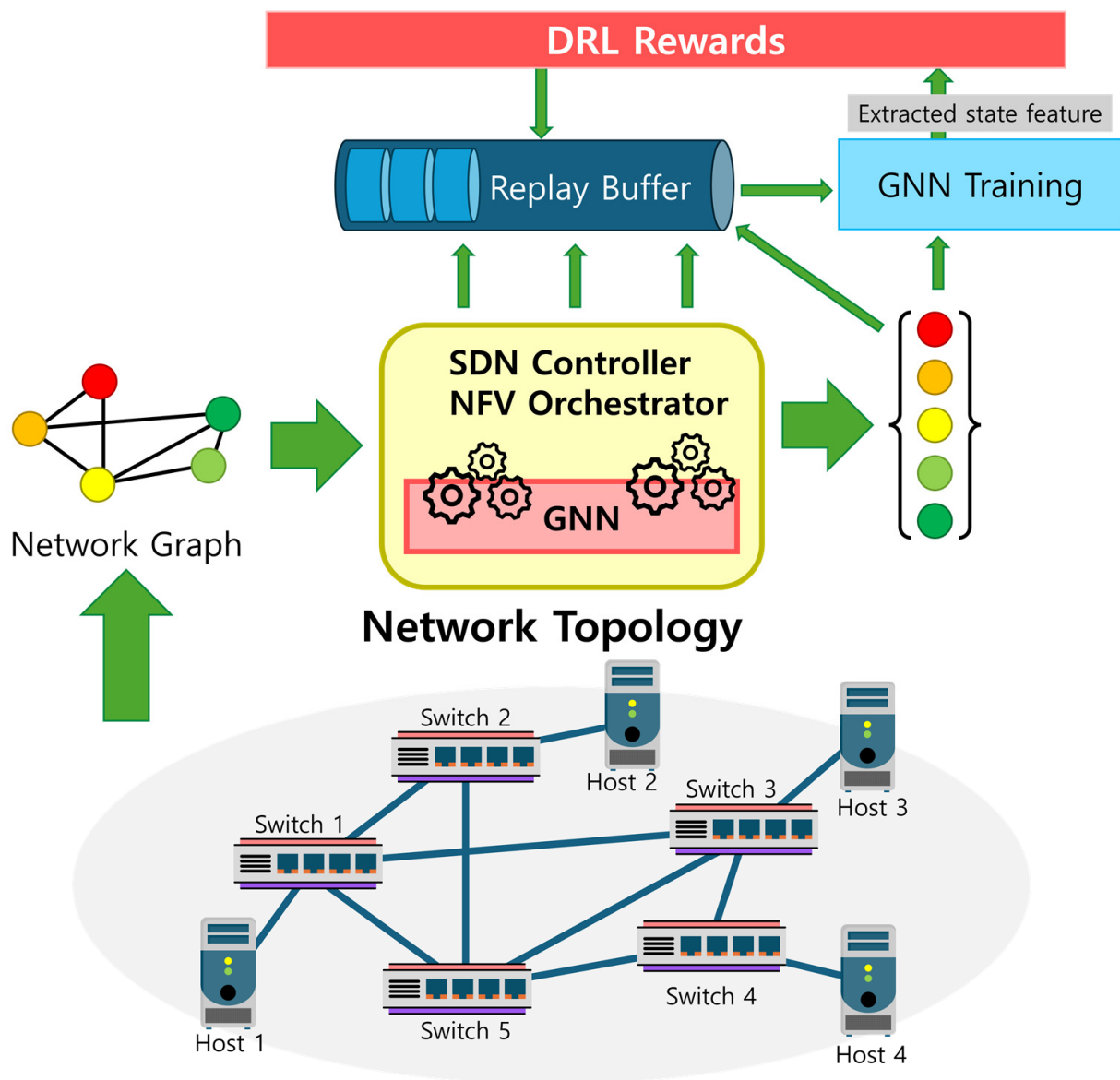


Figure 7. GNN enhances DRL with replay buffer-assisted training in SDN/NFV.

The DRL framework employed the REINFORCE algorithm to optimize the placement strategy by generating actions (node selections) and calculating rewards based on cost, penalties, and delay factors. The proposed approach ensures minimal resource consumption and adherence to QoS constraints, which showcased a low reject ratio of SFC requests at 0.22% compared to other conventional approaches.

In terms of optimization, Jalodia et al. [77] studied resource prediction of VNF with asynchronous DRL enhanced GNN in NFV-enabled system architecture. The authors addressed the complexity of NFV environments by considering the topology of SFC and employed multiple DRL agents to learn optimal resource allocation strategies asynchronously. The proposed approach improved prediction accuracy and operational efficiency by dynamically adjusting resources in real-time.

4.3.2. Adaptive SFC

Hara et al. [78] 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 that 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 8 presents the overview of logical adaptive SFC for slicing applications between high to low-mission-critical.

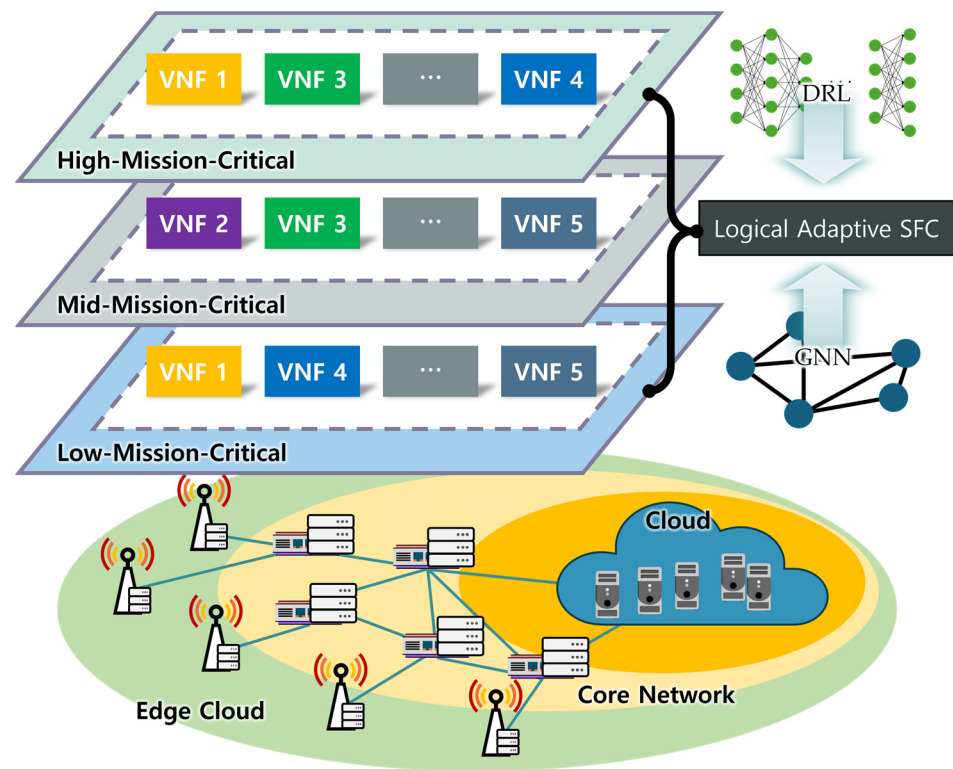


Figure 8. GNN+DRL for orchestrating service chains.

Qi et al. [79] studied graph-structured SFCs and leveraged GCN capabilities to extract the deep hidden states, while representing as Q-networks. By integrating GCN with constrained DDQN for energy-efficient VNF deployment, authors addressed the SFC challenges by optimizing the E2E delay of SFC request with high successful ratios of deploying VNFs. The proposed approach introduced a mask mechanism in DDQN to ensure that resource constraints are met. In experimental results, the approach outperformed traditional DDQN and greedy algorithms, also achieved better performance in handling unseen SFC graphs.

4.3.3. Core Slicing

Tan et al. [80] 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. [81] 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

In this section, we address the collaboration between GNN and DRL to enable efficient technique of applications in three deployment scenarios, namely smart transportation, smart factory, and smart grids, for improving overall autonomy, scalability, and efficiency. Figure 9 illustrates the overview of applied models for efficient smart services.

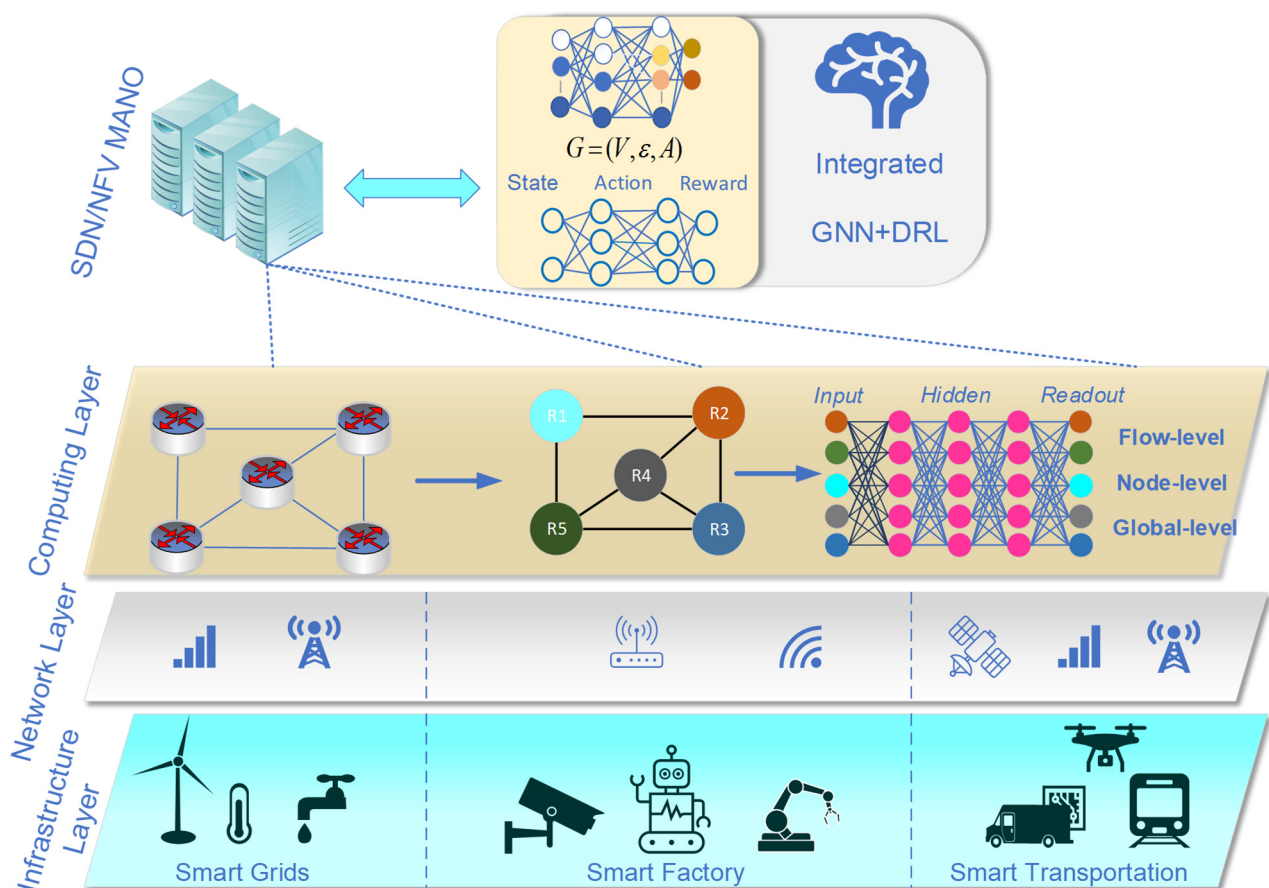


Figure 9. Application deployment on (1) smart transportation, (2) factory, and (3) grids.

5.1. Smart Transportation

In [82], 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, [42] 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. The authors of [60] discussed the utilization of VNF forwarding graph embedding to enable network services while formalizing VNF forwarding graph allocation problems as MDs solved by DRL to simplify the complexity of management and orchestration in telecommunication network service deployment. The study integrated heuristic algorithms and DDPG to enhance exploration in DRL agents.

Beyond that, [83] examined unmanned aerial vehicles (UAV) use in bolstering IoT edge network performance, addressing challenges such as limited computational capacity and energy availability. The authors proposed a multi-agent DRL-based strategy to minimize network computation costs while upholding the QoS of IoT devices. The authors of [84] introduced innovative solutions in IoV by employing a novel DRL method for the vehicle handover process, which aims to reduce handover failures. The paper also suggested a fuzzy-based GNN to navigate the network selection problem by incorporating fuzzy logic into the graph structure. The authors of [85] explored challenges in developing vehicular ad hoc networks and proposed DRL to optimize network caching and computing. Additionally, [86] outlined an efficient path planning scheme for UAVs in data gathering, which addressed scenarios without communication infrastructure by leveraging DRL for planning UAV hover points. A cluster-head searching algorithm with an autonomous exploration pattern was utilized to adapt to fluctuating positions. The authors of [87] focused on enhancing taxi service demand and providers' profits through GNN models, which address scalability challenges by proposing a heterogeneous GNN-LSTM algorithm. Figure 10 illustrates how BS and UAVs can be managed by core controller (using GNN+DRL modelling) to enhance the coverage with fault tolerance in smart transportation.

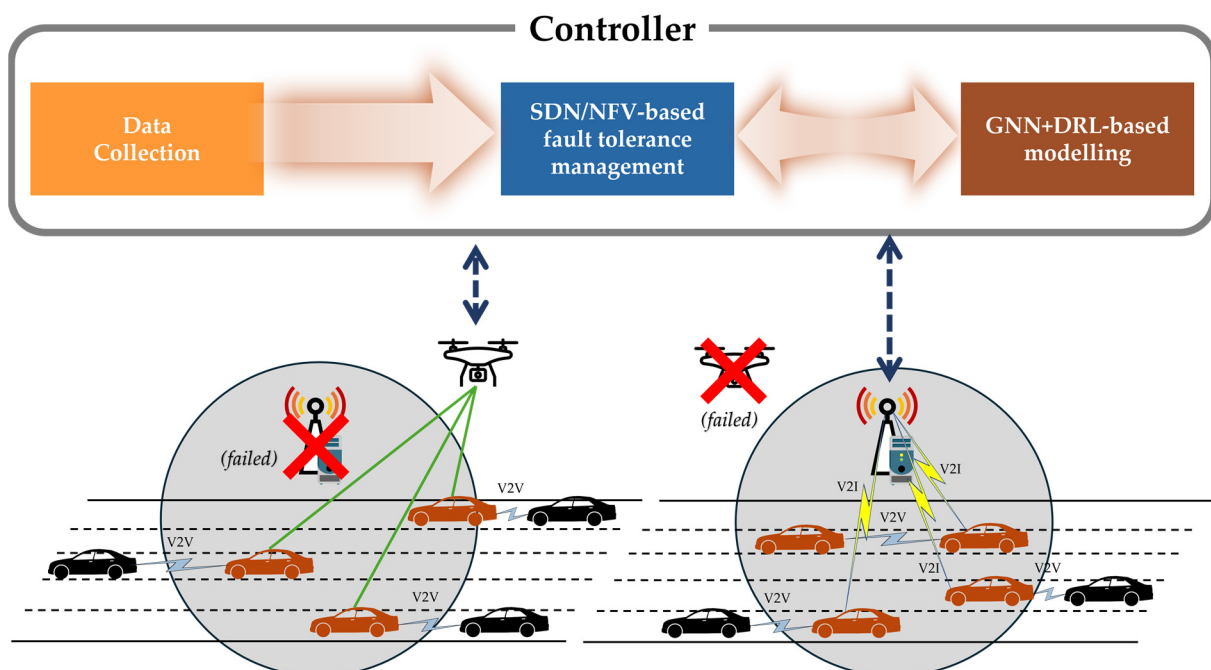


Figure 10. UAV-assisted coverage for fault tolerance in smart transportation.

5.2. Smart Factory

This section highlights the potential of GNNs and DRL in revolutionizing smart factory operations through intelligent decision making, efficient resource utilization, and adaptive control of manufacturing process. In [88], 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.

In [89], authors presented an interesting software-defined factory architecture with DRL-based QoS optimization. Double DQN approach was proposed to manage network load balance and dynamic traffic scheduling while meeting low-latency requirements. The method analyzed QoS threshold and focused on latency/bandwidth by utilizing double DQN to find optimal data flow paths. The proposed architecture included layers for interoperating heterogeneous networks, SDN centralized control, and DRL agents, which demonstrated improved latency, jitter, and throughput that offered a smart solution for dynamic traffic management in smart factories.

Moreover, in [90], authors presented a GNN-based approach with the time-series similarity-based GAT for predicting cellular traffic. By utilizing dynamic time warping and graph attention mechanisms, the proposed method effectively captured spatial-temporal relationships in cellular data, which enhanced prediction accuracy for smart factory environments. In [91], authors proposed a GNN-enhanced DQN algorithm for dynamic QoS flow path allocation in smart factory networks while addressing heterogeneous network challenges. The proposed method optimized traffic management by learning network states and allocation strategies, which improved agent learning efficiency with prioritized experience replay under sparse reward conditions. The authors showcased adaptability to network topology changes and autonomous operation 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 [92], 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.

In [93], authors focused on evaluating the cyber layer in systems, such as IEEE 14-bus and IEEE 39-bus test systems, where accurately predicting traffic demands is essential for efficient resource allocation by proposing a GNN and LSTM method for capturing spatial-temporal traffic patterns. By collecting data on 5G transmissions through power private LTE-G networks, including terminal locations and bandwidth requirements, the method used (1) LSTM to model temporal correlations and (2) GNN to capture spatial relationships between BS. The dual approach allows for accurate predictions of BS traffic, which supports efficient resource allocation and network optimization in smart grids.

Moreover, [94] highlighted the integration of green IoT within smart grid systems while addressing resource allocation in RAN slicing and mapping these challenges to DRL for enhanced flexibility and service delivery. And [95] aimed to solve energy efficiency and delay challenges in heterogeneous cellular networks for various task delay requirements, which suggested DRL approach for learning optimal policies based on predicted SINR states for enhancing decision making in executing high successful transmissions. The proposed approach dynamically adjusted power levels and access strategies to balance the trade-off between minimizing energy consumption and meeting the delay constraints

of smart grid data traffic, which demonstrated a practical approach to improving the performance of cellular networks and supporting real-world smart grid applications.

6. Potential Challenges and Future Directions

To the best of our knowledge, we point out these challenges as a potential guide, which we believe to pinpoint the primary directions for advancing GNN+DRL solutions that are not only highly efficient but also trustworthy, adaptable, and verifiable for ensuring the long-term viability in the ever-evolving optimization algorithms of handling future massive data traffic and merging towards network automation solutions. There are four primary challenges and directions as follows.

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 11 describes how explainable modelling interacts to stakeholders with understanding interfaces and outputs.

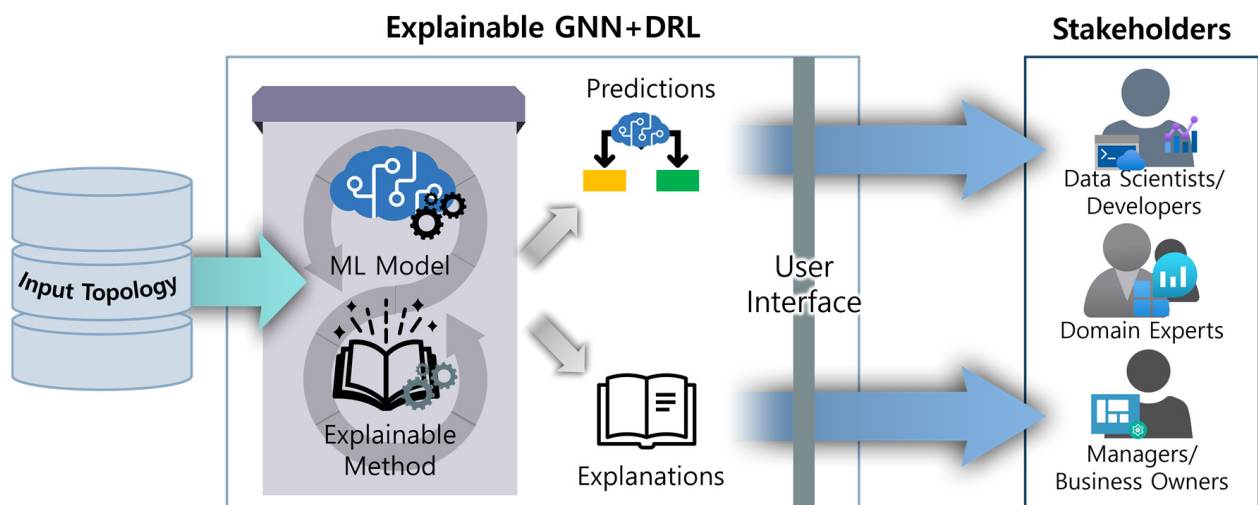


Figure 11. Explainable methods for explaining stakeholders with proper dashboard interfaces.

Therefore, in future studies, there are two essential aspects that can be addressed, including (1) attention mechanisms, which highlights the specific network features influencing GNN outputs that can provide insights into decision rationale, and (2) simulation-based explanations, which illustrates how different network states would affect DRL action selections and reward evaluations.

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. Therefore, we allow the deployment on network architecture of resource-constrained devices. By optimizing GNN+DRL for efficiency, we can obtain the benefits for ensuring its practical viability in latency-sensitive and energy-aware industrial IoT applications.

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 12 illustrates the overview of interoperating GNN+DRL in existing software-defined and virtualized infrastructures.

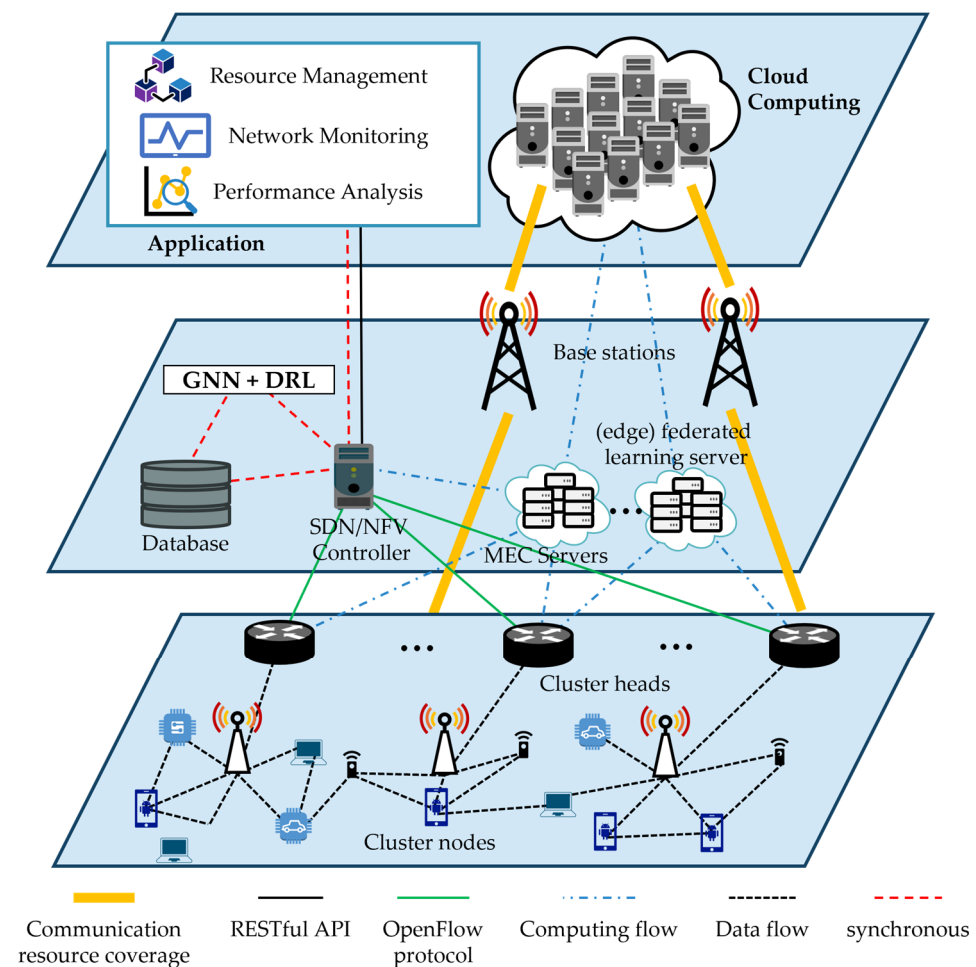


Figure 12. Interoperability of GNN+DRL with SDN, NFV, MEC, and federated learning.

Addressing interoperability is crucial for ensuring a smooth transition towards high-applicable GNN+DRL-powered networks and maximizing the benefits of these paradigms. Researchers can develop common formats for GNN representations that facilitate interoperability with diverse network protocols or build adapter modules that translate between GNN+DRL models and existing network protocols. Beyond these, federated learning can be a good choice for integration to offer collaborative training across different network architectures; therefore, creating interoperable models without compromising each network privacy [96,97].

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.

7. Conclusions

Our survey has delved into the integrating potential of GNN and DRL within E2E networking policies. We have showcased how GNN proficiency in encoding large-scale network topologies and DRL adaptability in continuous learning and agent decisions in order to pave the way for innovative network solutions. From optimizing access networks (RAN slicing, resource allocation, user association, and cluster-free NOMA) to enhancing transport networks (routing optimization, flow migration, traffic steering, and dynamic path reconfiguration), and improving core networks (VNF optimization, adaptive SFC, core slicing, and SLA management), the integrated GNN and DRL framework stands as one of the key backbone algorithms for activating zero-touch network automation. Deployments in smart transportation, smart factory, and smart grids demonstrate the practical benefits of these advancements. However, potential challenges in ensuring explainability, handling overhead consumption, achieving interoperability, and promoting reproducibility awareness appear as frontiers for future exploration. As our conclusion, pointing out these directions will ensure that GNN+DRL solutions remain robust and applicable in the dynamic aspects of future massive network states.

Author Contributions: Conceptualization, P.T. and S.K. (Seokhoon Kim); methodology, S.R. and P.T.; software, P.T. and S.R.; validation, S.K. (Seungwoo Kang), I.S. and S.R.; formal analysis, S.K. (Seungwoo Kang), I.S. and S.R.; investigation, S.K. (Seokhoon Kim); resources, S.K. (Seokhoon Kim); data curation, P.T.; writing—original draft preparation, P.T. and S.R.; writing—review and editing, P.T., S.R., I.S. and S.K. (Seungwoo Kang); visualization, I.S. and S.R.; supervision, S.K. (Seokhoon Kim); project administration, S.K. (Seokhoon Kim); funding acquisition, S.K. (Seokhoon Kim). All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government (MSIT) (No. RS-2022-00167197, Development of Intelligent 5G/6G Infrastructure Technology for The Smart City), in part by the

National Research Foundation of Korea (NRF), Ministry of Education, through Basic Science Research Program under Grant NRF-2020R1I1A3066543, in part by BK21 FOUR (Fostering Outstanding Universities for Research) under Grant 5199990914048, and in part by the Soonchunhyang University Research Fund.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Salh, A.; Audah, L.; Shah, N.S.M.; Alhammedi, A.; Abdullah, Q.; Kim, Y.H.; Al-Gailani, S.A.; Hamzah, S.A.; Esmail, B.A.F.; Almohammed, A.A. A Survey on Deep Learning for Ultra-Reliable and Low-Latency Communications Challenges on 6G Wireless Systems. *IEEE Access* **2021**, *9*, 55098–55131. [\[CrossRef\]](#)
- 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.
- 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).
- Yu, J.-H.; Zhou, Z.-M. Components and Development in Big Data System: A Survey. *J. Electron. Sci. Technol.* **2019**, *17*, 51–72.
- 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. [\[CrossRef\]](#)
- 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. [\[CrossRef\]](#)
- 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. [\[CrossRef\]](#)
- Li, G. Optimal Power Allocation for NOMA-Based Cellular Two-Way Relaying. *KSII Trans. Internet Inf. Syst.* **2023**, *17*, 202–215.
- 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.
- Mangipudi, P.K.; McNair, J. SDN Enabled Mobility Management in Multi Radio Access Technology 5G Networks: A Survey. *arXiv* **2023**, arXiv:2304.03346.
- 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.
- 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.
- 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.
- 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.
- 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. [\[CrossRef\]](#)
- 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.
- GENI Testbed. Available online: <https://github.com/GENI-NSF> (accessed on 2 February 2024).
- 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. [\[CrossRef\]](#)
- Rajab, M.E.; Yang, L.; Shami, A. Zero-Touch Networks: Towards Next-Generation Network Automation. *arXiv* **2024**, arXiv:2312.04159. [\[CrossRef\]](#)
- Mehmood, K.; Kravetska, K.; Palma, D. Intent-Driven Autonomous Network and Service Management in Future CellularR2 Networks: A Structured Literature Review. *Comput. Netw.* **2022**, *220*, 109477. [\[CrossRef\]](#)
- 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. [\[CrossRef\]](#)
- 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. [\[CrossRef\]](#)
- Jiang, W. Graph-Based Deep Learning for Communication Networks: A Survey. *Comput. Commun.* **2022**, *185*, 40–54. [\[CrossRef\]](#)
- 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. [\[CrossRef\]](#)
- 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.

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. [\[CrossRef\]](#)
27. Nie, M.; Chen, D.; Wang, D. Reinforcement Learning on Graphs: A Survey. *IEEE Trans. Emerg. Top. Comput. Intell.* **2023**, *7*, 1065–1082. [\[CrossRef\]](#)
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. [\[CrossRef\]](#)
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. [\[CrossRef\]](#)
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. [\[CrossRef\]](#) [\[PubMed\]](#)
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. [\[CrossRef\]](#)
35. Zhang, S.; Tong, H.; Xu, J.; Maciejewski, R. Graph Convolutional Networks: A Comprehensive Review. *Comput. Soc. Netw.* **2019**, *6*, 11. [\[CrossRef\]](#)
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. [\[CrossRef\]](#)
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.
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. [\[CrossRef\]](#)
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. [\[CrossRef\]](#)
41. Chen, T.; Zhang, X.; You, M.; Zheng, G.; Lambotaran, S. A GNN Based Supervised Learning Framework for Resource Allocation in Wireless IoT Networks. *IEEE Internet Things J.* **2021**, *9*, 1712–1724. [\[CrossRef\]](#)
42. 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.
43. Zhu, T.; Chen, X.; Chen, L.; Wang, W.; Wei, G. GCLR: GNN-Based Cross Layer Optimization for Multipath TCP by Routing. *IEEE Access* **2020**, *8*, 17060–17070. [\[CrossRef\]](#)
44. Ferriol-Galmés, M.; Paillisse, J.; Suárez-Varela, J.; Rusek, K.; Xiao, S.; Shi, X.; Cheng, X.; Barlet-Ros, P.; Cabellos-Aparicio, A. RouteNet-Fermi: Network Modelling with Graph Neural Networks. *IEEE ACM Trans. Netw.* **2023**, *31*, 3080–3095. [\[CrossRef\]](#)
45. Wang, H.; Wu, Y.; Min, G.; Miao, W. A Graph Neural Network-Based Digital Twin for Network Slicing Management. *IEEE Trans. Ind. Inform.* **2022**, *18*, 1367–1376. [\[CrossRef\]](#)
46. Kim, H.-G.; Park, S.; Heo, D.; Lange, S.; Choi, H.; Yoo, J.-H.; Hong, J.W.-K. Graph Neural Network-Based Virtual Network Function Deployment Prediction. In Proceedings of the 2020 16th International Conference on Network and Service Management (CNSM), Izmir, Turkey, 2–6 November 2020.
47. Li, Y. Deep Reinforcement Learning: An Overview. *arXiv* **2017**, arXiv:1701.07274.
48. Huang, Y.H. Deep Q-Networks. In *Deep Reinforcement Learning: Fundamentals, Research and Applications*; Springer: Berlin/Heidelberg, Germany, 2020; pp. 135–160.
49. Agarwal, A.; Kakade, S.M.; Lee, J.D.; Mahajan, G. On the theory of policy gradient methods: Optimality, approximation, and distribution shift. *J. Mach. Learn. Res.* **2021**, *22*, 4431–4506.
50. Schulman, J.; Wolski, F.; Dhariwal, P.; Radford, A.; Klimov, O. Proximal Policy Optimization Algorithms. *arXiv* **2017**, arXiv:1707.06347.
51. Tan, H. Reinforcement Learning with Deep Deterministic Policy Gradient. In Proceedings of the 2021 International Conference on Artificial Intelligence, Big Data and Algorithms (CAIBDA), Xi'an, China, 28–30 May 2021; pp. 82–85.
52. Xiang, H.; Zhang, M.; Jian, C. Federated Deep Reinforcement Learning-Based Online Task Offloading and Resource Allocation in Harsh Mobile Edge Computing Environment. *Clust. Comput.* **2023**. [\[CrossRef\]](#)
53. 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. [\[CrossRef\]](#)
54. 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. [\[CrossRef\]](#)

55. 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. [\[CrossRef\]](#)
56. 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. [\[CrossRef\]](#)
57. 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.
58. 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.
59. 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. [\[CrossRef\]](#)
60. 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. [\[CrossRef\]](#)
61. 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. [\[CrossRef\]](#)
62. 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.
63. Zhao, J.; Yang, C. Graph Reinforcement Learning for Radio Resource Allocation. *arXiv* **2022**, arXiv:2203.03906.
64. Yuan, S.; Zhang, Y.; Ma, T.; Cheng, Z.; Guo, D. Graph Convolutional Reinforcement Learning for Resource Allocation in Hybrid Overlay–Underlay Cognitive Radio Network with Network Slicing. *IET Commun.* **2022**, *17*, 215–227. [\[CrossRef\]](#)
65. Zhao, D.; Qin, H.; Song, B.; Han, B.; Du, X.; Guizani, M. A Graph Convolutional Network-Based Deep Reinforcement Learning Approach for Resource Allocation in a Cognitive Radio Network. *Sensors* **2020**, *20*, 5216. [\[CrossRef\]](#)
66. Ibtihal, A.; Alenazi, J.F.M. DQN-GNN-Based User Association Approach for Wireless Networks. *Mathematics* **2023**, *11*, 4286.
67. 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. [\[CrossRef\]](#)
68. 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. [\[CrossRef\]](#)
69. 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.* **2023**, *23*, 1444–1455. [\[CrossRef\]](#)
70. 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.
71. Sun, C.; Bi, J.; Meng, Z.; Yang, T.; Zhang, X.; Hu, H. Enabling NFV Elasticity Control with Optimized Flow Migration. *IEEE J. Sel. Areas Commun.* **2018**, *36*, 2288–2303. [\[CrossRef\]](#)
72. 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.
73. Rusek, K.; Suarez-Varela, J.; Almasan, P.; Barlet-Ros, P.; Cabellos-Aparicio, A. RouteNet: Leveraging Graph Neural Networks for Network Modelling and Optimization in SDN. *IEEE J. Sel. Areas Commun.* **2020**, *38*, 2260–2270. [\[CrossRef\]](#)
74. Xiao, D.; Zhang, A.; Liu, X.; Qu, Y.; Ni, W.; Liu, R.P. A Two-Stage GCN-Based Deep Reinforcement Learning Framework for SFC Embedding in Multi-Datacenter Networks. *IEEE Trans. Netw. Serv. Manag.* **2023**, *20*, 4297–4312. [\[CrossRef\]](#)
75. 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. [\[CrossRef\]](#)
76. 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. [\[CrossRef\]](#)
77. Jalodia, N.; Henna, S.; Davy, A. Deep Reinforcement Learning for Topology-Aware VNF Resource Prediction in NFV Environments. In Proceedings of the 2019 IEEE Conference on Network Function Virtualization and Software Defined Networks (NFV-SDN), Dallas, TX, USA, 12–14 November 2019.
78. 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. [\[CrossRef\]](#)
79. Qi, S.; Li, S.; Lin, S.; Saidi, M.Y.; Chen, K. Energy-Efficient VNF Deployment for Graph-Structured SFC Based on Graph Neural Network and Constrained Deep Reinforcement Learning. In Proceedings of the 2021 22nd Asia-Pacific Network Operations and Management Symposium (APNOMS), Tainan, Taiwan, 8–10 September 2021.
80. 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**. [\[CrossRef\]](#)
81. 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.

82. 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. [[CrossRef](#)] [[PubMed](#)]
83. Abegaz, M.S.; Boateng, G.O.; Mareri, B.; Sun, G.; Jiang, W. Multi-Agent DRL for Task Offloading and Resource Allocation in Multi-UAV Enabled IoT Edge Network. *IEEE Trans. Netw. Serv. Manag.* **2021**, *18*, 4531–4547.
84. Kumar, P.P.; Sagar, K. Reinforcement Learning and Neuro-Fuzzy GNN-Based Vertical Handover Decision on Internet of Vehicles. *Concurr. Comput. Pract. Exp.* **2023**, *35*, e7688. [[CrossRef](#)]
85. He, Y.; Yu, F.R.; Zhao, N.; Yin, H.; Boukerche, A. Deep Reinforcement Learning (DRL)-Based Resource Management in Software-Defined and Virtualized Vehicular Ad Hoc Networks. In Proceedings of the 6th ACM Symposium on Development and Analysis of Intelligent Vehicular Networks and Applications—DIVANet '17, Miami, FL, USA, 21–25 November 2017; pp. 47–54.
86. Liu, R.; Qu, Z.; Huang, G.; Dong, M.; Wang, T.; Zhang, S.; Liu, A. DRL-UTPS: DRL-Based Trajectory Planning for Unmanned Aerial Vehicles for Data Collection in Dynamic IoT Network. *IEEE Trans. Intell. Veh.* **2022**, *8*, 1204–1218. [[CrossRef](#)]
87. Nazzal, M.; Khreishah, A.; Lee, J.; Angizi, S. Semi-Decentralized Inference in Heterogeneous Graph Neural Networks for Traffic Demand Forecasting: An Edge-Computing Approach. *arXiv* **2023**, arXiv:2303.00524. [[CrossRef](#)]
88. 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. [[CrossRef](#)]
89. Xia, D.; Wan, J.; Xu, P.; Tan, J. Deep Reinforcement Learning-Based QoS Optimization for Software-Defined Factory Heterogeneous Networks. *IEEE Trans. Netw. Serv. Manag.* **2022**, *19*, 4058–4068. [[CrossRef](#)]
90. Wang, Z.; Hu, J.; Min, G.; Zhao, Z.; Chang, Z.; Wang, Z. Spatial-Temporal Cellular Traffic Prediction for 5G and Beyond: A Graph Neural Networks-Based Approach. *IEEE Trans. Ind. Inform.* **2022**, *19*, 1–10. [[CrossRef](#)]
91. Guo, Q.; Jin, Q.; Liu, Z.; Luo, M.; Chen, L.; Dou, Z.; Diao, X. Research on QoS Flow Path Intelligent Allocation of Multi-Services in 5G and Industrial SDN Heterogeneous Network for Smart Factory. *Sustainability* **2023**, *15*, 11847. [[CrossRef](#)]
92. 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. [[CrossRef](#)]
93. Zhong, L.; Tang, J.; Xu, C.; Ren, B.; Du, B.; Huang, Z. Traffic Prediction of Converged Network for Smart Grid Based on GNN and LSTM. In Proceedings of the 2022 3rd International Conference on Big Data, Artificial Intelligence and Internet of Things Engineering (ICBAIE), Xi'an, China, 15–17 July 2022.
94. Meng, S.; Wang, Z.; Ding, H.; Wu, S.; Li, X.; Zhao, P.; Zhu, C.; Wang, X.Z. RAN Slice Strategy Based on Deep Reinforcement Learning for Smart Grid. In Proceedings of the 2019 Computing, Communications and IoT Applications (ComComAp), Shenzhen, China, 26–28 October 2019.
95. Abdullah, A.F.; Bu, S.; Valente, K.P.; Imran, M.A. Channel Access and Power Control for Energy-Efficient Delay-Aware Heterogeneous Cellular Networks for Smart Grid Communications Using Deep Reinforcement Learning. *IEEE Access* **2019**, *7*, 133474–133484. [[CrossRef](#)]
96. Chen, F.; Li, P.; Miyazaki, T.; Wu, C. FedGraph: Federated Graph Learning with Intelligent Sampling. *IEEE Trans. Parallel Distrib. Syst.* **2022**, *33*, 1775–1786. [[CrossRef](#)]
97. Kang, S.; Ros, S.; Song, I.; Tam, P.; Math, S.; Kim, S. Real-Time Prediction Algorithm for Intelligent Edge Networks with Federated Learning-Based Modeling. *Comput. Mater. Contin.* **2023**, *77*, 1967–1983. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.