

# Chapter 5

## Machine Learning and Deep Reinforcement Learning in Wireless Networks and Communication Applications



Om Prakash, Prabina Pattanayak, Amrita Rai, and Korhan Cengiz

**Abstract** Wireless networks and communication of the future will have to manage an ever-increasing density of mobile users using a wide range of services and apps, as well as a surge in mobile data traffic. Meanwhile, networks are becoming increasingly dense, heterogeneous, decentralized, and ad hoc, encompassing a wide range of network elements. As a result, a number of service goals, such as high throughput and low latency, must be met, as a result, resource allocation must be established and optimized. Traditional service and resource management methods that need complete and perfect system information are inefficient or inapplicable in wireless network environments due to the inherent dynamic and uncertainty. This chapter covers similar research that employs Deep reinforcement learning (DRL) to address various difficulties in 5G networks after first going through the fundamental notions of DRL. Finally, we look at some of the DRL techniques that have been presented to deal with emerging communications and networking difficulties.

**Keywords** Reinforcement learning (RL) · Deep reinforcement learning (DRL) · Deep-Q Networks (DQN) · Q-function · Q-values

---

O. Prakash (✉)

Sri Venkateswara College of Engineering Technology, R V S Nagar, Chittoor, Andhra Pradesh, India

e-mail: [om4096@gmail.com](mailto:om4096@gmail.com)

P. Pattanayak

National Institute Of Technology QQ5R+3WM, NIT Road, Fakiratilla, Silchar, Assam 788010, India

e-mail: [prabina@ece.nits.ac.in](mailto:prabina@ece.nits.ac.in)

A. Rai

GL Bajaj Institute of Technology and Management, Plot No 2, Knowledge Park III, Greater Noida, Uttar Pradesh 201306, India

K. Cengiz

Istinye University, Istanbul, Turkey

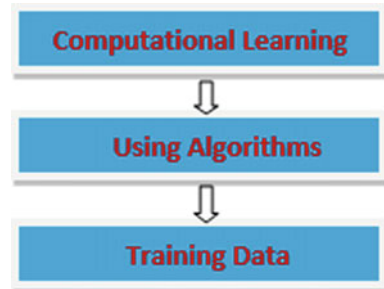
e-mail: [korhan.cengiiz@istinye.edu.tr](mailto:korhan.cengiiz@istinye.edu.tr)

## 5.1 Introduction

Machine learning (ML) is an elegant tool in designing many applications that can serve in the coming future from recommendation on e-commerce websites to content filtering on the social network by giving suggestion for making communication with each other (Qiu et al. 2016; Wang et al. 2021; Ahmed et al. 2022; Gronauer and Diepold 2022). ML techniques include conventional algorithms like Decision Tree and Regression Tree whereas Deep learning (DL) techniques include an automatic encoder and neural network. The basic block diagram of machine learning architecture is shown in Fig. 5.1. Machine learning is applied in advanced business and research for organizations today (Xiang and Foo 2021; Bellemare et al. 2017). It involves algorithms and neural network applications for improving their performance. Machine learning algorithms model mathematical simulations using referred data to take decisions.

In today's world of increasing complexity in mobile network technology, wireless access technologies have been evolving steadily. To ensure the lead for the future, Telematics has introduced 5G wireless technology to the future spectrum of mobile business and is replete with so many development opportunities that significantly improve on the previous generation and offer a whole number of new innovations. The ability to operate correctly on previously unknown inputs is the promise of applying deep learning technology in reinforcement learning. For example, image identification neural networks can recognize a bird in a photograph even if they have never seen the image or the bird. The requirement to predefine the environment is reduced because deep reinforcement learning allows raw data (e.g., pixels) as input, allowing the model to be used in a variety of situations. Deep reinforcement learning (DRL) approaches can be created to be universal because of this degree of abstraction, allowing the same model to be used for multiple tasks. Deep Reinforcement Learning blends reinforcement learning and deep learning to improve the generalization ability of policies developed with deep RL policies (DRL). It has the ability to learn end-to-end learning in robots, video games, natural language processing, computer vision, healthcare, and other domains from raw sensors or images Google Deep Mind in 2013 used Deep-Q Networks (DQN) to play Atari games, which was a watershed moment in value-based DRL.

**Fig. 5.1** Basics of Machine Learning and Deep Learning



Applying profound support learning (DRL) methods to improve remote enterprises' operations is becoming more and more popular. In this research, we examine three state-of-the-art DRL approaches for remote organization streamlining: Deep Deterministic Policy Gradient (DDPG), Neural Episodic Control (NEC), and Variance Based Control (VBC). We illustrate how the overall organizational improvement problem is structured as RL and provide specifics for the three remote systems administration strategies. A verifiable organization activity dataset is used for extensive analyses, and the presentation of the further developing rate and combination speed for these well-known DRL procedures is examined. While DDPG and VBC demonstrate significant promise for computerizing remote organization improvement, NEC has a significantly better union rate yet experiences the restricted activity space and doesn't perform seriously in its ongoing structure (Tanveer et al. 2022).

### ***5.1.1 Deep Learning***

A collection of inputs are transformed into a set of outputs by an artificial neural network in a process known as deep learning. Deep learning algorithms have been shown to handle more complex, high-dimensional raw input data, such as photos, with less manual feature engineering than earlier systems. This has allowed for substantial advancement in fields like computer vision and natural language processing.

### ***5.1.2 Reinforcement Learning (RL)***

Reinforcement learning teaches an agent to make judgments by allowing them to make mistakes. An agent in a Markov decision process (MDP) is in a state at each time step, takes action, and receives a scalar reward before transitioning to the next stage based on the dynamics of the environment (Bellemare et al. 2017). In order to maximize profits, the agent tries to figure out a policy, or a route from observations to actions (expected sum of rewards). In reinforcement learning, the dynamics are only sampled by the algorithm (as opposed to optimal control).

### ***5.1.3 Deep Reinforcement Learning (DRL)***

The states of the MDP are high-dimensional in many real-world decision-making problems. In order to solve MDPs and build specialized algorithms that operate well in this environment, deep reinforcement learning techniques usually represent the policy or other learned functions as a neural network (Li 2019).

### 5.1.4 From the RL to the DRL

For each state-action pair, Q-learning necessitates the agent keeping track of and updating a set of Q-values. Wireless networks, on the other hand, are likely to be large-scale in the future, diverse, and decentralized. As a result, the number of alternative system state values grows exponentially. Furthermore, due to the tremendous variety and unpredictability of system components and surrounding parameters. Hidden system states, or even an infinite number of system states, are possible (François-Lavet et al. 2018; Nguyen Cong Luong et al. 2019). The expense of computing and storing all Q-values becomes essentially unsustainable in this case. The “curse of dimensionality” is a term used to describe this condition. DRL blends RL and deep learning methodologies to solve this problem. Deep Q-learning, which makes use of a DQN, is a common feature of DRL models.

As the data source, all input states are routed through several layers of neural networks. each with its own set of weight factors. Finally, the DQN generates a set of Q-values for each of the actions that can be taken. The purpose of learning in the DQN is to use historical data such as Q-values, actions, and state changes to train and pick the most viable weight components. Because the underlying neural network is linear in complexity, a multilayer perceptron is used to calculate the Q-values and actions for a DQN. Furthermore, the DQN’s number of inputs is solely defined by the type of each state. For each input, several state values, such as distinct channel states without, can be conveyed (Packer et al. 2019; Mnih et al. 2016; Haarnoja et al. 2018; Lillicrap et al. 2016).

In order to take into account a dynamic multichannel access problem, where clients choose the channel to exchange information and several corresponding channels follow a mysterious combined Markov model, finding a method that increases the typical long-distance number of efficient transmissions is the objective. The problem is conceptualized as a somewhat recognized Markov choice cycle with enigmatic framework components. Apply the concept of support learning and implement a thorough Q-organization to primarily overcome the challenges of obscure elements and constricting calculation (DQN). Many researchers first examine the optimal approach for fixed design channel exchanging with known framework components, then demonstrate through simulations that DQN can achieve similarly optimal execution without being aware of the framework insights. Then, using both larger reproductions and actual information, compare the demonstration of DQN with a Myopic method and a Whittle Index-based heuristic, and demonstrate that DQN achieves almost optimal execution in more perplexing conditions. Finally, we suggest a flexible DQN technique that can adapt its learning in time-varying circumstances (Li and Li 2020; Zhao et al. 2022; Nair et al. 2015; Hessel et al. 2018; Zikria et al. 2020; Du et al. 2020).

The fundamental distinctions between RL and DRL when the Q-function is created using neural network are the following two features.

- (1) The purpose of learning in the DQN is to use historical data such as Q-values, actions, and state changes to train and pick the most viable weight components.

Because the underlying neural network is linear in complexity, a multilayer perceptron is used to calculate the Q-values and actions for a DQN.

- (2) DRL reduces the Q-values of the target and the estimated Q-values by adjusting the weight every few time steps, resulting in a more stable DRL learning process.

DRL reduces model complexity greatly when compared to RL, especially when used to handle a variety of challenges in future communication systems. DRL also excels RL when it comes to extracting system properties and predicting optimal rewards and behaviours. RL, on the other hand, considers and transfers from states to mini batches. The deep network weight factors are then trained using the historical data collected by DRL. According to this, with the support of system and environment knowledge, which provides additional information, the training process can be hastened (Table 5.1).

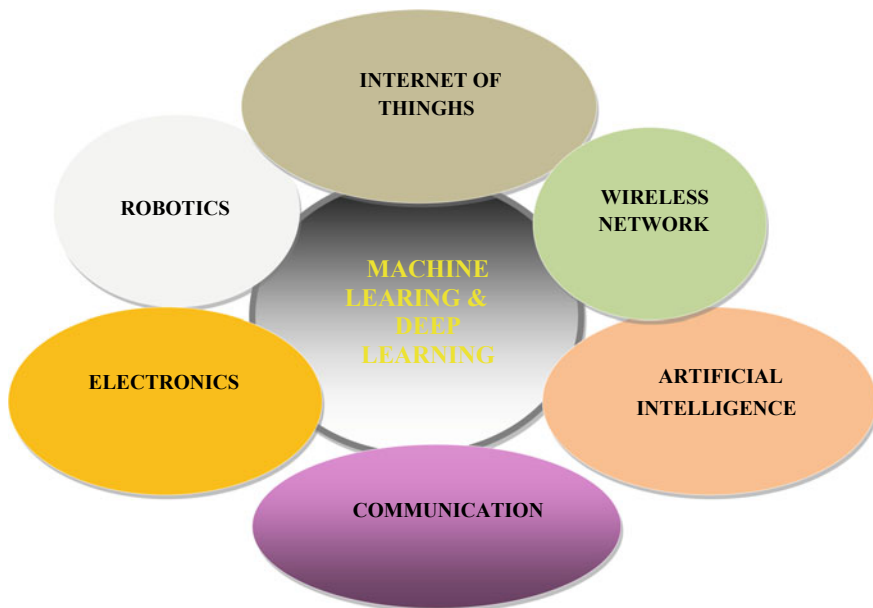
**Table 5.1** Various DRL algorithms for different areas of application

DRL algorithms	Main techniques	
Value-based	DQN (Qiu et al. 2016)	Experience replay, target network, clipping rewards, and skipping farms
	Double DQN (Wang et al. 2021)	Double Q-Learning
	Dueling DQN (Ahmed et al. 2022)	Dueling neural network architecture
	Prioritized DQN (Rai et al. 2020)	Prioritized experience replay
	Bootstrapped DQN (Qin et al. 2018)	Deep exploration with DNNs
	Distributional DQN (Tanveer et al. 2022)	Distributional Bellman equation
	Noisy DQN (Li and Li 2020)	Parametric noise added to weights
	Rainbow DQN (Hasselt et al. 2016)	Combine 6 extensions to the DQN
	Hierarchical DQN (Wang et al. 2016)	Hierarchical value functions
	Gorila (Xiang and Foo, 2021)	Asynchronous training for multi-agents
Policy based	TRPO (Bellemare et al. 2017)	KL divergence constraint
	PPO (Yang et al. 2020)	Specialized clipping in the objective function
Actro critic	Deep DPG (Wang et al. 2018)	DNN and DPG
	TD3 (Li 2019)	Twine Delayed DDPG
	PGQ (Gronauer and Diepold 2022)	Policy gradient and Q-learning
	Soft Actro Critics (SAC) (Lazaridis, et al. 2020)	Maximum entropy RL framework
	A3C (François-Lavet et al. 2018)	Asynchronous Gradient Descent

### 5.1.5 Machine Learning (ML)

Machine learning is an area where a computer does not perform tasks with explicit instructions. It includes statistical models and algorithms that give patterns and inferences obtained from the data. Machine Learning algorithms use mathematics, statistics, probability, linear algebra, and calculus deeply. The next generation of wireless communication networks is becoming more and more complex due to a wide number of services, and different characteristics of applications in devices. These approaches do not support future complex wireless systems with regards to the operation and optimization, and low cost of the networks. The exponentially increasing demand in the number of people using wireless communication technologies requires innovation and research into new technologies in the coming future (Rai et al. 2020; Qin et al. 2018; Tanveer et al. 2022). Machine learning (ML) is emerging as one of the most intelligent technologies that will help researchers conceptualize and implement technologies in virtual and augmented real life as shown in Fig. 5.2.

Nowadays, the architecture of machine learning involves major industrial interaction in the automation process. This helps in the optimization of the available resources and gives results based on the data available (Li and Li 2020; Hasselt et al. 2016; Wang et al. 2016). The machine learning architecture explains different layers involved in the machine learning process. The following major steps for the transformation of data into training data for the functioning of the intelligence of a system.



**Fig. 5.2** Application area of machine learning and deep learning

## 5.2 Applications Deep Reinforcement Learning Techniques

Deep reinforcement learning is widely employed in a range of disciplines, as seen in Table 5.2, which is based on (Hasselt et al. 2016; Lillicrap et al. 2016; Nair et al. 2015; Du et al. 2020). This chapter will cover contemporary developments in game development, robots, Transportation, industrial applications, communication, and networking are all examples of natural language processing (NLP), and other areas. Aside from DRL, a variety of methodologies can be utilized in a programme, for example, Alpha Go is taught via supervised learning and reinforcement learning. POMDP issues are given more attention in this study, However, establishing whether an application is a firm POMDP problem is difficult. Not only do we focus on POMDP issues, but we also look at deep reinforcement learning applications in general.

### 5.2.1 Application in Wireless Network

Over the past 20 years, advancements in AI have been greatly impacted by reinforcement learning, one of the most popular machine learning research subjects (Yang et al. 2020). With reinforcement learning, an agent makes consistent judgments, evaluates the results, and then modifies its approach to produce the best possible policy. Although it has been demonstrated that this learning process converges, it is

**Table 5.2** Tabular form of different application areas based on Various DRL

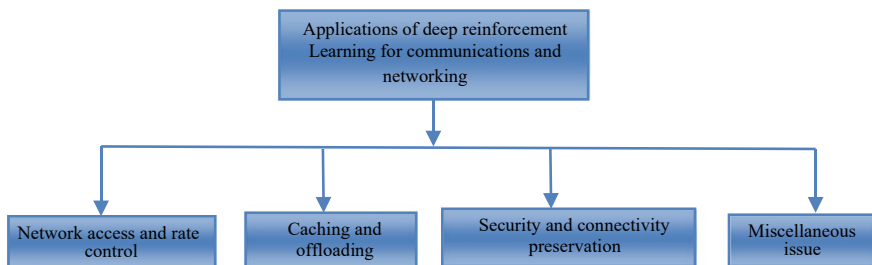
Domains	Application
Healthcare	DTRs, HER/EMR, Diagnosis
Education	Educational games, recommendation, proficiency estimation
Transportation	Traffic control
Energy	Decision control
Finance	Trading, risk management
Science, engineering, art	Math, physics, music, animation
Business management	Recommendation, customer management
Computer systems	Resource management, security
Games	Board games, card games, video games
Robotics	Sim-to-real, control
Computer vision	Recognition, detection, perception
NLP	Sequence generation, translation, dialogue

improper and inapplicable to large-scale networks since it requires to explore and learn about the entire system in order to arrive at the optimum policy. The practical application of reinforcement learning is thus constrained.

Deep learning (Yang et al. 2020; Packer et al. 2019; Haarnoja et al. 2018) has been regarded as a game-changing technique recently. It has the ability to go beyond reinforcement learning’s restrictions, ushering in a new era of reinforcement learning progress known as DRL. DNNs are used in DRL to train the learning process, allowing reinforcement learning algorithms to learn quicker and perform better. As a result, DRL has been used in a variety of reinforcement learning applications, including robotics, computer vision, speech recognition, and natural language processing (see Fig. 5.3).

The DRL techniques have the following advantages in general:

- DRL is capable of resolving difficult network optimization issues. As a result, without having extensive and precise network knowledge, network controllers in current networks, base stations, for example, are capable of resolving non-convex and difficult problems including joint user association, communication, and transmission scheduling
- DRL enables network entities to better understand and learn about the communication and networking environment. Without knowing the channel model or mobility pattern, network entities, such as a mobile user, DRL can be used to discover the optimum base station, channel, handover, caching, and offloading rules.
- Independent decision-making is possible with DRL. Network entities can use DRL approaches to observe and obtain the best policy locally.
- DRL dramatically accelerates learning, particularly in circumstances with large state and action spaces. In order to manage dynamic user association, spectrum access, and transmit power for a sizable number of mobile users and IOT devices, such as IoT systems with thousands of devices, DRL offers network controllers or IOT gateways.
- Other communication and networking challenges, Games, such as the non-cooperative game, can be used to represent cyber-physical threats, interference



**Fig. 5.3** Applications of deep reinforcement learning in communications and networking (Yang et al. 2020)



control, and data dumping. DRL has lately been used to solve games that need imperfect knowledge, such as Nash equilibrium determination.

- Wireless networks of the future are predicted to be large-scale, diverse, and decentralized (Wang et al. 2018; Xiong et al. 2019). When it comes to system parameters and states, The Q-function for the optimum actions to maximize system services and resources is calculated and updated by the agent of a network entity. The rewards earned by the agent can differ when the same behaviours are conducted with the same system states due to the uncertainty of mobile communication systems and the agent's short-sighted perspective. In this case, an appropriate updating rate, or learning rate, must be used to update the Q-function iteratively until convergence is attained. The process of updating the Q-function is referred to as "training."

### 5.3 DRL Applications for Future-Generation Mobile Networks

The majority of future-generation network optimization challenges have been solved using centralized optimization techniques. Such approaches are, on the other hand, based on fundamental hypotheses on the entire information requirements of network conditions. These presumptions are getting less and less useful as mobile networking scenarios become more unexpected. Unlike centralized optimization approaches, the DRL methodology makes no fundamental assumptions about the target system, making it a practical technique for dealing with a variety of difficulties in mobile 5G and beyond (François-Lavet et al. 2018; Fujimoto et al. 2018; Sharma et al. 2021; Schulman et al. 2017).

#### 5.3.1 *Power Management and Power Control*

Because of the increasingly varied and dense environment in 5G networks, traditional intercell interference coordination will be inefficient, if not impossible. To reduce interference, a transmitter can lower its transmit power. The data rate, however, may deteriorate as a result. In mobile networks, interference reduction is frequently viewed as a power control optimization problem. Power management for network devices increases energy effectiveness while reducing expenses and carbon footprint. DRL enables network organizations to gain knowledge about their networks and make the best power control and management decisions possible.

### **5.3.1.1 Controlling Power in Cellular Networks**

A DRL-based distributed dynamic power-allocation system was created by the authors of Qin et al. (2018) to take into consideration the unpredictable nature of 5G networks. Each transmitter performs the role of a system agent, and all agents are coordinated and carry out their tasks simultaneously. In order to predict the impact of current actions on future neighbour behaviour, each agent can study the consequences of its neighbours' prior activities throughout the current decision period before acting. With the aid of DRL, each agent chooses a policy that maximizes the discounted expected future benefit.

### **5.3.1.2 Ultra Dense Network Power Management**

Ultradense networks (UDNs) consume an excessive amount of electricity due to the dense deployment of tiny BSs (SBSs). As a result, dynamically turning SBSs off is a cost-effective technique to save energy. The ON/OFF challenge in energy-harvesting UDNs has been transformed into a dynamic optimization problem. DRL is taught in order to learn a method for selecting SBS ON/OFF modes that maximize energy efficiency while accounting for the unpredictability of energy charging, channel condition information, and traffic arrivals. The ON/OFF scheduling method based on DRL consumes less energy than Q-learning.

### **5.3.1.3 In mm-wave Communications, Power Control is Important**

Non-line-of-sight (NLOS) transmission will be a key challenge with the introduction of mm-wave communications in 5G. To increase NLOS transmission performance, the authors suggested a dynamic transmission power regulation technique. Within the restrictions of transmission power and QoS standards, In the 5G network, the overall data rate achieved by all UEs should be maximized. To forecast the Q-function, convolutional neural networks are first trained offline. The DRL is then utilized to conduct activities such as UE association and power allocation.

### **5.3.1.4 Computation Offloading at the Mobile Edge**

A UDN with a single UE and a high number of BSs was examined. Depending on channel quality, the UE's computing responsibilities may be delegated to one of the BSs. The optimal offloading problem is tackled as a Markov decision process in order to lower the long-term anticipated cost (MDP). The aforementioned MDP challenge was addressed by developing a DRL-based online strategic computational offloading technique due to the dynamics of time-varying channel quality and the unexpected nature of task arrivals. The numerical outcomes demonstrate that the suggested strategy significantly lowers average cost. This system only has one server,

as opposed to a single-server edge computing system that enables several UEs to offload computation to the server using wireless channels. The overall goal of the optimization is to reduce total delay costs and energy consumption across all UEs. The computation offloading problem is solved using a DRL-based technique to discover the best strategy and avoid the dimensionality curse.

### **5.3.1.5 Caching at the Edges**

Edge caching can reduce transmission costs by lowering latency and easing traffic congestion on backhaul connections. Content caching, on the other hand, poses the challenge of policy control: in the face of continually changing content popularity distributions, the contents to be saved in the cache must be determined by the network edge device containing the cache. A single BS acts as a cache node for a large number of mobile users who make frequent requests for content from the BS. The authors employed DRL for the BS's cache replacement decisions to handle the requests efficiently. The objective, which is supported by evidence, is to raise the long-term cache hit rate.

### **5.3.1.6 Caching and Computing at the Edge**

An integrated framework for coordinating computing, networking, and caching resource management in vehicular networks, in contrast to previous studies that looked at edge computing and caching concerns individually. The computational capabilities of the cache nodes' edge servers and statuses, such as the BS channel conditions, change dynamically. Furthermore, because the system state spaces are enormous, determining which resources should be allocated to which vehicle is difficult. DRL must therefore choose the best resource allocation strategy. In a similar vein, integrate computation, caching, and communication design to boost automobile networks' performance.

### **5.3.1.7 Transportation that is Intelligent**

Vehicle networks may be able to give real-time intelligent services as 5G technology advances. Vehicular networks are anticipated to play a significant role in the development of intelligent transportation systems by enhancing the design of reliable and effective transportation systems by utilizing improved communications and data-collection methods. Based on DRL, a decentralized resource allocation system for vehicle networks. The system is used to figure out how to map incomplete observations from each car to the most efficient resource allocation. Without the necessity for precise prior knowledge, this approach can meet rigorous latency requirements for vehicle-to-vehicle links.

### 5.3.1.8 Slicing a Network

Network slicing is seen as the essential paradigm for satisfying varied service requirements, thanks to the greater softwarization provided in 5G networks. The necessity for slicing in mobile 5G networks is driven by the complexity of network resource management, considering the exponential increase of wireless data services, as well as various service requirements and heterogeneous wireless settings. Slicing divides, the network into a number of segments (slices), each of which is created and optimized to meet unique service demands. Right now, network slicing is a hot topic in both academia and business. An economic evaluation for distributing requests for network slices while taking into account the 5G infrastructure provider's maximum income.

## 5.4 Future Prediction of the Wireless Networks

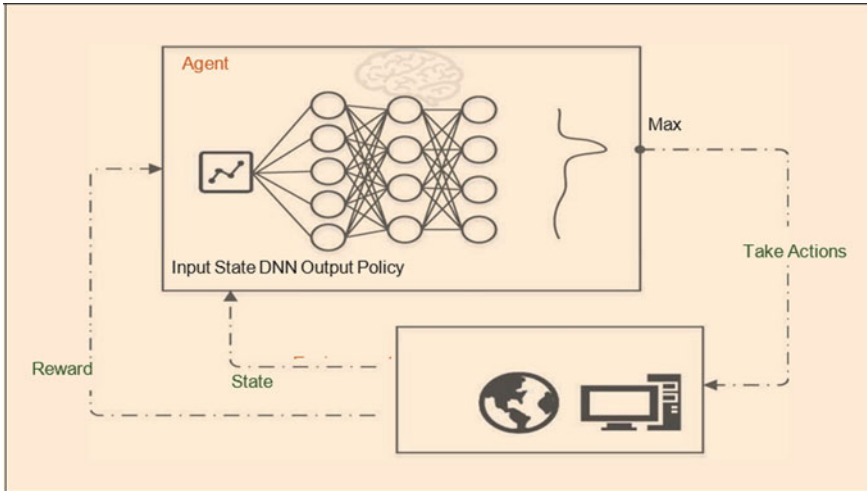
Wireless networks of the future are predicted to be large-scale, diverse, and decentralized. When it comes to system parameters and states, The Q-function for the optimum actions to maximize system services and resources is calculated and updated by the agent of a network entity. The rewards earned by the agent can differ when the same behaviours are conducted with the same system states due to the uncertainty of mobile communication systems and the agent's short-sighted perspective. In this case, an appropriate updating rate, or learning rate, must be used to update the Q-function iteratively until convergence is attained. The process of updating the Q-function is referred to as "training".

For each state-action pair, Q-learning necessitates the agent keeping track of and updating a set of Q-values. Wireless networks, on the other hand, are likely to be large-scale in the future, diverse, and decentralized. As a result, the number of alternative system state values grows exponentially. In addition, due to the extreme diversity and unpredictability of system parts and environmental variables (Priya et al. 2021; Jagannath et al. 2019; Akhtar et al. 2021; Du et al. 2020; Liu and Wang 2017). There could be an infinite number of system states or even concealed system states. In this case, it becomes practically impossible to justify the expense of computing and storing all Q-values. This phenomenon is referred to as the "curse of dimensionality".

DRL combines RL and deep learning methods to overcome this issue. DRL models frequently have deep Q-learning, which makes use of a DQN.

As the data source, all input states are routed through several layers of neural networks. each with its own set of weight factors. Finally, the DQN generates a set of Q-values for each of the actions that can be taken as shown in Fig. 5.4.

The purpose of learning in the DQN is to use historical data such as Q-values, actions, and state changes to train and pick the most viable weight components. Because the underlying neural network is linear in complexity, a multilayer perceptron is used to calculate the Q-values and actions for a DQN. Furthermore, The



**Fig. 5.4** This image depicts a DQN. The letters DNN and Max stand for deep neural network and maximum, respectively

DQN's number of inputs is solely defined by the type of each state. For each input, several state values, such as distinct channel states without, can be conveyed.

The fundamental distinctions between RL and DRL when the Q-function is created using a neural network are the following two features.

- (1) The purpose of learning in the DQN is to use historical data such as Q-values, actions, and state changes to train and pick the most viable weight components. Because the underlying neural network is linear in complexity, a multilayer perceptron is used to calculate the Q-values and actions for a DQN.
- (2) DRL reduces the Q-values of the target and the estimated Q-values by adjusting the weight every few time steps, resulting in a more stable DRL learning process.

DRL reduces model complexity greatly when compared to RL, especially when used to handle a variety of challenges in future communication systems. DRL also excels RL when it comes to extracting system properties and predicting optimal rewards and behaviours. RL, on the other hand, considers and transfers from states to mini batches. The deep network weight factors are then trained using the historical data collected by DRL. According to this, with the support of system and environment knowledge, which provides additional information, the training process can be hastened.

## 5.5 Wireless Mobile Communications and the Future of the Indian Cellular Market

Since its debut, wireless technology has advanced significantly. Over the past 10 years, there has been a remarkable increase in the number of wireless customers worldwide. Today, there is a large demand for wireless technology across a wide range of applications, including Internet and online browsing, video, and other text- and multimedia-based ones (Qiu et al. 2016). The usage of wireless technology is no longer solely restricted to wireless telephony. With more than 3 billion subscribers in over 215 countries, GSM is the most commonly used 2nd generation digital cellular standard. It gains about 1000 new users every minute. GSM, which provides wide area voice communications utilizing 200 kHz carriers and employs both TDMA and FDMA methods, was initially introduced in 1991 after being developed in the 1980s. Later, it evolved into 2.5 G standards with the introduction of packet data transmission technology (GPRS) and faster data speeds using higher order modulation methods (EDGE). For GSM/EDGE-based radio access network, they are now collectively referred to as GERAN (Wang et al. 2021). Mobile communications have quickly evolved into an essential symbol of our times since the development of the first genuinely worldwide mobile telecommunications system, GSM, in 1992. In order to protect the confidentiality of user-related information, some security mechanisms were incorporated into the original GSM. In addition to audio communication, GSM offers mobile phone communication services based on digital data exchange at up to 9.6 kbps (Ahmed et al. 2022). The majority of GSM standards were developed with operators in mind in order to minimize fraud and network abuse; operators were given the duty of integrating features relevant to user privacy (Rai et al. 2020). Even though they represent a significant evolutionary step, new modes like GPRS and EDGE will seamlessly integrate into the current GSM architecture and only slightly alter the network and base stations. In order to meet the criteria of UMTS, It's intriguing to consider how the evolutionary path can be expanded to even higher bit rates, enabling an even wider variety of services (Qin et al. 2018). There are issues with the current GSM:

- (1) The end systems of GSM, a circuit-switched, connection-oriented technology, are set aside for the duration of each call. This results in inaccurate bandwidth and resource utilization.
- (2) High data speeds are not supported by GSM-enabled systems (Ahmed et al. 2022).
- (3) The fact that several users share the same bandwidth is GSM's biggest drawback. If there are many users, interference may occur during the broadcast. In order to get around these bandwidth restrictions, speedier technologies, such as CDMA, have been developed on networks other than GSM.
- (4) According to [Int.Technology.com](http://www.Technology.com), another drawback of GSM is that it can interfere with some human electronics, such as pacemakers and hearing aids. Pulse-transmission technology used by GSM is what's causing this disturbance.

Therefore, it is required to turn off cell phones in many locations, such as hospitals and aircraft (Tanveer et al. 2022). Today's GSM users take for granted the ability to use their cell phones' voice roaming features to place and receive calls while travelling to other countries. The GSM mobile operators' packet-based General Packet Radio Service (GPRS) network provides a variety of services like video conferencing, multimedia apps, etc. However, this new network also makes it more difficult to accommodate the packet-based roaming scenario. Due to the projected volume of services, operators, and customers, it will be hard for operators to test all services for GPRS roaming across all partner networks. In order to ensure service continuity, consideration must be taken into account that the network capabilities of a visited GPRS network may differ from those of the home network. This takes us to a further consideration for mobile data service developers. This is especially important now that mobile operators are starting to provide 3G networks and support 3G to 2G network roaming agreements (Li and Li 2020).

### ***5.5.1 The Growth Factor of the Telecom Sector in India***

With 1.16 billion customers, India is currently the second-largest telecoms market in the world and has had impressive development in recent years. According to a report by the GSM Association (GSMA) and Boston Consulting Group, India's mobile economy is expanding quickly and will have a significant impact on the country's GDP (BCG). The US was surpassed by India in 2019 to become the second-largest market for app downloads.

Along with high consumer demand, the liberal and reformist policies of the Indian government have been essential to the sector's quick rise. The government has made it feasible for a fair and proactive regulatory structure to exist, as well as easy market access for telecom equipment, guaranteeing that customers may receive telecom services at fair prices. Due to the relaxation of Foreign Direct Investment (FDI) standards, the sector is among the five fastest growing and largest employment in the country.

### ***5.5.2 Methodology Used in the Overall World***

It is based on secondary information acquired from sources such as the Telecom Regulatory Authority of India, the Ministry of Communication, the Department of Telecommunication, government reports, and other sources. In order to assess the established targets, statistical tools such as annual growth rate, percentage, and year-by-year market share of different service providers were calculated.

### ***5.5.3 Market Size Especially in India***

In terms of overall internet users, India is the second-largest market in the world. 757.61 million people utilized the internet overall in January 2021. The number of wireless or mobile phone subscribers increased from 1,153.77 million in December 2020 to 1,163.41 million in January 2021. The second-largest telecom market in the world is found in India. As of January 2021, there were 1,183.49 million users nationwide.

In the third quarter of FY21, the telecom industry's gross revenue was Rs. 68,228 crores. Over the next five years, 500 million more Indians will have access to the internet because of rising mobile phone adoption and declining data costs, creating new business opportunities.

### ***5.5.4 Growth Factor of Telecommunication in India***

The second-largest telecoms market in the world is in India. The wireless, wireline, and internet services categories make up the telecom market. As of January 2021, there were 1,183.49 million users nationwide. Rural customers' teledensity increased from 59.05% in December 2020 to 59.50% in January 2021, indicating that there is still a significant amount of room for demand to rise in this sector. The number of wireless or mobile phone subscribers increased from 1,153.77 million in December 2020 to 1,163.41 million in January 2021.

In terms of internet users, India is the second-largest nation. One of the top data consumers worldwide is India. According to TRAI, each wireless data subscriber used an average of 11 GB of data each month in FY20. From 12.07 billion in 2017 to 19 billion in 2019, app downloads nationwide increased significantly, and it is predicted that this figure would reach 37.21 billion by 2022F. In the third quarter of FY21, India's overall wireless data usage increased 1.82% quarterly to 25,227 PB. Data consumption on mobile devices increased overall by 2.81% and 96.48%, respectively, in the third quarter of FY21 thanks to 3G and 4G data usage. In the identical quarter, 2G data usage made up 0.71% of all data usage.

Gross revenue for the telecom sector in the third quarter of FY21 was Rs. 68,228 crores. The sector's development has been dependent on the government's significant policy support. The maximum amount of FDI allowed in the telecom sector has increased from 74 to 100%. FDI totaling US\$37.62 billion was invested in the telecom sector between April 2000 and December 2020.

India is predicted to have the fastest-growing telecom advertising market between 2020 and 2023, with an average annual growth rate of 11%, according to a Zenith Media report. According to its National Digital Communications Policy, the Indian government anticipates investing \$100 billion in the telecom sector by 2022. International manufacturers of telecom networks like Ericsson, Nokia, Samsung, and



Huawei are being enticed by the government to create all of their equipment in India using only locally available materials.

The TEPC organized India Telecom 2021 in March 2021 as a venue for the integration of technologies and business exchange (Telecom Equipment Export Promotion Council).

The Union Cabinet approved an Rs. 12,195 crores production-linked incentive (PLI) package for telecom and networking devices under the Department of Telecom. Petabytes and Forecast are abbreviations for each other.

### ***5.5.5 Major Market Players or Companies of Telecommunication in India***

Bharti Airtel Limited, a major international telecommunications firm with operations in 16 nations throughout Asia and Africa, is known as Airtel. Its main office is located in New Delhi, India. In terms of subscribers, the firm is one of the top three mobile service providers internationally. The company provides 2G, 3G, and 4G cellular services, mobile commerce, fixed line services, high-speed home internet, DTH, and both domestically and internationally long-distance services to carriers in India. In the remaining regions, it provides mobile shopping in addition to 2G, 3G, and 4G wireless services.

In the most recent spectrum auction held by the Department of Telecom in March 2021, Bharti Airtel paid a total of Rs. 18,699 crores for 355.45 MHz of spectrum in the Sub GHz, mid-band, and 2300 MHz bands (Government of India).

Bharti Airtel entered the advertising sector in February 2021 with the introduction of Airtel Advertisements, a productive brand interaction solution. Bharti Airtel and Qualcomm Technologies, Inc. hastened the rollout of 5G in India in February 2021. Airtel just made history in Hyderabad by testing 5G on a LIVE commercial network.

Reliance Jio's ecosystem as a whole allows Indians to completely live their digital lives. Strong broadband networks, useful applications, top-notch services, and smart devices are all components of this ecosystem, which is accessible to every Indian home. The most extensive collections of recorded and live music, sports, live and catch-up television, movies, and events are among Jio's media offerings. Jio wants to unleash the power of a young nation by developing linked intelligence for the 6 billion minds on the earth. Thanks to a triple strategy of concentrating on broadband networks, reasonably priced mobile devices, and the availability of high-quality content and applications, Jio has been successful in creating an integrated business model from the start.

Jio has the ability to provide a unique combination of media, fast bandwidth, digital commerce, telephony, and financial services. Jio Business unveiled an integrated service for micro, midsize, and medium-sized companies (MSMB) in March 2021 with the goal of providing integrated fibre connection and digital solutions. Reliance Jio announced the purchase of the spectrum in March 2021, earning the ability to

utilize it across all 22 circles in India. More than 100 million feature phone customers in India were successfully upgraded to the Jio Phone network in February 2021, ushering in a new era of change for the country's feature phone consumers.

**BSNL:** BSNL is a technology-driven firm that offers a wide range of telecom services, including landline, mobile, and wireless local loop (WLL) telephone services, as well as broadband, leased circuits, and long-distance telecom service. With a switching network made entirely of digital technology, the firm has remained at the forefront of technology. All district, sub-divisional, tehsil, and nearly all block headquarters are covered by the nationwide telecommunications network of BSNL.

## 5.6 Summary

By 2020, it is anticipated that the telecom equipment industry would generate US\$26.38 billion in revenue. By 2021, it's anticipated that there will be 829 million more internet members in the nation, and overall IP traffic will have increased fourfold at a CAGR of 30%.

India is predicted to have the fastest-growing telecom advertising industry between 2020 and 2023, with an average annual growth rate of 11%, according to a Zenith Media report.

IoT will be crucial in the development of the 100 smart city projects that the Indian government plans to implement. According to the National Digital Communications Policy of 2018, by 2022, the telecoms industry was expected to receive investments totaling \$100 billion USD. There will be 18.11 billion app downloads in India in 2018 and 37.21 billion in 2022, according to forecasts.

## References

- Ahmed QW, Garg S, Rai A, Ramachandran M, Jhanjhi NZ, Masud M, Baz M (2022) AI-Based resource allocation techniques in wireless sensor internet of things networks in energy efficiency with data optimization. *Electronics* 11(13):2071
- Akhtar T, Tselios C, Politis I (2021) Radio resource management: Approaches and implementations from 4G to 5G and beyond. *Wirel. Netw.* 27:693–734
- Du Z, Deng Y, Guo W, Nallanathan A, Wu Q (2020) Green deep reinforcement learning for radio resource management: architecture, algorithm compression, and challenges. *IEEE Veh Technol Mag* 16:29–39
- François-Lavet V et al (2018) An Introduction to deep reinforcement learning. *Found Trends Mach Learn* 11:219–354
- Fujimoto S, Hoof HV, Meger D (2018) Addressing function approximation error in actor-critic methods. *ArXiv, abs/1802.09477*
- Gronauer S, Diepold K (2022) Multi-agent deep reinforcement learning: a survey. *Artif Intell Rev* 55:895–943
- Haarnoja T, Zhou A, Abbeel P, Levine S (2018) Soft actor-critic: off-policy maximum entropy deep reinforcement learning with a stochastic actor. *ICML*

- Hessel M, Modayil J, van Hasselt H, Schaul T, Ostrovski G, Dabney W, Horgan D, Piot B, Azar M, Silver D (2018) Rainbow: Combining improvements in deep reinforcement learning. In: Proceedings of the AAAI Conference on Artificial Intelligence, 32(1). <https://doi.org/10.1609/aaai.v32i1.11796>
- Jagannath J, Polosky N, Jagannath A, Restuccia F, Melodia T (2019) Machine learning for wireless communications in the Internet of Things: A comprehensive survey. *Ad Hoc Netw* 93:101913
- Lazaridis, Aristotelis et al. (2020): Deep reinforcement learning: A state-of-the-art walkthrough. *J. Artif. Intell. Res.* 1421–1471
- Li Y (2019). Deep reinforcement learning. ArXiv, abs/1810.06339
- Li M, Li H (2020) Application of deep neural network and deep reinforcement learning in wireless communication. *PLoS ONE* 15(7):e0235447. <https://doi.org/10.1371/journal.pone.0235447>
- Lillicrap TP, Hunt JJ, Pritzel A, Heess NM, Erez T, Tassa Y, Silver D, Wierstra D (2016) Continuous control with deep reinforcement learning. *CoRR*, abs/1509.02971
- Liu Y, Wang DL (2017) Speaker-dependent multipitch tracking using deep neural networks. *J Acoust Soc Am* 141(2):710–721. <https://doi.org/10.1121/1.4973687>. PMID: 28253703
- Marc G. Bellemare, Will Dabney, Rémi Munos, (2017). A distributional perspective on reinforcement Learning. In: Proceedings of the 34th International Conference on Machine Learning, PMLR 70:449–458
- Mnih V, Badia AP, Mirza M, Graves A, Lillicrap TP, Harley T, Silver D, Kavukcuoglu K (2016) Asynchronous methods for deep reinforcement learning. *ICML*
- Nair A, Srinivasan P, Blackwell S, Alciçek C, Fearon R, Maria AD, Panneershelvam V, Suleyman M, Beattie C, Petersen S, Legg S, Mnih V, Kavukcuoglu K, Silver D (2015) Massively parallel methods for deep reinforcement learning. ArXiv, abs/1507.04296
- Nguyen Cong Luong Et Al (2019.), Applications Of deep reinforcement learning in communications and networking: A survey, *IEEE communications surveys & tutorials*, 21(4), Fourth Quarter <https://doi.org/10.1109/COMST.2019.2916583>
- Packer, Charles; Gao, Katelyn; Kos, Jernej; Krähenbühl, Philipp; Koltun, Vladlen; Song, Dawn (2019) Assessing generalization in deep reinforcement learning 2019. In: 33rd Conference on Neural Information Processing Systems
- Priya A, Rai A, Singh RP (2021) Internet of things: architecture, applications and future aspects. In *Advances in Smart Communication and Imaging Systems* (pp 183–190). Springer, Singapore
- Qin M, Yang Q, Cheng N, Zhou H, Rao RR, Shen X (2018) Machine learning aided context-aware self-healing management for ultra-dense networks with QoS provisions. *IEEE Trans Veh Technol* 67:12339–12351
- Qiu J, Wu Q, Ding G et al (2016) A survey of machine learning for big data processing. *EURASIP J Adv Signal Process* 2016:67. <https://doi.org/10.1186/s13634-016-0355-xA>
- Rai A, Sehgal A, Singal TL, Agrawal R (2020). Spectrum sensing and allocation schemes for cognitive radio. *Mach Learn Cogn Comput Mob Commun Wirel Netw* 91–129
- Schulman J, Wolski F, Dhariwal P, Radford A, Klimov O (2017) Proximal policy optimization algorithms. ArXiv, abs/1707.06347
- Sharma D, Singhal S, Rai A, Singh A (2021) Analysis of power consumption in standalone 5G network and enhancement in energy efficiency using a novel routing protocol. *Sustain Energy, Grids Netw* 26:100427
- Tanveer J, Haider A, Ali R, Kim A (2022) Machine learning for physical layer in 5G and beyond wireless networks: a survey. *Electronics* 11:121. <https://doi.org/10.3390/electronics11010121>
- Van Hasselt H, Guez A, Silver D (2016). Deep reinforcement learning with double Q-Learning. In: Proceedings of the AAAI Conference on Artificial Intelligence, 30(1). <https://doi.org/10.1609/aaai.v30i1.10295>
- Wang S, Liu H, Gomes PH, Krishnamachari B (2018) Deep reinforcement learning for dynamic multichannel access in wireless networks. *IEEE Trans Cogn Commun Netw* 4(2):257–265
- Wang Z, Schaul T, Hessel M, van Hasselt H, Lanctot M, de Freitas N (2016). Network architectures for deep reinforcement learning. In Proceedings of the 33rd International Conference on Machine Learning, New York, NY, USA

- Wang S, Chen M, Liu X, Yin C, Cui S, Vincent Poor H (2021). A machine learning approach for task and resource allocation in mobile-edge computing-based networks. *IEEE Internet Things J* 8(3), 1358–1372. [9146372]. <https://doi.org/10.1109/JIOT.2020.3011286>
- Xiang X, Foo S (2021). Recent advances in deep reinforcement learning applications for solving partially observable markov decision processes (POMDP) Problems: Part I—Fundamentals and applications in games, robotics and natural language processing. *Mach Learn Knowl Extr*, 3, 554–581. <https://doi.org/10.3390/make3030029>
- Xiong Z, Zhang Y, Niyato D, Deng R, Wang P, Wang L-C (2019) Deep reinforcement learning for mobile 5G and beyond: Fundamentals, applications, and challenges. *IEEE Veh Technol Mag* 14(2):44–52. <https://doi.org/10.1109/MVT.2019.2903655>
- Yang K, Shen C, Liu T, (2020) Deep Reinforcement Learning based Wireless Network Optimization: A Comparative Study. In: *IEEE INFOCOM 2020—IEEE Conference on Computer Communications Workshops*, pp 1248–1253, doi: <https://doi.org/10.1109/infocomwkshps50562.2020.9162925>
- Zhao X, Yang R, Zhang Y, Yan M, Yue L (2022) Deep reinforcement learning for intelligent Dual-UAV reconnaissance mission planning. *Electronics* 11:2031. <https://doi.org/10.3390/electronics1113203>
- Zikria YB, Afzal MK, Kim SW, Marin A, Guizani M (2020) Deep learning for intelligent IoT: Opportunities, challenges and solutions. *Comput Commun* 164:50–53