[Array 14 (2022) 100136](https://doi.org/10.1016/j.array.2022.100136)

|  |  |  |
| --- | --- | --- |
|  | Contents lists available at [ScienceDirect](http://www.elsevier.com/locate/array) |  |
| Array |
| journal homepage: [www.elsevier.com/locate/array](http://www.elsevier.com/locate/array) |
|  | | |

Knowledge-defined networking: Applications, challenges and future work Sepehr Ashtaria,∗, Ian Zhoua, Mehran Abolhasana, Negin Shariatia, Justin Lipmana, Wei Nib a *University of Technology, Sydney, Australia*   
b *The Commonwealth Scientific and Industrial Research Organisation, Sydney, Australia*

|  |  |  |
| --- | --- | --- |
| A R T I C L E | I N F O | A B S T R A C T |
| *Keywords:*  6G  Knowledge-defined networking (KDN) Software-defined networking (SDN) Open-radio access network (O-RAN) Localization  Machine learning  Mobility coordination  Network management  Resource management | | Future 6G wireless communication systems are expected to feature intelligence and automation. Knowledge-defined networking (KDN) is an evolutionary step toward autonomous and self-driving networks. The building blocks of the KDN paradigm in achieving self-driving networks are software-defined networking (SDN), packet-level network telemetry, and machine learning (ML). The KDN paradigm intends to integrate intelligence to manage and control networks automatically. In this study, we first introduce the disadvantages of current network technologies. Then, the KDN and associated technologies are explored with three possible KDN architectures for heterogeneous wireless networks. Furthermore, a thorough investigation of recent survey studies on different wireless network applications was conducted. The aim is to identify and review suitable ML-based studies for KDN-based wireless cellular networks. These applications are categorized as resource management, network management, mobility management, and localization. Resource management applications can be further classified as spectrum allocation, power management, quality-of-service (QoS), base station (BS) switching, cache, and backhaul management. Within network management configurations, routing strategies, clustering, user/BS association, traffic classification, and data aggregation were investigated. Applications in mobility management include user mobility prediction and handover management. To improve the accuracy of positioning in indoor environments, localization techniques were discussed. We classify existing research into the respective KDN architecture and identify how the knowledge obtained will enhance future networks; as a result, researchers can extend their work to empower intelligence and self-organization in the network using the KDN paradigm. Finally, the requirements, motivations, applications, challenges, and open issues are presented. |

**1. Introduction**

Fifth-generation (5G) wireless communication systems provide higher data rates, massive connectivity, and low-latency communi-cation. However, the current 5G cellular architecture lacks intelli-gence and sufficient flexibility to handle massive machine-type com-munication (mMTC), low-latency, and enhanced mobile broadband (eMBB) [1]. The sixth-generation (6G) cellular network is a promising technology to address 5G shortcomings. To achieve this, 6G will enable greater intelligence within the network to overcome a number of challenges and improve performance [2]. As a result, an architectural transformation is required for 5G to 6G cellular networks.

The concept of the knowledge plane (KP) was introduced by Clark et al. [3]. As shown in Fig. 1 KP is an additional plane over a network with inbuilt machine learning (ML) capabilities. The incorporation of KP in software-defined networking (SDN) architecture is referred to as knowledge-defined networking (KDN) [4], where knowledge is the pro-cessed network information using an ML algorithm. Knowledge is used

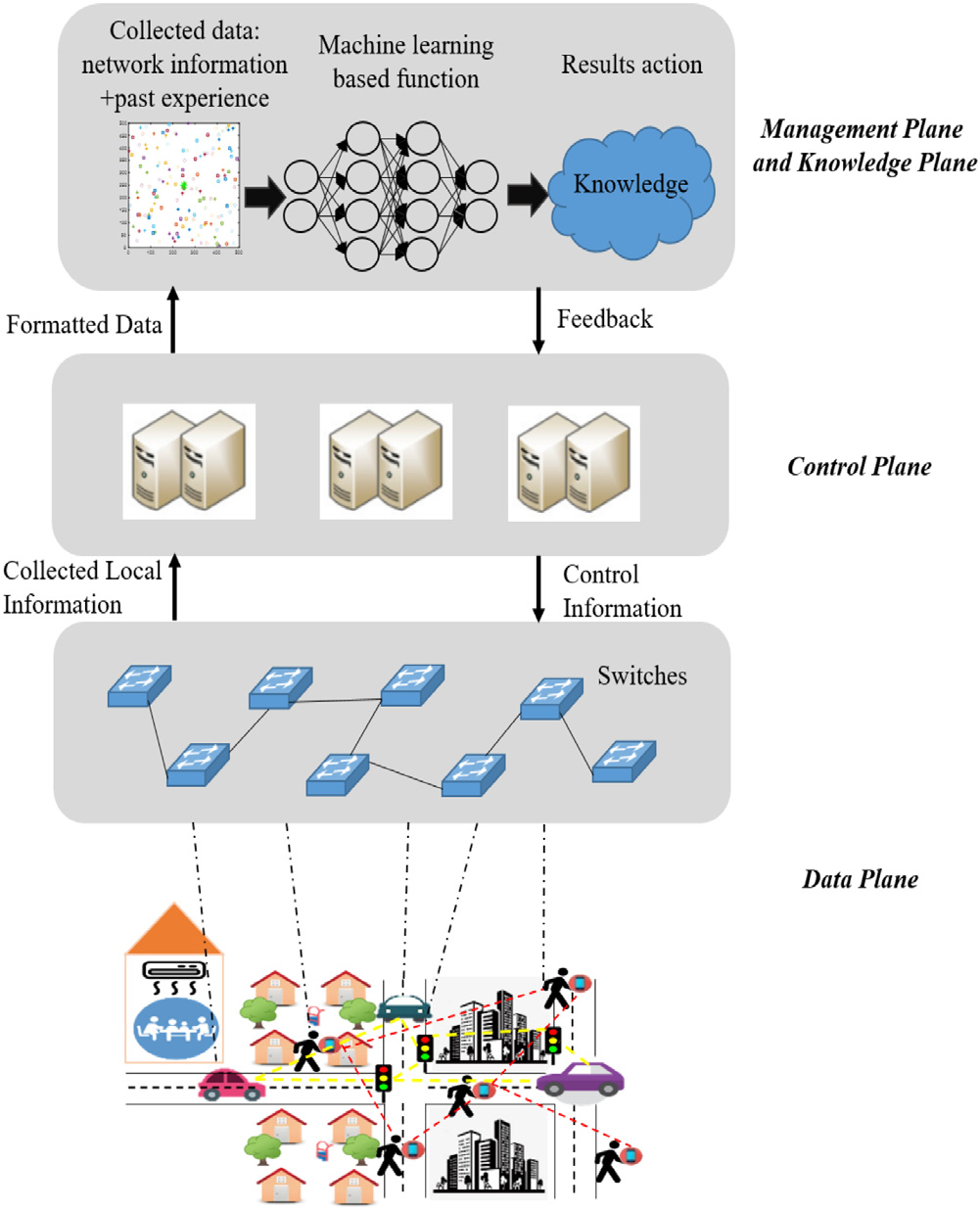
∗ Corresponding a[uthor.](mailto:sepehr.ashtarinakhaei@student.uts.edu.au)

*E-mail address:* [sepehr.ashtarinakhaei@student.uts.edu.au](mailto:sepehr.ashtarinakhaei@student.uts.edu.au) (S. Ashtari).

for recommendation and automation across different applications in wired and wireless networks. Therefore, knowledge can be referred to as intelligence over a network, and having intelligence over a network with different environmental characteristics can be a breakthrough in network performance. For instance, in resource management problems, parameters such as bandwidth, quality of service (QoS), and power can be obtained and processed using an ML algorithm in different network situations. The output of ML is then stored as knowledge for network automation. Moreover, in networking applications, routing decisions can benefit from knowledge for better route discovery while the network is overpopulated. Further, user information, including mo-bility patterns and velocity, can be used as an initial stage to generate knowledge to improve the accuracy of localization and handover. KDN can also be referred to as an application of autonomous networking. The concept of an autonomous network comes from the growth of ML and artificial intelligence (AI); for instance, in self-driving cars, an ML agent will run the car without a human operator. Similarly,

<https://doi.org/10.1016/j.array.2022.100136>  
[Received 29 September 2021; Accepted 22 F](https://doi.org/10.1016/j.array.2022.100136)ebruary 2022   
Available online 4 April 2022   
[2590-0056](http://creativecommons.org/licenses/by-nc-nd/4.0/)/© 2022 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

*S. Ashtari et al.*  *Array 14 (2022) 100136*



**Fig. 1.** Knowledge-defined orchestration in wireless networks.

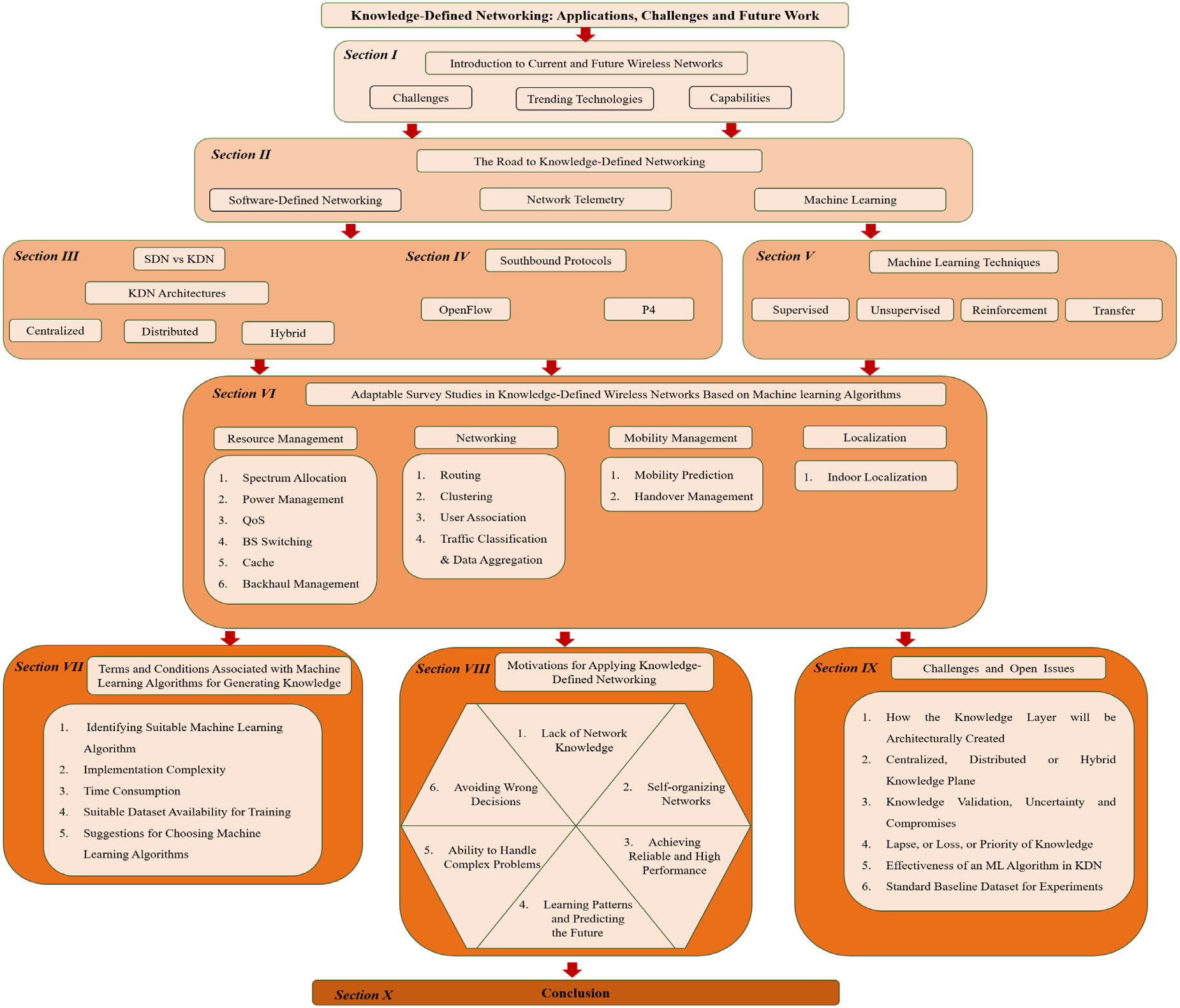
an autonomous network can manage and optimize network applica-tions without human intervention. In autonomous networks, network information or telemetry is collected and used by ML techniques to automatically troubleshoot, instruct, or manage the network. Hence, monitoring and retrieval of network telemetry data in real time will provide an opportunity for ML-based optimization algorithms to enable intelligence for 6G networks.

The fundamental building blocks in KDN are network telemetry, SDN, and ML. Network telemetry is network information, such as Net-Flow data, sFlow data, queue occupancy, policy rules, and processing time. However, to fully auto-mate the network, information such as hop latency, link utilization, packet drop, and queue congestion states are also required. This temporary data is available through in-band net-work telemetry (INT) or packet-level network states. The packet-level network state information can be collected using a new southbound domain-specific language called P4 [5], which allows the collection of information directly from the data plane. As shown in Fig. 2 the southbound and northbound interfaces are part of the SDN communi-cation protocol, where the southbound interface allows communication between controllers and switches, and other low-level network com-ponents. The northbound interface enables communication with high-level components [6]. The next building block of KDN is SDN, which enables the global network view, network programmability functions, and flexibility to manage the network. The combination of network analytics and SDN provides a foundation for the KDN paradigm. How-ever, an ML algorithm will be the heart of KDN, meaning that an ML technique can provide an efficient and optimized strategy to operate the network autonomously.ML is a key component of the KDN paradigm that provides a solution given the telemetry data.

One of the indications of the KDN concept is the open radio access network (O-RAN). O-RAN is an emerging technology that enables service heterogeneity, on-demand service deployment, and simulta-neous coordination of heterogeneous devices [7]. Over the past few

2

*S. Ashtari et al.*  *Array 14 (2022) 100136*



**Fig. 3.** The paper structure.

KDN, there is no survey that comprehensively studies applications of ML in different layers of wireless networking with their relationship and usage in the KDN architecture. Moreover, there exist some limitations in recent works and the overall challenges in applying KDN, which also need attention for KDN to be fully functional. Furthermore, the benefits of having a KDN in the network are clearly stated in the pro-posed research studies. For instance, the authors of [4,15] investigated the general advantages and use cases of the KP and introduced the KDN switch operator along with its architecture. More concretely, the work in [12] examined deep reinforcement learning (DRL) for QoS-aware routing. The literature in [13] proposed a solution for mitigating data center congestion issues by deploying KDN in the network. Such works demonstrate the benefits of adapting knowledge before making decisions in wireless networks.

The motivation of this study comes from the lack of intelligence and automated networks. In the near future, network congestion and smart applications will force the network to have self-managing systems. Con-sidering the shortcomings of 5G and the continuous research activities in wireless networks to enable intelligence and automation in wireless networks, specifically in 6G, a more structured and comprehensive survey is required to understand the potential advantages of KDN-based ML networks. Specifically, we first introduce three possible KDN architectures adapted in wireless networks, including centralized, dis-tributed, and hybrid. Then, we concentrate on the latest achievements in applications of ML from the MAC layer up to the application layer, including resource management, networking, mobility management,

3

*S. Ashtari et al.*  *Array 14 (2022) 100136*

Where, packet-level programming is used to perform ML algo-rithm on collected data and apply the ML-based instructions after the learning procedure.

(6) An overview of supervised learning (SL), unsupervised learn-ing (UL), reinforcement learning (RL), transfer learning (TL), and neural networks (NNs) with basic principles and common applications in wireless networks is provided.

(7) A thorough review of the applications of ML within the KDN paradigm is presented, covering resource management, network-ing, mobility management, and localization. This survey focuses on the MAC layer, network layer, and application layer (the PHY layer is outside the scope of this study). To cover every aspect of a wireless network, resource management is broken down into resource allocation, power management, QoS, base station (BS) switching, cache, and backhaul management. Networking appli-cations are further classified into route selection, clustering, user association, traffic classification, and data aggregation. Mobility prediction and handover management lie within the mobility section, and finally, indoor localization.

(8) Finally, the terms and conditions associated with ML algorithms for generating knowledge and motivations for applying KDN were identified. Then, a summary of the challenges and current issues in KDN is discussed, followed by the conclusion.

The remainder of this paper is organized as follows: In Section 2 KDN, the enabling technologies for KDN are explained, such as SDN, network telemetry, and ML. Then, in Section 3, SDN and KDN are compared, and the potential architecture of the KDN is illustrated. In Section 4, the southbound interfaces for packet-level programming are thoroughly studied, OF and P4, in addition to some of the disadvan-tages of current southbound interfaces and applications of P4. Section 5 describes some of the popular ML techniques utilized in KDN with a brief introduction of their applications in wireless communication networks. In Section 6, a complete study of resource management, networking, mobility management, and localization is presented. Later, in Section 7, terms and conditions associated with ML algorithms for generating knowledge are discussed, including identifying suitable ML algorithms, implementation complexity, time consumption, suit-able dataset availability for training, and suggestions for choosing ML algorithms. Moreover, in Section 8, motivations for applying KDN are presented, including a lack of network knowledge, self-organizing networks, achieving reliable and high performance, learning patterns and predicting the future, ability to handle complex problems, and avoiding a wrong decision, followed by challenges and open issues in KDN in Section 9. Finally, Section 10 concludes the study. Fig. 3 depicts the survey’s structure as a graphical representation for a better understanding of how the survey is shaped and analyzed according to each section. Further, for readers’ convenience, Table 1 lists the abbreviations used throughout this survey.

**2. The road to knowledge-defined networking**

It is estimated that by 2030, the number of connected devices in cellular networks will reach 100 billion [16]. Furthermore, the majority of network occupancy is due to the demand for high-definition video data, which leads to massive data traffic [17–19]. Hence, many existing algorithms cannot process traffic flows, resulting in a loss of informa-tion. Moreover, algorithms are incapable of offering optimum system performance when the network environment is dynamic and random. Therefore, these algorithms are unable to meet the requirements of the 6G cellular networks. To overcome these problems and achieve better performance, researchers have developed optimization methods to attain effective solutions to get closer to optimal and suboptimal performance. However, many studies presume a static network environ-ment rather than considering the random nature of networks [20,21]. Additionally, traditional centralized algorithms for network manage-ment and simultaneous collection of global data are affected as the

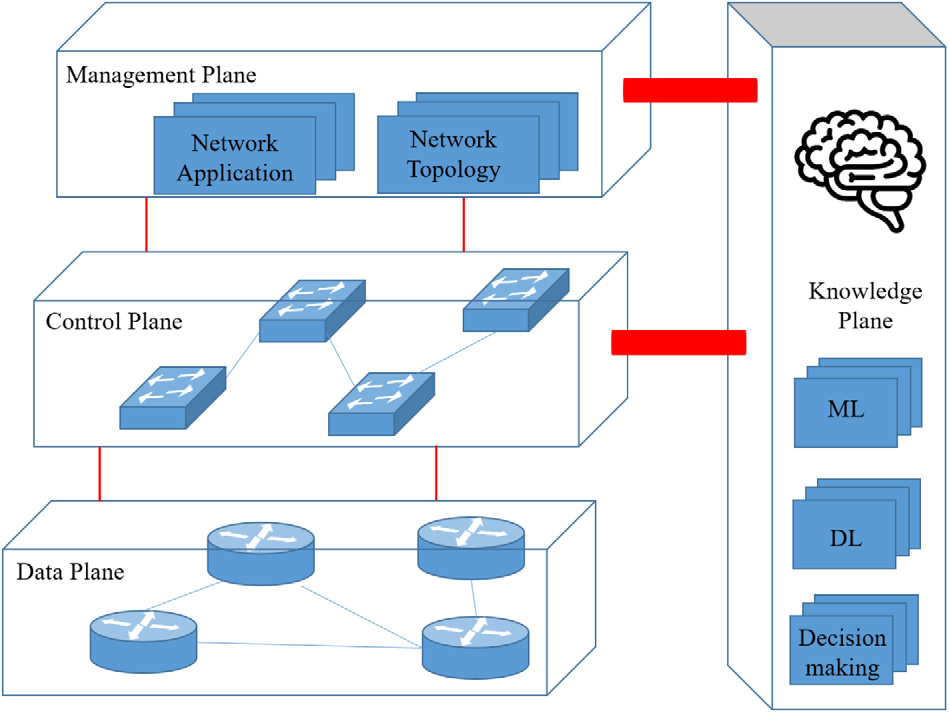
4

|  |  |  |  |
| --- | --- | --- | --- |
| *S. Ashtari et al.* |  |  | *Array 14 (2022) 100136* |
| **Table 1**  Abbreviations. |
| 5G  6G  AI  ANN  AoA  API  ASIC  BBU  BM  BS  CHO  CNN  CP  C-RAN  CRE  CRN  CSI  D2D  DCN  DL  DNM  DNN  DPDK  DPPO  DQN  DRL  DSL  DT  ECR  ELM  eMBB  eNB  ESN  ETL  FBS  FC  FFNN  GF  gNB  GPU  HAL  HetNet  ICIC  IDE  INT  IoT  IP  ISP  IVN  JSON  KDN  KD-NO  KNN  KP  LFU  LOS  LRU  LTE  LTE-U  M2M  MAB  MAC  MBM | 5th generation  6th generation  artificial intelligence  artificial neural network  angle of arrival  application programming interface  application specific integrated circuit baseband unit  behavioral model  base station  conditional handover  convolutional neural network  control Plane  cloud radio access network  cell range extension  cognitive radio network  channel state information  device-to-device  data center network  deep learning  deep neural model  deep neural network  data plane development kit  distributed proximal policy optimization deep Q-network  deep reinforcement learning  domain-specific language  decision tree  energy consumption ratio  extreme learning machine  enhance mobile broadband  evolved node B  echo state network  extract-transform-load  femto-base station  femtocell  feed-forward neural network  gradient follower  gigabit node B  graphics processing unit  hardware abstraction librar  heterogeneous network  inter-cell interference coordination  integrated development environment in-band network telemetry  Internet of things  Internet protocol  Internet service provider  intelligent vehicular networks  javaScript Object Notation  knowledge-defined networking  knowledge defined-network orchestrator k-nearest neighbors  knowledge plane  least frequently used  line-of-sight  least recently used  long-term evolution  LTE-unlicensed  machine-to-machine  multi-armed bandit  medium access control  modified Bushand Mostelle | MBS  MDP  ML  mMTC  mmWave  MPLS  MRE  MS  NE  NFV  NLOS  NN  OF  ONF  O-RAN  OSPF  OVS  PBS  PCA  PISA  POF  PU  PVS  QoE  QoS  RAN  RB  RIC  RL  RNN  RRH  RRU  RSRP  RSRQ  RSS  RSSI  RSU  SBS  SCN  SDN  SINR  SL  SMU  SNR  SOM  SON  SSU  SVD  SVM  SU  TL  TLV  TTP  UAV  UDN  UE  UL  V2I  V2V  VANET  V-RAN  WMMSE  WSN | macro-base station  markov decision process  machine learning  massive machine type communication millimeter wave  multiprotocol label switching  modified Roth–Erev  mobile station  Nash equilibrium  network function virtualization  non-line-of-sight  neural network  open flow  open network foundation  open-radio access network  open shortest path first  open vSwitch  pico-base station  principal component analysis  protocol independent switch architecture protocol oblivious forwarding  primary user  POF switch  quality of experience  quality of service  radio access network  resource block  RAN intelligent controller  reinforcement learning  recurrent neural network  remote radio head  remote Radio unit  reference signal received power  reference signal received quality  received signal strength  received signal strength indicator  roadside unit  small base station  small cell network  software-defined networking  signal-to-interference-plus-noise  supervised learning  spectrum selection utility  signal-to-noise ratio  self-organizing map  self-organizing networks  spectrum selection utility  singular value decomposition  support Vector Machine  secondary user  transfer learning  type-length-value  table type patterns  unmanned aerial vehicle  ultra dense networks  user equipment  unsupervised learning  vehicle to Infrastructure  vehicle-to-vehicle  vehicular ad hoc network  virtual radio access network  weighted minimum mean square error wireless sensor network |

the network performance once the features are fed to the algorithm. UL assists the network operator by following the correlations in the data. For example, ML may predict the mobility effect on a user’s network link. Moreover, in RL, the learning algorithm will discover the best action that leads to an optimal configuration in a network. As RL adapts to the environment, it eventually determines the target policy. This ap-proach returns the optimal action based on the target strategy and then applies them to the KDN controller to enable the best configuration. RL provides extensive benefits for resource management [27]. As a result of network softwarization, network telemetry, and integration of ML in the KP, the foundation of KDN has emerged, which will provide

5

*S. Ashtari et al.*  *Array 14 (2022) 100136*



**Fig. 5.** KDN architecture.

architecture enables programmability, more straightforward configura-tions, and network management. To go one step forward in software-defined cellular networks, KDN introduced intelligence in addition to programmability and centralized control of SDN.

*3.1. Software-defined networking versus knowledge-defined networking*

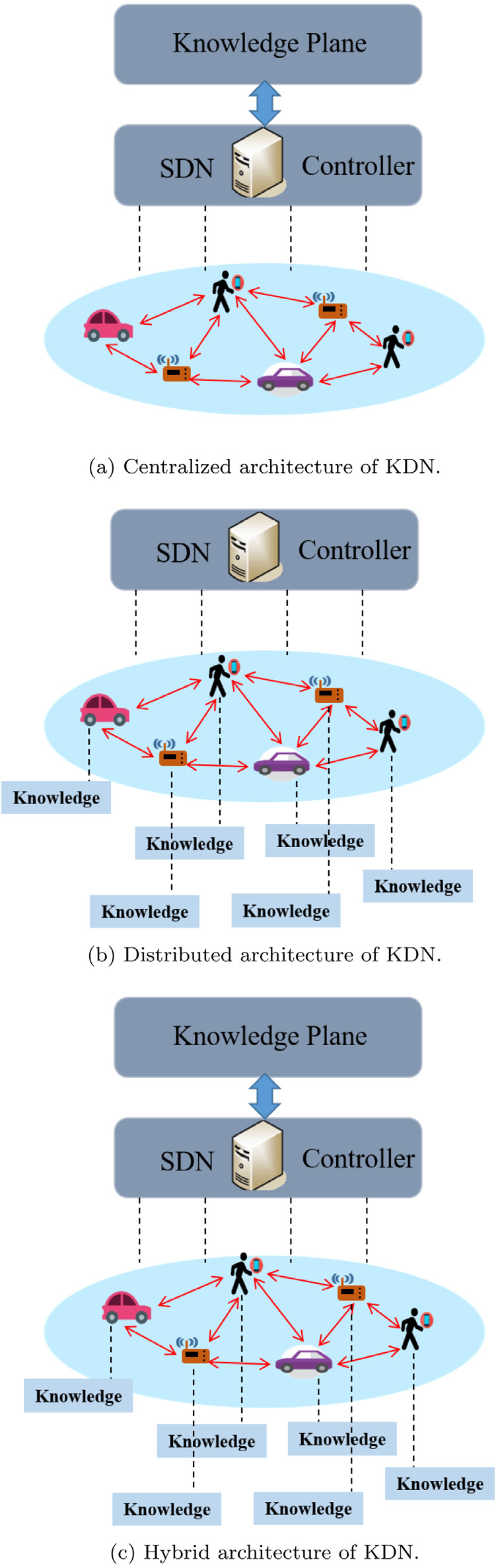
The integration of the KP in SDN is a new concept called knowledge-defined networking (KDN). The concept of KP is to add one more plane to the traditional two planes of SDN. This new paradigm incor-porates SDN, data analytics, and ML. The KDN paradigm has several advantages: first, it has a global view of the network, and second, it enables telemetry data to be collected by the management plane to transform the data into knowledge via ML. The knowledge will later turn into decisions by nodes to achieve efficient network operations [18,19]. The benefit of having the KDN over traditional networks is that it automatically operates based on the knowledge obtained from the network. Fig. 5 illustrates the KDN architecture, including the data plane, control plane, management plane, and KP.

The data plane in KDN is responsible for forwarding, dropping, processing, and packet modification. This layer works precisely like the data plane in SDN, where it consists of physical and virtual device elements. This layer operates unaware of the rest of the network and relies on the instructions and control rules coming from other planes. The control plane exchanges information and updates the data plane processing and matching strategy rules. The logically centralized controller exchanges data and updates policies using a southbound application programming interface (API). The controller gathers the data and network state from the data plane and updates the flow tables to perform actions. In KDN, the data are also utilized to allow the KP to know which appropriate action is required. Then, the controller receives an action from the KP and updates the flow tables accordingly. These actions are usually used for forwarding and routing packets, while the data plane is populated.

The management plane facilitates network topologies, support ser-vices, and configuration of the network devices. This layer must ensure that the network operates fully with the maximum performance. This functionality of the network in KDN is handled by the centralized controller as well, being responsible for monitoring the data plane and observing network analytics. The network analytics will then be col-lected and stored as a network state and telemetry. This information is also monitored by the KP for possible updates of the network topology. KP is the brain of the architecture and responsible for modeling network behavior and decision-making. These decisions are made for different applications in the network, such as resource management,

6

*S. Ashtari et al.*  *Array 14 (2022) 100136*



**Fig. 6.** Proposed KDN architectures.

*4.1. The road to software-defined networking*

There are two main problems with the traditional IP network, mainly the complexity and difficulty of network management [30]. The complexity is due to the configuration of individual network devices, such as routers and switches. The difficulty in managing the network is due to the close attachment of the data and control planes. Hence, the

7

*S. Ashtari et al.*  *Array 14 (2022) 100136*

forwarding tables of switches, and it can also add and delete forwarding entries of almost 50 different header types. Accordingly, the vendors dictate the control plane to which the header they support using the table-type patterns (TTPs) provided by ONF. To manage large-scale networks, SDN operates with OF standards to provide simpler configuration options. While SDN separates the control and data planes, OF only applies a fixed set of protocols to populate the rules in the data plane by using the control plane. However, these protocols are understandable by the fraction of available hardware routers and switches [40]. Therefore, the OF does not control the switch’s behavior of the supported protocols. It only provides a way to populate the tables in the switch. Moreover, the current OF has specific protocol headers for forwarding a packet. Forwarding a packer requires forwarding tables, which are known as flow tables in the OF standard. These tables define how a frame is forwarded out of a switch, where the tables operate by matching specific header fields. These tables have grown from 12 to 41 fields in just a few years, presenting a huge challenge. This increased the complexity of the specifications without providing any flexibility by adding new headers. P4 is a tool that reduces the complexity of OF. The necessity of P4 alongside (or op-erating separately without OF) OF for improving network functionality is promising [5,41]. Fig. 7 depicts the hybrid network utilization of the OF and P4 [40].

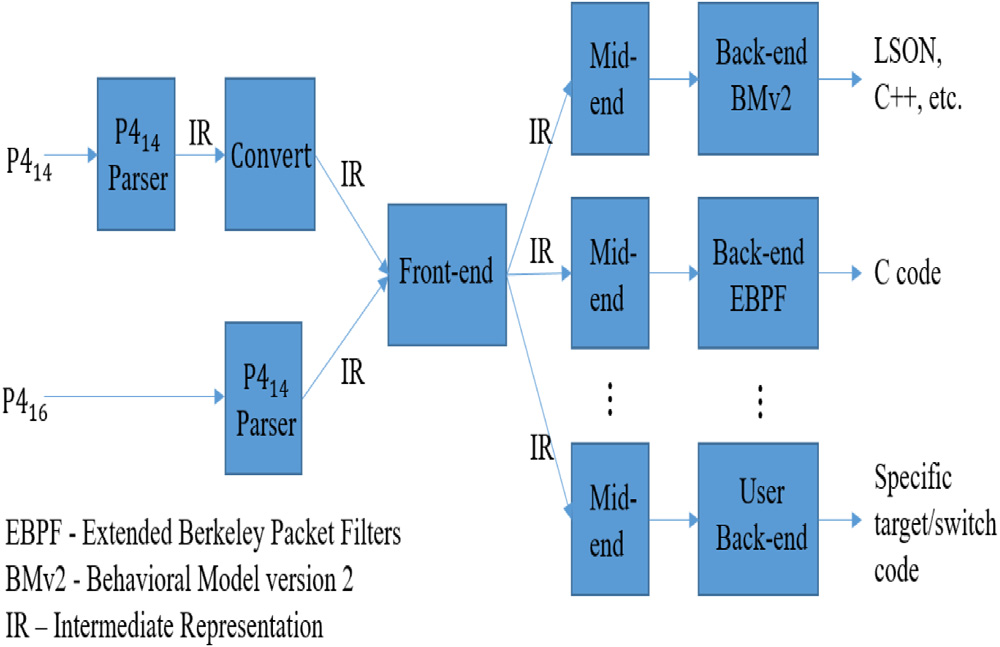
*4.2. P4*

The programming protocol-independent packet processor language is abbreviated as P4. In 2013, the P4 high-level language for pro-grammable protocol-independent packet processors was developed through the collaboration of Barefoot Networks, Intel, Stanford Uni-versity, Princeton University, Google, and Microsoft [5]. P4 enables the programmability of the data plane and allows switches to process the packet. Hence, vendors and enterprises will be able to develop their own application-oriented software for a programmable switch chip, resulting in several benefits to the network, such as reducing the packet processing time, modifiable packet headers, and switch protocol independence. These programmable switch chips are based on a protocol-independent switch architecture (PISA).

P4 was first introduced in 2014 to address the limitations of the data plane by providing flexibility in programming the data plane in network switches that support OF standards [42,43]. The original P4 language was called *𝑃* 414 which only assumed distinct/specific device capabilities and was able to program a subset of programmable switches [40]. With the evolution of the language, *𝑃* 416 brought new functionalities, such as stable language definition, supporting many switches, and removed the assumption of device capabilities. Generally, P4 processes packets by the pipeline of networking elements, including switches, routers, etc. It is based on a fundamental forwarding model that uses a parsing of the packages and applies match+action table recourses to ingress [41], where the abstract packet forwarding model is shown in Fig. 8. The parser is the process of identifying the headers from each incoming packet. As the header is identified, a lookup func-tion is performed to find the appropriate match according to the header fields, and then applies the action corresponding to the match within the table. The P4 programming language focuses on the specifications of these two procedures and the control flow through the pipelines. These specifications are controlled by programmers who write and execute P4 files. Translation tools are required to execute a file from the P4 program. There are two options for translation: one is by the interpreter on every cycle of execution, and the second is by a compiler once the program is executed. Both approaches possess advantages; while the former has the ability to minimize the error in optimization operations, the latter method can reduce the translation time of each cycle between the development and runtime of the program. The P4 compiler consists of different parts, as depicted in Fig. 9. Some of the most commonly used compilers available to execute a P4 file are summarized as follows:

8

*S. Ashtari et al.*  *Array 14 (2022) 100136*



**Fig. 9.** Inside P4 compiler.

PISCES is a software switch in which forwarding behavior is obtained by high-level DSL or P4. PISCES is derived from Open vSwitch (OVS) and is configured by P4. In this protocol, the pro-grammer must specify how to process the packets. For instance, if in P4 we assign PISCES to process IPv6 packets, the program-mer needs to introduce IPv6 packets, including the format and fields of the IPv6 headers. PISCES brings several benefits, such as a personal protocol header, adding/removing a standard header, and easy to add new features [46].

Most of the available P4 programming translators are compilers. However, there exists research for developing interpreters [47], which still needs further attention. Currently, most of the focus is on pro-grammable hardware switches with PISCES that are compiled to a customized software-based POF switch (PVS), where POF stands for protocol oblivious forwarding.

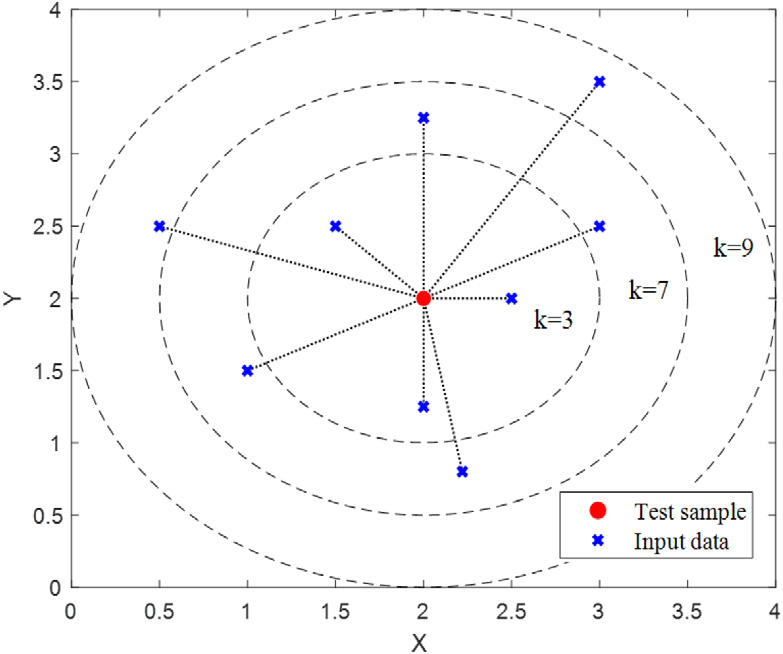
Therefore, P4 enables a new functionality for controlling the for-warding behavior of the switch by populating the tables. Moreover, P4 compilers allow us to use different APIs for switch chips. For most of the current switches in the market, this chip has an on-board invariant programmed module. These APIs are auto-generated by the P4 compilers to populate the switch tables. The new capability to utilize programmable switches will change the way switches operate, but never the less OF is still useful in networking for the old fixed-function switches. Therefore, a program called openflow.p4 has the ability to program switch chips with the support of OF. Hence, P4 and OF can work together for networks such that P4 is the language and OF is the program. Additionally, while OF was essentially designed for SDN networks where the control and data planes were separated, P4 was designed to program the behavior of the switch or the router with no restriction on whether controlled locally by a switch operating system or automatically by an SDN controller. Network information can be obtained using both OF and P4, and can be processed in the KDN to inform the data plan for the new policy. Researchers can choose their own plan toward how and which protocol suits their work, and based on that, they can identify the appropriate translation tools for packet-level programming.

**5. Overview of machine learning techniques**

Artificial intelligence (AI) is a progressive branch of computer sci-ence that deals with automation across various fields, and ML is an application of AI. ML is applied to developing systems to learn from patterns and data without explicitly being programmed [48]. The KDN architecture requires adapting an ML technique to optimize and create intelligence for the network. ML applications have been successfully utilized for network analysis, online customer support, search engines,

9

*S. Ashtari et al.*  *Array 14 (2022) 100136*



**Fig. 10.** An abstract model of the KNN algorithm.

dataset in the feature space. The *𝐾* value is the hyperparameter and should be selected carefully, but the optimal K-value can be chosen after running the KNN several times with different K-values. The resulting KNN outputs can be used for both classi-fication and regression. In classification, the basic principle is to decide the category of a test point based on the decision made by the majority votes of K nearest neighbors. In the regression, the test point is categorized in a class by the average values of its K nearest neighbors. Additionally, this method can be adapted by assigning weights to neighbors such that the nearest points are more involved than the distant points. The KNN method is illustrated in Fig. 10. A more detailed explanation of KNN can be obtained in [55].

(5) *Decision Tree*: One of the most common ML algorithms is DT, which is a statistical model for classification problems. This ML technique is used to classify data and formulate a dataset in a hierarchical structure, such as a flowchart representation. This flowchart is usually a tree-based structure, and the algorithm starts from the root and classifies the dataset until it reaches the leaf. The representation of one class DT is as follows: Each split of the tree is based on a condition on a particular attribute, and leaves are the classes. For a simple example of DT, the iterative Dichotomiser 3 (ID3) is explained as follows:

(i) First, it assigns the original dataset (*𝐷*) to the root node. (ii) Iteratively calculate the entropy *𝐻*(*𝐷*) or the information gain *𝐼𝐺*(*𝐷*) of every attribute of set *𝐷*.

(iii) Then, it partitions set *𝐷* into subsets using the smallest resulting entropy from step ii.

(iv) Making a DT node consisting of that attribute.

(v) Recurse on *𝐷* using the remaining attributes.

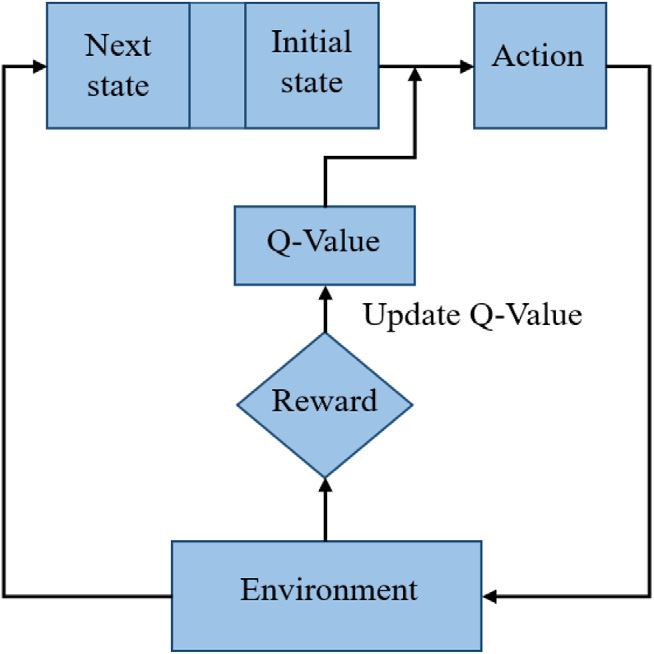
The DT method works for both regression and classification problems with many advantages, such as ease of interpretation, requires fewer data points to learn, and works well with a large dataset. However, minor changes to the large dataset require a new training sequence, and overfitting is the algorithm’s limita-tion. For implementation and further information, refer to [55]. Fig. 11 provides a simple example of a decision tree.

*5.2. Unsupervised learning*

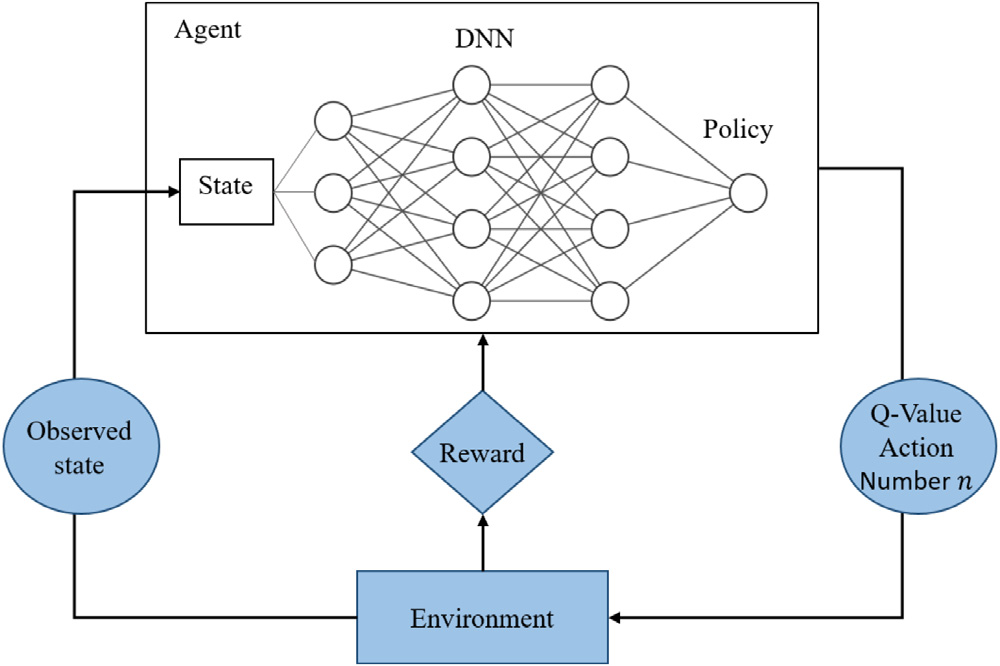
In unsupervised learning (UL), the input data are unlabeled data, where the algorithm has to find patterns and hidden structures to learn a useful function. The enormous data collection by devices and sensors results in a lack of labeling due to the unavailability of funds to pay for manual allocation or the nature of the data itself. The UL is extensively used in clustering and data aggregation [56]. The following algorithms are the most common UL techniques:

10

*S. Ashtari et al.*  *Array 14 (2022) 100136*



**Fig. 13.** An abstract view of Q-learning method, where every action taken in any state will be observed by the environment and generates a reward to provide insights about how good was the agent’s action.



**Fig. 14.** The procedure of deep reinforcement learning.

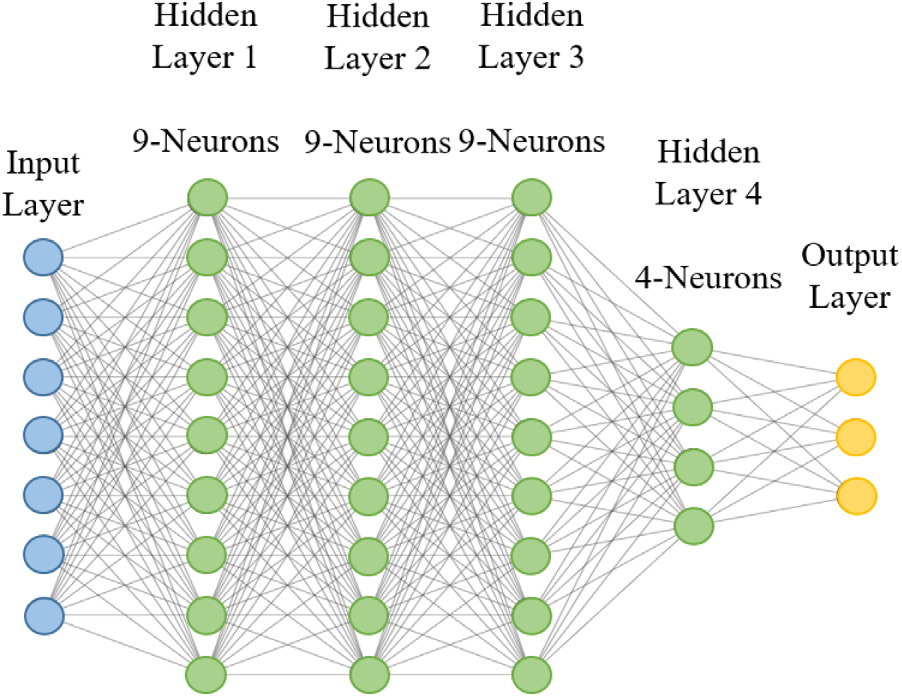
(1) *Q-learning*: Q-learning is the most popular RL technique that uses Q-value and Q-function to find optimum action policies. Specifi-cally, the agent interacts with the current given environment to continuously learn the Q-values and maximize this value. This algorithm starts with an initial state and an action, followed by the epsilon-greedy policy. For each performed action, the Q-value is learned through the optimal (greedy) policy, enabling the agent to take any action based on the largest Q-value under the current state. Details of the Q-learning method can be found in [26,60]. Moreover, fuzzy Q-learning is used to deal with con-tinuous state spaces with a certain number of given rules [61].

A flowchart of the Q-learning method is presented in Fig. 13. (2) *Deep Reinforcement Learning (DRL)*: Deep RL is a combination of deep learning and RL, which correlates the value function with corresponding actions and states [62]. DRL uses these two principles to approximate the optimal Q-values. Moreover, DRL techniques facilitate an NN with an RL architecture to enable agents to learn the best action in a specific environ- ment. DRL is mostly used in complex decision-making tasks with unstructured environments and can handle large datasets.

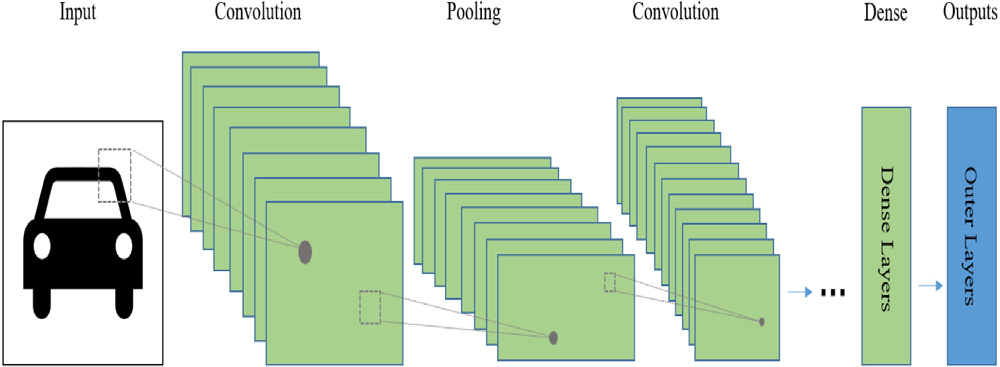
Recently, DRL has made great strides in vehicle-to-vehicle (V2V) communication [63], wireless communications [64], and video games [65,66]. More details on DRL can be found in [67] for the readers. An overview of this concept is shown in Fig. 14.

11

*S. Ashtari et al.*  *Array 14 (2022) 100136*



**Fig. 17.** The architecture of DNN.



|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Fig.**  **18.** An  architecture. | overview | of | training | process | of | a | convolutional | neural | network |

when the reward is immediately calculated locally; for instance, the transmission rate between the transmitter and receiver. This algorithm is often referred to as an equilibrium concept in game theory in many survey papers, such as coarse correlated equilibrium and logit equilibrium. The concept of this algorithm is illustrated in Fig. 16.

*5.4. Neural network*

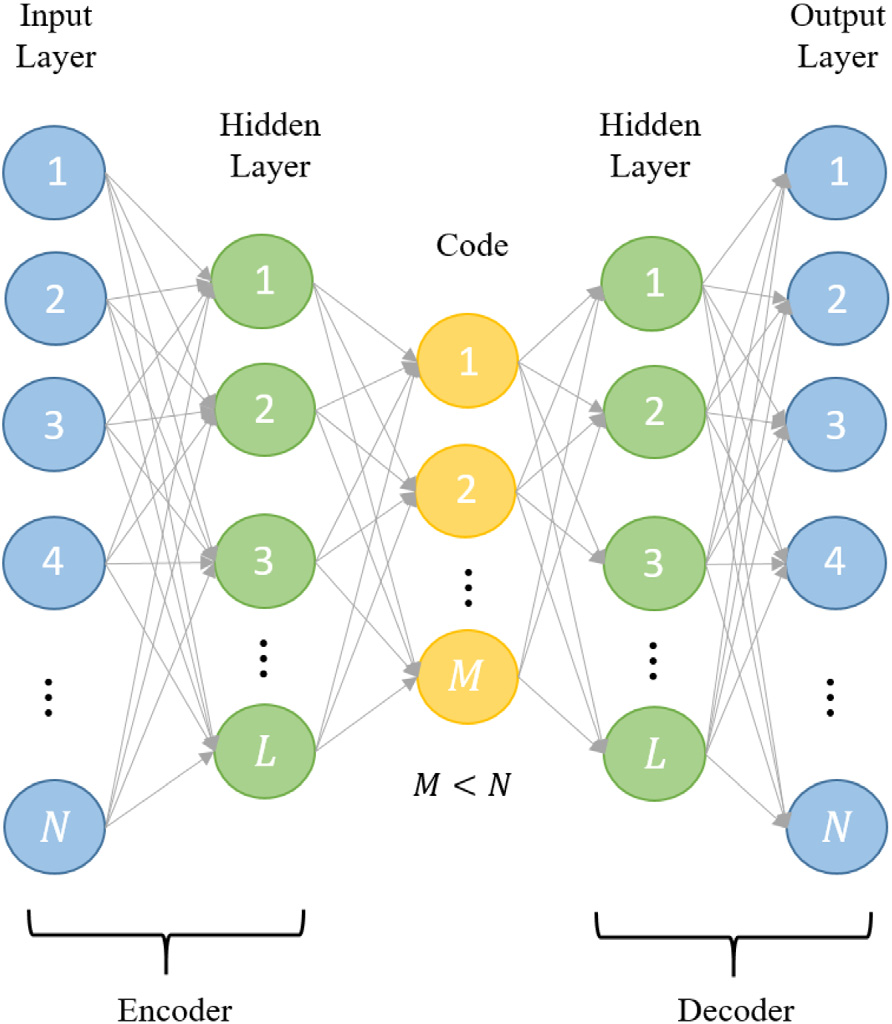
A neural network is a type of artificial intelligence for information processing that imitates the human brain. The neural network structure consists of thousands of closely connected, simple processing nodes. Neural networks are organized into layers, and in each layer, many nodes move the data. With the development of graphics processing units (GPUs) to accelerate the processing time, NN has attracted con-siderable attention from researchers and companies [72]. The following NN techniques are explained in this study.

(1) *Deep Neural Network (DNN)*: As shown in Fig. 17, the compo-nents of the structure are densely connected by the neurons in the network layer. Each neuron in a layer is connected to the rest of the neurons, resulting in the structure of the DNN. Each neuron corresponds to a weight for the input and an activation function for the output. The input data are transformed from layer to layer, with no direct connection between the two non-consecutive network layers. The main advantage of a DNN is that it can automatically deduct and tune the features to obtain the desired output. For the optimization of network parameters, DNN uses backpropagation (one of the most popular learning techniques for multilayer neural networks) and various gradient descent algorithms, such as Adam and Momentum [73].

(2) *Convolutional Neural Network (CNN)*: The convolutional neural network is a class of artificial NN that was developed during the

12

*S. Ashtari et al.*  *Array 14 (2022) 100136*



**Fig. 20.** Illustration of the structure of an autoencoder.

assist learning more robustly. For deep learning and RL, knowledge can be defined as weights and Q-values, respectively. For instance, when deep learning is adapted for resource management, the ML process can use the weights that have been trained for other resource management tasks as the initial weight. For example, a similar network with the same number of nodes and similar behavior with a trained ML and a fully operational resource management policy can provide knowledge for other similar networks. Moreover, in RL, the Q-values learned from an environment can be used in a similar environment to make better decisions during the initial stage of learning. The specifications for utilizing TL with RL can be further studied in [77]. While there are advantages of having prior knowledge for learning a pattern, there are some limitations and negative impacts on the performance, which needs further attention that exceeds the scope of this study.

Most of the above ML techniques can be deployed in several net-work applications for automation and optimization. ML algorithms provide information and knowledge for different tasks. In the next section, the application of ML in KDN for wireless networks is dis-cussed. First, a general introduction to each application was explored. Second, the important characteristics of each surveyed study with supporting ML algorithms were investigated. Finally, a possible KDN architecture that can be adapted for this specific study is presented. This study was expanded over various parts of a wireless network, including resource management, networking, mobility management, and localization. Therefore, it is crucial to identify the studies within these parts that can be potential use cases in generating knowledge in the KDN paradigm. These studies are the building blocks for achieving a fully knowledge-based network in future wireless networks. Table 2 presents the selected number of surveys with different applications supporting various parts of wireless networks.

**6. Application of machine learning for knowledge defined net-working**

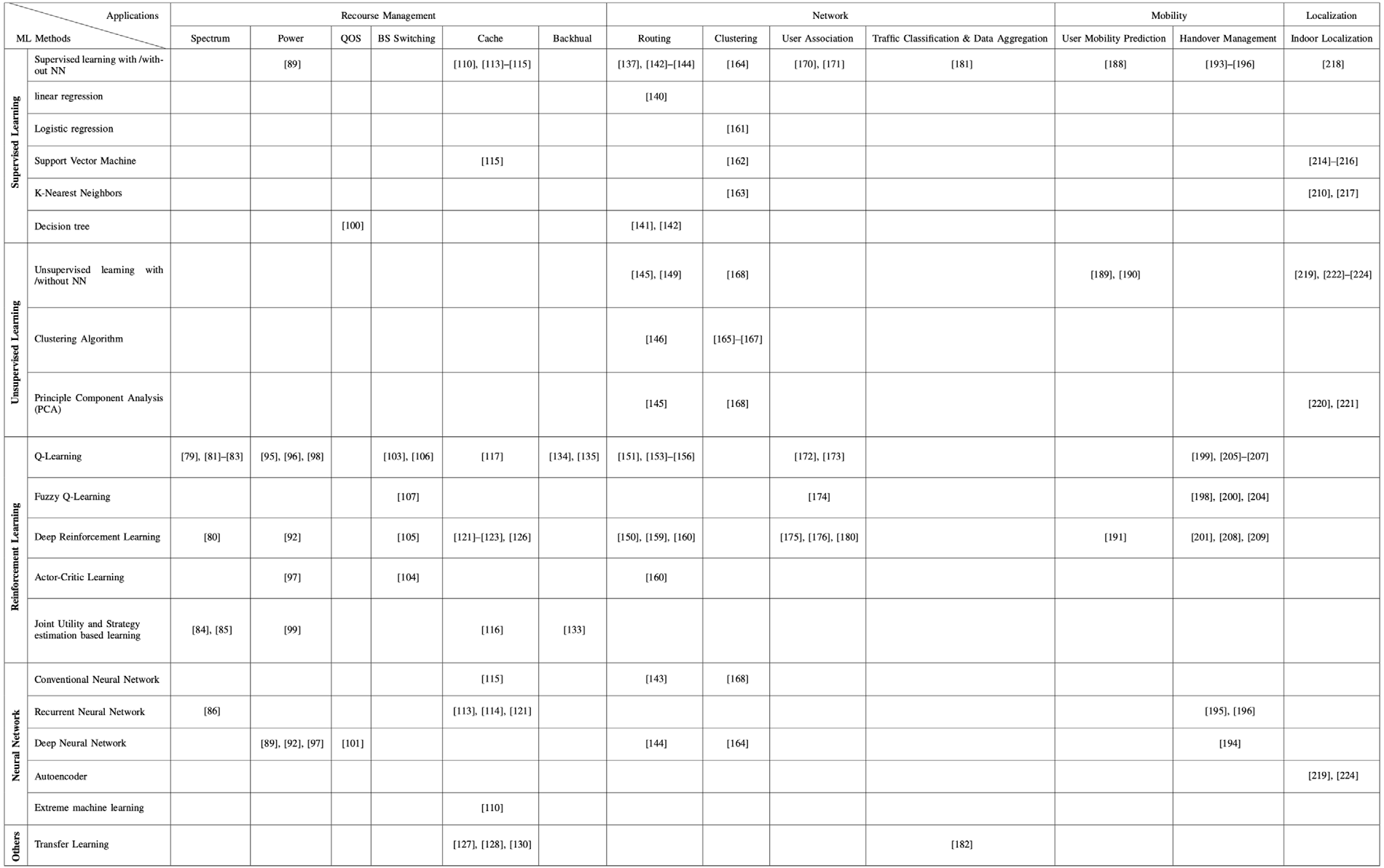
Owing to the exponential growth of data traffic, wireless networks will require advanced technical solutions in the near future. As a result, the traffic load among BSs, the complexity of wireless chan-nels, the emergence of self-driving vehicles, device-to-device (D2D),

13

*S. Ashtari et al.*  *Array 14 (2022) 100136*

**Table 2**

The application of machine learning methods on different areas of telecommunication.

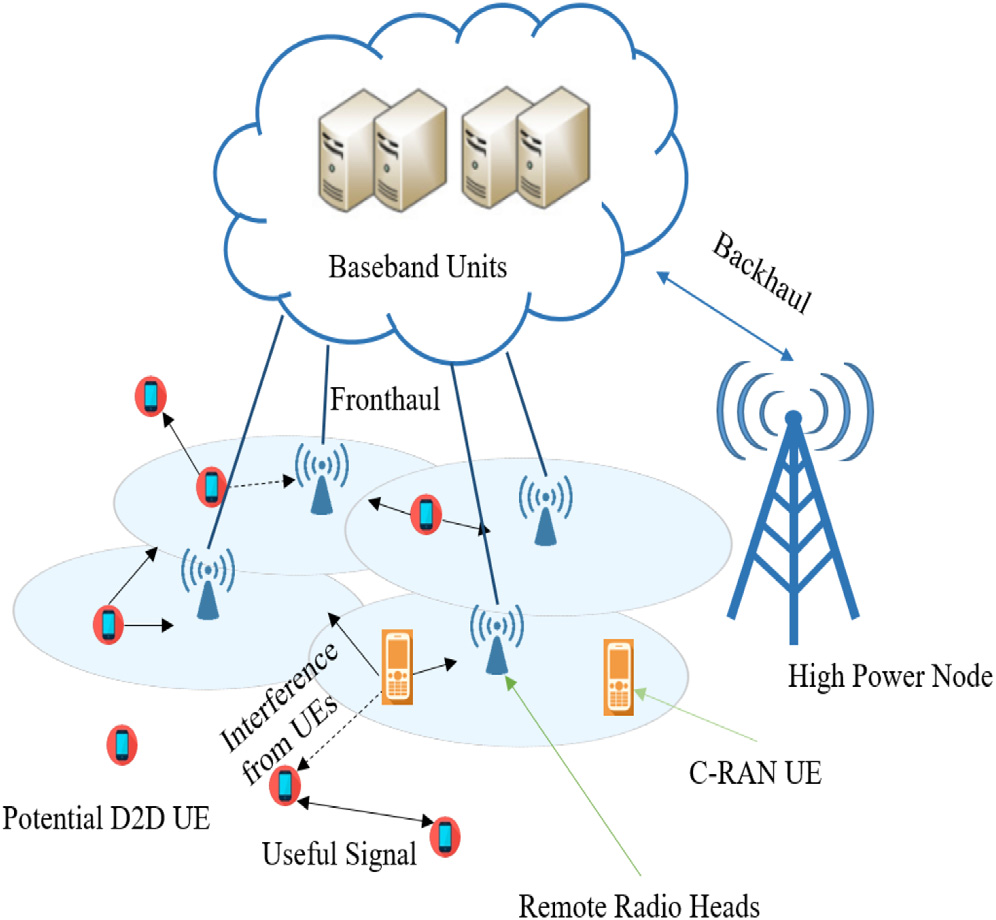


each user receives an acknowledgment of whether the packet has been received. Subsequently, each user maps its current state to an action based on the spectrum access obtained by a trained deep Q network. Their algorithm proves that users can learn the best policy based on the objective network utility. The proposed algorithm provides twice the channel throughput compared to slotted-aloha with optimal probabil-ity, which requires full knowledge of the number of users. Hence, this suggests that the output of the Q-learning algorithm in their study can be used as knowledge for KP in wireless networks, where the number of nodes is not visible by the controller, and the controller needs to perform the best action and allocate an efficient channel to users. The proposed algorithm suggested a distributed adaptation of transmission parameters, where the knowledge obtained from a network by users can be used in a distributed manner in the KDN architecture.

As data traffic increases in mobile networks and fixed spectrum allo-cation of operators becomes a major problem, inter-operator spectrum sharing has been proposed as a solution. This solution brings benefits but also introduces new challenges. The authors of [81] introduced a new paradigm called inter-operator proximal spectrum sharing (IOPSS) to intelligently offload users from an overloaded BS to the neighbor-ing BSs based on the spectral proximity. A Q-learning framework is equipped with each BS to dynamically determine the network load and efficiently utilize its spectrum in a self-organizing manner. The state of the system depends on the network load experienced by the BS and is defined as a discretized value. At the same time, the action of each state is determined by the tuple of spectral sharing parameters associated with each neighboring BS. The spectral sharing parameter for a BS includes the required spectrum resources (where spectral resources in this study are considered as resource blocks (RBs)) from a neighboring BS, the probability of a user that the neighboring BS can support with the strongest value of the signal-to-interference-plus-noise ratio (SINR), and the fraction of RBs that need to be reserved as requested. Simulation results demonstrate that the IOPSS-based Q-learning algorithm can intelligently offload users from congested BSs

14

*S. Ashtari et al.*  *Array 14 (2022) 100136*



**Fig. 21.** An abstract system model of D2D enabled C-RAN.

theory, optimal channel allocation is obtained. The simulation results suggest that the proposed policy outperforms other policies in the literature based on temporal–spatial spectrum reuse. The study in [82] can assist in knowledge extraction and maximize the network benefits for a distributed KDN architecture. In contrast, in [83] the spectrum allocation problem is formulated into centralized and distributed parts, which can best fit our hybrid architecture in KDN.

To improve resource management, D2D communication was intro-duced to decrease the load on the BS. In [84], the authors proposed a distributed approach for resource allocation and communication mode selection for potential D2D pairs using a joint utility and strategy estimation algorithm. The proposed system model is shown in Fig. 21 consisting of potential D2D pairs, BS, C-RAN, and user equipment (UE) that are unable to perform D2D communication. They investi-gated D2D-enabled C-RANs to improve the spectral efficiency using RL techniques. The action of the learning algorithm consists of the communication mode selection and subchannel allocation for each D2D pair. After each pair selects its action, the remote radio head (RRH) association and power control of the D2D pairs is modeled. Then, based on the reinforcement algorithm, the system obtains the instantaneous utility (known as the received spectral efficiency), which is updated for each action continuously. This study aims to enable D2D pairs to self-optimize their resource allocation and perform mode selection under different practical constraints, including fronthaul capacity and inter-tier interference constraints. The numerical simulations demonstrated near-optimal performance and better spectral efficiency. From the pro-posed method, it is evident that this study best suits the distributed architecture of the KDN, where D2D pairs will maximize their resources using the proposed algorithm.

In heterogeneous networks, the authors of [85] proposed a multi-objective fully distributed strategy based on RL for self-configuration and optimization in LTE small cells. The proposed algorithm uses a joint utility and strategy estimation under QoS constraints to minimize the received intra- and inter-tier interference for femtocells (FCs) and reduce inter-tier interference from FCs to eNBs. Their algorithm utilizes RL techniques to simplify their problem formulation compared to works where global knowledge and complete CSI are unavailable. Hence, we can utilize the RL ability to gather information from the interaction between BSs and users. They identified two sequential learning levels. During the first phase of learning, available spectrum resources were

15

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| *S. Ashtari et al.* | | | |  |  | *Array 14 (2022) 100136* |
| **Table 3**  Knowledge-based strategies for spectrum management. | | | |
| Article | Knowledge objective | | Architecture | ML algorithm | Deliverable | Conclusion |
| [79] | Utilizing optimal  values of Q-table for similar CRNs | | Centralized | Q-learning | Self-tuning spectrum  allocation, near-optimal solution and improving network capacity | Handles high computational complexity, however  performance is highly  dependent on the control parameters |
| [80] | Providing a general real-world solution for spectrum access using a trained  multi-user DQN | Distributed | | Deep reinforcement learning | Better channel throughput compared to slotted Aloha and optimal spectrum  allocation | Offline learning and online real-time spectrum access, decreasing routing overhead but subjected to time  constraint for mobile users |
| [81] | Intelligent user  offloading and  self-organizing  spectrum allocation | Hybrid | | Distributed  Q-learning | Maximizing the user’s QoE | Efficiently share the spectrum between BSs to offload users but challenges of user mobility have not been studied yet |
| [82] | Self-adaptation of SUs based on  trained Q-learning algorithm | Distributed | | Distributed  Q-learning | Achieving NE | Maximizing transmission  efficiency while coordinating interference and suitable for 5G communications but  include redundant information |
| [83] | Efficient channel  reuse for SUs once the optimal policy is learned | Hybrid | | Multi-armed Bandit techniques | Enhancing spectrum  utilization with  temporal–spatial spectrum reuse | The channel allocation policy has fewer regrets than other policies, but it works work static networks |
| [84] | Using the built-in strategy profile for intelligent  communication by using mode  selection (referred to as selecting D2D or C-RANs data  transmission) and resource allocation of D2D pairs | Distributed | | Joint utility and  strategy estimation based learning | optimizing spectral  efficiency | D2D pairs can self-optimize their mode selection and  resource allocation, although it is time consuming |
| [85] | Autonomously  identifying and  optimizing spectrum usage in femtocells using the learning  strategy | Distributed or  Hybrid | | Joint utility and  strategy estimation based learning | Minimizing the  intra/inter-tier interference and increasing the  throughput | Reduces waste of memory and the unnecessary information |
| [86] | Self-organizing  framework to  optimize resource  allocation in both  uplink and downlink LTE networks | | Distributed or Hybrid | Multi-agent  reinforcement  learning based on  echo state networks | Load balancing and efficient spectrum allocation | Choosing the optimal  resources by SBS given  minimum information about users and network |

(ii) *Knowledge derived from reinforcement learning*: Research in [92] introduced a deep Q-learning algorithm for dynamic power al-location based on collected channel state information (CSI) and QoS. Their proposed distributed algorithm is based on the model described in [93], which does not rely on the network size, and it uses a robust technique for the dynamic changes of the network. They considered a single frequency band for transmis-sion with synchronized time slots. A classical power allocation method is utilized in the initial stage of the network. Then, an RL agent interacts with the environment and learns by ob-serving the rewards by trial and error over time [59,94]. To optimize the system and mitigate the problems with Q-learning, a deep Q-network (DQN) is used to estimate the Q-function. Their algorithm showed a fast and suboptimal power-allocation technique for various dynamic wireless networks. This will have a massive opportunity in the future for heterogeneous networks to be implemented and used as prior knowledge and facilitate the users such that optimum power is allocated to all the users quickly and efficiently. This method is more reliable than the other methods because of its response to dynamic changes in the network. A distributed knowledge plan can benefit from this

16

*S. Ashtari et al.*  *Array 14 (2022) 100136*

(with two different carrier frequencies). Therefore, using Q-learning will enable PBS and MBS to self-optimize system per-formance using ICIC. Moreover, a fully automated multi-agent Q-learning technique was developed to facilitate heterogeneous cellular networks and to model the channel and power levels of D2D pairs [96]. Each pair attempts to maximize the value obtained by the difference between the throughput and power consumption cost, which is achieved via the defined SINR con-straint. The proposed model is formulated using a stochastic non-cooperative game, where each pair of devices becomes a learning agent to learn the best policy from locally observed information. Their simulation results showed an acceptable con-vergence rate and near-optimal performance with a few learning iterations. The study proposed in [95] suggests a distributed policy for KDN to incorporate self-healing, self-optimizing, and self-configuration of the network. Similarly, in [96] users act selfishly to choose the wireless channel and power level to max-imize their throughput, which in fact represents a distributed manner for the knowledge plane. However, the information can be used at the centralized controller to make optimal decisions when congestion in the network is high. Hence, both the dis-tributed and hybrid architectures of the KDN are suitable for this study.

In [97], the authors integrated two DRL techniques for power control between primary users (PUs), SUs, and wireless sensors. Their work is based on an asynchronous variant of the actor–critic learning algorithm, where an A3C-based power allocation method and distributed proximal policy optimization (DPPO)-based power control are utilized. These two methods are based on the actor–critic learning mechanism for optimizing power control policies and making spectrum sharing more accurate. The aim is to model the information interaction between users and wireless sensors to learn the simultaneous power allocation scheme and optimize power consumption. In their algorithm, first, each SU gathers information from a centralized controller, and based on the experiences of other SUs, it performs power control strategies and simultaneous power control management policies at the controller. PUs cannot obtain power allocation policies using other PUs. They only adjust their transmission power based on a power-control scheme. Finally, the simulation results indicate that the proposed power control scheme per-forms better than the DQN-based power allocation in terms of power efficiency, spectrum sharing, and network convergence. To allow spectrum sharing, the authors of [98] proposed an optimization-based algorithm for power management at small BSs (SBS) with a Q-learning method to reduce the interference of each RB. Initially, power is randomly allocated by SBS, based on the assumption that the SBS is concerned about maximizing its expected data rate in the long term, Q-learning is used to find the optimal power for SBS. The state of each SBS is a binary value that indicates whether the QoS has been violated, while the action of the learning algorithm is the selection of the optimal power level. The system reward is represented as the instantaneous rate of the SBS. Simulation results show the ability of Q-learning to increase the long-term expected data rate of SBSs. The centralized power control mechanism in [97] represents a centralized KDN, where the latest power control strategies are updated. The power allocation scheme in [98] relied only on local information at the SBSs. Local data are not shared within the SBSs. Hence, the algorithm can be applied to the distributed architecture of the KDN.

In addition, a research study in [99] utilized the RL technique to self-organize the transmission power in femto-BS, pico-BS, and micro-BS. Their goal is to mitigate interference in SCNs and increase the spectral efficiency. Here, the interaction between SBS and macro-BS is modeled as a non-cooperative game, where

17

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| *S. Ashtari et al.* | | | |  |  | *Array 14 (2022) 100136* |
| **Table 4**  Knowledge-based strategies for power management. | | | |
| Article | Knowledge objective | | Architecture | ML algorithm | Deliverable | Conclusion |
| [89] | Real-time power allocation | Centralized | | Supervised learning with DNN | Power control and  interference management | Decreasing the computational cost but cannot be utilized for directional antennas |
| [92] | Adaptable to large networks with  real-world scenarios | Distributed | | Deep Q-learning | Suboptimal power  allocation with faster  convergence compared to WMMSE and FP | Robust to unpredictable  changes of the wireless  medium as well as  delayed/incomplete CSI,  although it lacks global  knowledge that might degrade system performance |
| [95] | Smartly offload  traffic and  autonomous  optimization of cell range expansion  (CRE) in  heterogeneous  networks | Distributed | | Reinforcement  learning | Significant improvement in throughput | Joint power allocation and cell association by achieving target SINR for each UE but some backhaul constraints |
| [96] | Intelligent joint  selection of power level and spectrum channel by D2D  users in a multi-cell network | Distributed or  hybrid | | Multi-agent  Q-learning | Reducing the power  consumption and  maximizing throughput | Near-optimal and fast  convergence for D2D pairs, although it does not support multi-hop D2D |
| [97] | Simultaneous power allocation in  wireless sensor  networks | Centralized | | Actor–critic based  learning with DNN | Efficient power  management with less frequent updates | Guarantees QoS but does not involve mobility |
| [98] | Trained Q function is capable of  adjusting the SBS transmission power | Distributed | | Distributed  Q-learning | Maximize the data rate of SBSs while keeping  acceptable QoS | Optimization of power  strategies for maximizing the long term data rate |
| [99] | Self-organization  and optimization of throughput by the femtocell user  equipment (FUE)  under QoS  constraint | Distributed | | Joint utility and  strategy estimation based learning | Achieving NE with less overhead | High spectral efficiency when utility function is the network performance of FBSs, but the network is static during the learning process |
| **Table 5**  Knowledge-based strategies for QoS. | | | |  |  |  |
| Article | Knowledge objective | | Architecture | ML algorithm | Deliverable | Conclusion |
| [100] | Automatic QoS  prediction for  correction and  suggestion for 5G networks | Centralized | | Decision tree | reliable QoS | A new architecture for QoS prediction and reaction to dynamic changes of network |
| [101] | Efficient QoS  prediction strategy | Centralized | | deep neural network | High QoS prediction accuracy | Proposed algorithm has a quick response-time, which crucial for autonomous  systems |

*6.1.4. Base station switching*   
 One significant difference between 5G and 4G is the network ar-chitecture and deployment of a large number of SBSs. Because of mmWave signals in 5G cellular networks, BSs need to be closer to users to reduce the propagation loss and improve the channel capacity. However, deploying a large number of SBSs comes at a price and a significant increase in the total energy consumption of the wireless network. One of the promising solutions is BS on/off switching, which saves approximately 36 million kWh per year [102]. Considerable effort has been dedicated to finding the best strategy for on/off switching mechanisms in 5G wireless networks. Among them, ML algorithms have attracted attention for their self-optimization and self-management abilities. The extracted knowledge from a trained ML can be used in a centralized KDN architecture to manage the entire network BS on/off switching. Table 6 summarizes the studies surveyed in this section.

18

*S. Ashtari et al.*  *Array 14 (2022) 100136*

active users and SBS. Finally, the reward of the learning algorithm is achieved through a switching action in any particular state. The reward of the algorithm is entitled to energy consumption and transmission gain constraints. The procedure is iterative, and the calculated reward updates the Q-value until it converges. From the simulation results, it is concluded that the proposed HAIIC algorithm minimizes energy consumption. Similarly, studies were conducted, such as the work in [104], which used actor–critic learning to control the on/off BS switching. The RL technique defines the BS switching operation as the action and the amount of traffic load on the controller as the state. The controller decides an action based on the traffic load in a stochastic manner to minimize the overall energy usage. The authors of [103] used the coverage of the BS to switch off the SBSs in their vicinity, which suggests a centralized KDN architecture at the main controller. Similarly, in [104], the overlap between the coverage areas of BSs is considered to turn on/off a BS, which also suggests a centralized architecture of KDN.

Although the studies mentioned above used on/off procedures and significantly decreased power consumption, they lack the on/off state transition between BSs. To this end, the authors of [105] included the transition overhead in the cost function of the DRL-based framework for downlink scenarios in cloud radio access networks. The state in DRL consists of two components: the on/off state of the BS, which is a binary value (0 for off, 1 for on), and users’ data rate demands. The action is taken based on the state of the iteration, which leads to the activation of a BS. The reward is calculated based on power consumption and user demand constraints. Simulation results indicate that the proposed advanced technique can be adapted to dynamic environments and can provide power consumption optimization while satisfying user demands. Moreover, the study in [106] considered the on/off transition state, where it uses the Q-learning method and a novel dual-threshold-based sleep-mode control for SBSs. The use of a dual threshold for controlling the BS sleep mode minimizes energy consumption and avoids frequent BS transitions. There is an upper user and a lower user threshold, which defines the action of the learning algorithm. The state of each SBS is the number of users under coverage. When the number of users in the cell passes the upper threshold, the small cell is switched on, and once the number of users becomes less than the lower threshold, the small cell is switched off. Based on the simulation, the algorithm achieves near-optimal performance by reducing the energy consumption with a minimum BS on/off transition state.

Moreover, a BS active/sleep scheduling scheme is proposed for k-tier heterogeneous networks, guaranteeing coverage, QoS, and through-put [107]. They used a fuzzy Q-learning method to put the BS in sleep, while there was no user to serve or activate a BS once the user was detected in the cell. To save energy, the algorithm uses an optimal sensing probability strategy for user detection. An SBS is in the sleep mode state when there is no active flow, and it randomly activates to scan the coverage area for possible users based on the tuned output sensing probability action. It is observed that the proposed algorithm can efficiently handle user population fluctuations and increase the energy efficiency. All three [105–107] proposed schemes can be used in a centralized architecture of KDN to provide an energy-efficient algorithm to reduce energy consumption.

*6.1.5. Cache management*   
 As predicted by Cisco [108], wireless networks, especially cellular networks, will produce about 30.6 exabytes of data traffic each month. This is due to the proliferation of smart devices and the appear-ance of high-tech applications, such as ubiquitous social networking, augmented reality, and high-definition live streaming. Faced with un-precedented data traffic, intelligent learning-based caching strategies have been introduced to alleviate backhaul traffic and shorten la-tency [109]. In this section, we investigate ML algorithms to assist in creating knowledge in the KDN paradigm for cache management. Table 7 summarizes the studies surveyed in this section.

19

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| *S. Ashtari et al.* | |  |  |  | *Array 14 (2022) 100136* |
| **Table 6**  Knowledge-based strategies for BS switching. | |
| Article | Knowledge objective | Architecture | ML algorithm | Deliverable | Conclusion |
| [103] | Offline trained  solution to minimize energy consumption in heterogeneous  networks | Centralized | Q-learning | Reducing the cumulative energy consumption | Optimal small cell switch  policy but startup energy cost and user mobility are not  considered |
| [104] | The trained data is utilized at the  controller to choose the active BS based on the traffic load in RANs | Centralized | Actor critic based learning | improving the energy efficiency | Robust BS switching solution but eliminated inter-cell  interference |
| [105] | Dynamic resource allocation in cloud RANs | Centralized | Deep reinforcement learning | Minimizing power  consumption | Adaptable to dynamic  environments while satisfying user demands |
| [106] | Using the trained  data for  energy-efficient and QoS-aware in SBSs | Centralized | Q-learning | Minimizing network energy consumption while  avoiding frequent mode  transition in BS switching | Effective multi-objective  energy optimization between BS switching and QoS, no user mobility is considered |
| [107] | Self-organizing the active or sleep  mode of SBS in  heterogeneous  networks | Centralized | fuzzy Q-learning | Improving energy  consumption while  maintaining network  capacity and coverage area | A thorough BS switching  based on user activity  fluctuations, QoS, channel  propagation, and interference |

update scheme among different small cells based on joint utility and strategy estimation. The SBSs can optimize the time-varying caching probability distribution using the received instantaneous utility update. The proposed algorithm is capable of minimizing the service delay when serving user requests at the SBS. Sim-ilarly, Wei et al. focused on distributed caching design at BSs to reduce the traffic load via D2D communications [115]. BSs can cooperate and exchange information regarding their locally missing content from other BSs through the backhaul channel to make the scheme more cost-efficient. A Q-learning algorithm is utilized to improve the system transmission cost and enable D2D devices to offload traffic using cache utilization. The action of Q-learning depends on the adjustments of the cache contents taken in a specific state for observation. The convergence of the proposed cache replacement strategy is tested by the sequential stage game model, meaning that the decision-making process of each BS on cache placement is a cooperative game (every state can be considered as a different stage). The simulation results have shown superior performance compared to conventional strategies, including the least recently used (LRU) strategy [125], least frequently used (LFU) strategy [126] and randomized re-placement strategy [127]. In [114], each SBS optimizes the cache policy via a decentralized cache strategy, which represents a distributed architecture of the KP for future optimization via KDN. In [115] the traffic is offloaded from the cellular chan-nel to the WiFi channel using D2D communication between users, which provides a distributed architecture for users in the network to exchange data traffic.

Cache-enabled D2D communication technology is expected to lower the requested content and congestion at the BS by en-abling devices to request content from nearby users. The authors of [116] focused on joint content delivery policy and cache content placement. Cache content placement determines the amount of traffic unloaded from the BS to the D2D. This study uses RNN methods, specifically ESN and long short-term mem-ory (LSTM), to predict the users’ mobility patterns and content popularity. Therefore, the algorithm realizes where to cache and which content to cache. Once the user’s local cache content cannot satisfy the content request of the user, the user will establish a D2D link with one of the neighboring users. The process of selecting the most appropriate user was performed

20

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| *S. Ashtari et al.* | | | |  |  | *Array 14 (2022) 100136* |
| **Table 7**  Knowledge-based strategies for cache management. | | | |
| Article | Knowledge objective | | Architecture | ML algorithm | Deliverable | Conclusion |
| [110] | Adaptive caching  scheme for cellular networks | Centralized | | Extreme learning machine | Reduces the network  traffic and increases the users’ QoE | The proposed scheme  outperforms industry standard caching schemes |
| [111] | Optimal  cache-enabled  method for UAVs | Centralized | | Echo state networks | Maximizes users’ QoE and minimizes the total  transmission power | The proposed algorithm  locates the UAVs and predicts the content to cache at UAV |
| [112] | Proactive cache  approach in  cloud-based radio access networks | Centralized | | Echo state networks | Maximizes the sum  effective capacity | The proposed method also  decreases the delay and traffic in the network |
| [113] | Applying  intelligence-based content-aware in cellular networks | Centralized | | 3D CNN, support  vector machine and regression model | Improves the backhaul load | The proposed scheme is content-aware and  cost-effective |
| [114] | Optimization and  continuous update  of caching policies in decentralized  manner in small cell networks | | Distributed | Joint utility and  strategy estimation | Minimizes the service delay | The proposed scheme shows 15% and 40% gain compared to other baselines |
| [115] | Dynamic  improvement of  cache efficiency in cellular networks | Distributed | | Q-learning | Minimizes the transmission cost | The algorithm is capable of D2D offloading and cache replacement |
| [116] | Intelligent joint  content placement  and content delivery using D2D  communication | | Distributed or hybrid | Deep reinforcement learning and  recurrent neural  network | Improves the cache hit ratio | The proposed strategies and approaches can efficiently reduce the energy  consumption and delivery delay |
| [117] | Intelligent and  realistic  cache-enabled and interference  alignment for next generation wireless networks | Centralized | | Deep reinforcement learning | Improving energy  consumption and the network’s sum rate | Saving the limited backhaul capacity using cache-enabled opportunistic interference  alignment |
| [118] | Content catching with a single BS scenarios | Centralized | | Deep reinforcement learning | Maximizes both long-term and short-term cache hit rate | Reducing the computational complexity, but it only  incorporates one BS |
| [119] | Intelligent joint  optimization of  networking,  computing and  caching resources in the next generation vehicular networks | Centralized | | Deep reinforcement learning | Improves the traffic control and network efficiency | Joint optimization of resources and cache content |
| [120] | Knowledge transfer from D2D  interactions to  improve content  popularity matrix in the target domain | Centralized | | Transfer learning | Increasing the users’ QoE and improving backhaul capacity | Optimal cache strategy at  small cells for estimating  content popularity, traffic load and backhaul capacity |
| [121] | Sharing prior  knowledge from  D2D communication | Centralized | | Transfer learning | Maximizes cache-hit ratio | Improves the prediction task and caching performance |
| [122] | Efficient caching mechanism for  heterogeneous  networks | Centralized | | Transfer learning | Reduces the training time | The proposed scheme can efficiently estimate the  popularity profile |

resources in vehicular networks. The DRL agent assigns vehicles to BSs and decides whether to cache the requested content by the vehicle at the BS. The proposed algorithm jointly optimizes the problems associated with resource allocation and caching. Simulation results with various system parameters show that the DRL system performs much better than existing methods, such as edge caching, mobile edge computing (MEC) offloading, and virtualization. Based on the requested content in [118] the DRL agent acts as a controller to decide whether to store the

21

*S. Ashtari et al.*  *Array 14 (2022) 100136*

and content viewing history, is observed to execute a TL at each SBS in the network. Their method improves backhaul offloading based on the popularity of the content. This information comes from source domain sharing and accessing D2D data between users in a social community. The information is later used as prior knowledge for a content popularity matrix for estimation in the target domain. Moreover, they utilize the traffic load and backhaul capacity as feedback to the system for further improvements. From the simulation results, it is clear that the proposed TL cache procedure enables wireless systems to have prior knowledge instead of starting from scratch and also deals with data sparsity. Generally, TL acts as a knowledge plane and provides essential information to networks with similar behavior. Using caching optimization using TL is a significant step toward KDN, and one of the significant breakthroughs in this method is the maximization of QoE. Similar work was conducted in [121] to extract knowledge from the interaction between users by accessing, sharing, and recommending files. Instead of learning from scratch, the information obtained via D2D communication among users is used for caching content at the network edge. The proposed TL approach increases the cache-hit ratio and outperforms classical collaborative filtering (CF) methods [130]. Both studies provide centralized knowledge for the network to reduce backhaul traffic and improve content management.

In heterogeneous networks, SBSs are assumed to have high storage capabilities to cache popular files, such that the user can capture the file in a faster and more efficient manner. In particular, SBSs must distinguish the popularity of files and esti-mate the user demand for that file within a specific time interval. The authors of [122] proposed a TL-based approach to increase the performance of estimating a popularity profile. The central-ized approach uses prior knowledge to compute and estimate the popularity of a file based on requests during a predefined observation period. The estimation was then used to optimize the catching probability. The proposed approach reduced the convergence time of the training phase. In this study, centralized knowledge is generated based on the popularity profile of the cache content.

*6.1.6. Backhaul management*   
 Content caching at SBSs requires backhaul management, and be-cause of the heterogeneity of backhaul, both wired and wireless back-haul must work together to handle the massive traffic. Wired links use fiber cables, and wireless connections are now deploying mmWave frequencies. Owing to the heterogeneity of backhaul links, the manage-ment of backhaul has attracted significant attention. Different solutions have been proposed to reduce the complexity of backhaul [131,132]. However, new studies have concentrated on ML for reliable backhaul management. In the following section, ML studies that can be deployed in the KDN are investigated. Table 8 summarizes the studies surveyed in this section.

Knowledge derived from reinforcement learning: The deployment of low-power and short-range heterogeneous SBSs and the exponential increase in wireless traffic will cause congestion in backhaul for BSs to communicate with each other. The authors of [133] designed a distributed backhaul management model from a game-theoretic per-spective. They utilized the RL technique to solve the gaming problem using joint utility and strategy estimation. Each SBS is responsible for predicting files to download without compromising the required transmission rate. In this study, different SCNs with several coexisting backhaul systems are connected using wired links, mmWaves, and sub-6 GHz frequency bands that can only support a few files per time slot. This work delivers a balance between downloading files and pro-vides an equal chance for SBSs to access backhaul. The self-organizing nature of the RL algorithm allows SBSs to reach a Boltzmann–Gibbs

22

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| *S. Ashtari et al.* | | |  |  | *Array 14 (2022) 100136* |
| **Table 8**  Knowledge-based strategies for backhual management. | | |
| Article | Knowledge objective | Architecture | ML algorithm | Deliverable | Conclusion |
| [133] | Enabling SBSs to  autonomously  manage and  optimize the traffic in 5G networks | Centralized | Joint utility and  strategy estimation based learning | Minimizing the traffic  congestion in the backhaul | The proposed solution  converges to proper mixed  Nash equilibrium (PMNE) (an approximate version of NE) |
| [134] | Running  self-optimization  mechanism on small cells for  backhaul-aware cell range extension | Centralized | Q-learning | Backhaul management and optimized cell range  extension | The proposed approach  improves QoE and mitigates the backhaul traffic conflict in macrocell |
| [135] | Intelligent load  balancing using  backhaul resources | Centralized | Q-learning | Minimizing the probability of the user’s outage and increasing the backhaul  resource utilization | The proposed scheme improves the average throughput |

which consist of heuristic algorithms. After the training pro-cess, the algorithm can identify the path by satisfying system QoS. Once the controller receives new routing requests, the trained ML instantly provides heuristic-like results. The system performance is compared to the classical max–min ant system (MMAS) [154] which has been proven to be an appropriate approach for examining the performance of routing frameworks. MMAS is an upgraded version of the ant colony optimization method taken from ant routing [155]. The proposed study [137] can be adapted to the centralized architecture of the KDN for this particular study. The reason is that the routing framework uti-lizes the SDN functionality to gather global information and then predict a route. Therefore, the proposed protocol is inherently centralized and suitable for the centralized architecture of O-RAN networks. The proposed technique enables the knowledge plan in the KDN to identify routes that maximize the network QoS.

An autonomous vehicle or self-driving car can communicate with other vehicles, roadside units, and infrastructure. This ca-pability is known as vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) communication to exchange essential in-formation, such as speed, location, environmental conditions, etc., to nearby vehicles and the controller. Authors of [138] proposed a delay-bounded routing framework for vehicular ad hoc networks (VANETs). They focused on delivering messages with user-defined delay parameters and minimum usage of the radio spectrum. The delay-bounded routing protocol uses linear regression to predict the traveling distance and available time for forwarding a message. Their algorithm has two schemes, the greedy and centralized schemes, which are both based on linear regression. The greedy strategy predicts the available time by using current sampling data, and the centralized scheme uses global statistical information to choose the optimal path for routing. The simulation results illustrate that the radio usage is greatly reduced. Moreover, the functionality of using both greedy and centralized-based techniques establishes a connec-tion for the hybrid architecture of KDN to enable the installation of routing protocols in VANETs.

In [139], the authors combined two SL classifiers; decision tree learner and rule learner, for routing optimization in a wireless sensor network. They proposed a MetricMap based on MintRoute, which collects the routing protocol to obtain the link quality. MetricMap uses two components, the first component updates the features for the learning strategy when a packet arrives. The second component controls the link classification with input from the features, and the output values indicate the link quality. For their performance measurements, they considered data latency, data delivery rate, and fairness index. From the evaluation of the 30 sensor nodes in the network, the

23

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| *S. Ashtari et al.* | |  |  |  | *Array 14 (2022) 100136* |
| **Table 9**  Knowledge-based strategies for routing. | |
| Article | Knowledge objective | Architecture | ML algorithm | Deliverable | Conclusion |
| [137] | Real-time route discovery by the trained dataset | Centralized | Deep neural  network | Improves network  performance by efficient traffic engineering | Simple framework and  computationally efficient, but it lacks some network features including node mobility,  backhaul traffic, etc. |
| [138] | Routing strategy can be used for similar VANETs | Hybrid | Linear regression | Efficient spectrum usage and high packet delivery ratio | Introducing delay-bound  routing protocol with accurate route prediction with high  efficiency |
| [139] | Optimized route  selection in wireless sensor networks | Centralized | Decision tree and rule learner | Significant improvement in data delivery ratio | Efficient sensor  communication via user link quality |
| [140] | Enabling efficient route selection | Centralized | Decision tree and neural networks | Decreasing average latency, improving overhead and  packet delivery ratio | Highly efficient route selection based on various network  parameters |
| [141] | Real-time updating and routing  judgments in  heterogeneous  networks | Centralized | Supervised learning and deep CNN | Minimize the average  delay and improves packet loss ratio | Intelligent traffic control |
| [142] | Self-driving  networks by  learning traffic  control mechanism in heterogeneous  networks | Distributed | Supervised deep  neural networks | Better signaling overhead, delay, and throughput  compared to OSPF | Capable of learning complex patterns and functions to  predict the least cost path |
| [143] | Efficient and  intelligent route  decisions in wireless mobile networks  based on KDN  architecture | Centralized | Principle component analysis and neural network | Lower packet loss ratio  and acceptable throughput and E2E delay | Provides a  dimension-reduction vector  matrix to reduce the algorithm response time but it must be verified over larger networks |
| [144] | Trained algorithm  can select next hop with a least total  average hop count and successful  delivery probability in wireless networks | Centralized | K-mean clustering | Less dropped packets and network overhead | Simple, efficient routing  protocol but needs  improvement in average  message latency and should involve energy consumption as a node feature |
| [145] | Optimal route  selection and  prediction capability using global  information for  VANETs | Centralized | Unsupervised  learning-based  algorithm | Better transmission delay compared to existing  VANET routing protocols | The proposed scheme is robust to varying mobility rates |
| [146] | Smart routing  decision in  IoT-based smart cities | Centralized | Deep reinforcement learning | Mitigating the network congestion and load  balancing | Simultaneous QoE satisfaction and crowd management |
| [147] | Finalized updated route table can be used as the routing policy in wireless sensor networks | Distributed | Q-learning | Extending the network lifetime | Acceptable communication  and computational overhead, it can be extended with more node characteristics, such as mobility and traffic |
| [148] | Identifying stable routes in cognitive radio networks | Distributed | Q-learning | Minimizing the  interference between SUs and PUs and less frequent route discovery | Boosting network scalability and functionality |
| [149] | Route selection for multi-hop cognitive radio networks | Hybrid | Reinforcement  learning | Improves the QoS | Selecting the best possible route in terms of throughput and packet delivery ratio |
| [150] | Efficient route  selection in dense cognitive radio ad hoc networks | Distributed | Reinforcement  learning | Minimizing the  interference | Stable protocol when the network size increases but might be degraded in high mobility scenarios |
| [151] | Intelligent  QoS-aware route selection | Distributed | Q-learning | Energy efficient,  QoS-aware and mobility tolerance | Reliable, stable and extended lifetime network |

(*continued on next page*)

24

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| *S. Ashtari et al.* | |  |  |  | *Array 14 (2022) 100136* |
| **Table 9** (*continued*). | |
| Article | Knowledge objective | Architecture | ML algorithm | Deliverable | Conclusion |
| [152] | Intelligent traffic  routing control with KDN approach for  next generation of wireless networks | Centralized and hybrid | Deep reinforcement learning | Minimizing the overhead | The proposed paradigm  combines distributed and  centralized intelligence to achieve highest performance |
| [153] | Automatic  adaptation based on the traffic  conditions via KDN paradigm | Centralized,  distributed,  or hybrid | Deep reinforcement leaning | Traffic engineering | Near optimal solution after one single training step |

proposed an intelligent routing scheme using a deep CNN. Their method learns from the previous experience based on conges-tion, and uses this information to train a two-phase procedure, namely cold start period and intelligent running period. The cold start period is the initialization of the training set, where the algorithm only defines a route with a minimum hop path. After this period, the algorithm switches to the intelligent running period, where it performs real-time updating and routing judg-ments. More importantly, a CNN is constructed for each routing decision, which takes the collected information based on traffic patterns from routers, including traffic generation rate, to pre-dict whether the selected routing strategy can cause congestion in the network. This process is periodically updated until it is predicted that the chosen route will cause no congestion. Simu-lation results prove that the proposed algorithm performs much better in terms of E2E delay and packet loss ratio compared to conventional routing strategies, where there is no intelligence.

The proposed method is a real-time intelligent network traffic control method that can be adopted in a centralized KDN. A centralized KDN can collect traffic congestion information from the network and tune the deep CNN to converge to optimum per-formance within the timeframe threshold. It is not recommended to use the distributed knowledge plane due to the algorithm’s real-time updates, which require the maximum amount of data from a large number of nodes to make appropriate routing decisions.

In [142], a supervised DNN was proposed for routing optimiza-tion in heterogeneous networks to predict the path from the source to the destination node. Each router in the network uses a DNN to predict the next hop; the DNN takes traffic patterns as in-puts, and based on these inputs, it generates the desired output.

The output of the deep learning structure significantly improved the network traffic management. There are three phases to ob-tain a fully functional ML. The first phase is the initial phase, where the traditional routing protocols, such as OSPF, provide the network route, and the network starts to operate. At the same time, the second phase is the training phase to train the deep learning system from the collected information based on the traditional operating system. Finally, the running phase is the stage in which the machine is fully trained and can provide real-time routing strategies. This method has been proven to have a higher throughput and lower overhead compared to OSPF. The proposed study suggests a greedy-based distributed architecture over a knowledge-based network to increase the throughput.

(ii) *Knowledge derived from unsupervised learning*: In [143], the au-thors focused on load balance routing based on PCA and NN for dimension reduction and prediction of the network load status.

To obtain intelligence from the network, they combined SDN with ML and data analytics. The use of these algorithms has led to efficient and intelligent routing decisions. This article aims to address the shortcomings associated with the next generation of wireless mobile networks, such as video streaming and online gaming, to mitigate the delay caused by traffic. They proposed a routing strategy based on an ML scheme, where PCA was

25

*S. Ashtari et al.*  *Array 14 (2022) 100136*

BS. Hence, in the near future, centralized routing protocols for mobile devices can use the knowledge plane to decide whether the BS or other devices must route the packet. Centralized VANET protocols have better performance compared to the de-centralized topologies. In centralized VANET protocols, vehicles are supported by the BS and road-side units (RSUs). On the other hand in purely distributed systems, vehicles must relay any vital information across the network using multi-hop communication.

Therefore, there is a risk of link failure in highly mobile environ-ments, which might have catastrophic consequences. Generally, VANET protocols are supported by a centralized controller to collect data and inform vehicles. The proposed protocol in [145] uses a centralized controller to collect data and process the data through ML algorithms. Consequently, centralized KDN architecture is usually more suitable for VANET protocols.

(iii) *Knowledge derived from reinforcement learning:* With the massive growth of IoT devices connected to the edge network, the de-sign of routing strategies is complicated. In particular, in smart cities, routing is significantly more difficult owing to the dis-tribution of the crowd and network congestion. The authors of [146] designed a DRL algorithm for smart routing decisions for load balancing and mitigating network congestion when massive crowds are moving around the city for daily activities.

They adapted a DRL agent to directly use the NN and generate Q-values. First, the network state information is collected by the SDN controller on top of the network. Then, the DRL agent makes an action (routing decisions) based on the current state, and finally, the agent receives a reward. The objective of the reward function is to maximize the successful service access rate, minimize the data transmission delay, and balance the network load. The algorithm performance was better than that of the OSPF and enhanced-OSPF (EOSPF). Fig. 1 in this article repre-sents the network architecture of the proposed algorithm, which shows a solid connection to the KDN centralized architecture, where all the information is collected at the controller and is processed by ML in KP. The knowledge/reward is later used as an action for the following routing strategy.

In [147], the study tackles the energy-aware routing in wireless sensor networks (WSNs) to transmit data packets using efficient paths within the shortest time such that the lifetime of the network increases. Specifically, they used Q-routing algorithms and extended them to propagate information faster with lower energy consumption [158]. The algorithm uses Q-learning to save the energy levels of the nodes in matrices after a sensor sends a feedback message. When the sensor receives the feed-back messages from neighboring nodes, it modifies the Q-values in the achieved routing table. Once a node has a packet to transmit, it selects the next node from the routing table with the best Q-value in a greedy manner to relay the packet to the des-tination. Their technique has proven optimal routing decisions for low-energy nodes. The greedy-based approach allows nodes to individually select the best route, which suggests a distributed KDN.

Recently, CRN has attracted considerable attention owing to its importance in future wireless communication systems. This technology overcomes the scarcity of the channel spectrum by allowing secondary users or unlicensed users to benefit from underutilized licensed channels. However, the dynamic nature of CRNs makes routing a complicated task. The authors of [148] proposed a clustering mechanism or cluster-based routing to boost network scalability and functionality. Once the cluster heads are identified in the network, each cluster head estimates the Q-value of each neighboring node. The routing table is constructed based on the Q-values, and the largest Q-value is the next chosen node for the next hop. During the learning procedure, the state of the network represents the destination

26

*S. Ashtari et al.*  *Array 14 (2022) 100136*

network overhead. The network mind acts as a global controller for connection-oriented tunneling-based routing protocols, in-cluding segment routing and multiprotocol labeled switching. A DRL-based routing strategy is deployed in the network mind for tunneling-based routing to ensure QoS. The state of the RL algorithm is represented by the network traffic characteristic information and device information, and the action is the for-warding path. The reward it acquires is the effectiveness of an action with respect to the delay requirements and optimiza-tion targets. Their proposed RL method converges to the global minimum, and the routing strategy shows better performance compared to other routing strategies in congested areas. The pro-posed scheme has a centralized architecture with an intelligent control plane. In our proposed framework, this plane controls and installs the rules in the network. Stampa et al. proposed a DRL algorithm for optimizing routing in a centralized knowledge plane [153]. The actor–critic learning method is used, where the state of the learning algorithm is calculated by the traffic matrix (which is defined by the bandwidth request between pairs of source and destination), and the action is the path taken to transmit data (obtained using link weights). Finally, the reward of the algorithm is based on the average network delay. Their method provides operational advantages compared to traditional optimization algorithms for routing strategies. This algorithm can be used in both distributed and centralized KDN architectures, suggesting a hybrid structure.

*6.2.2. Clustering*   
 In wireless networks, nodes/sensors/users have always been clus-tered to describe their distinctive features or differentiate based on their mobility rate, coordinates, etc. Clustering different nodes for different purposes improves the overall performance of the network. It is evident from the introduction of ML techniques that clustering problems are naturally solved using K-mean algorithms. However, other cluster-ing methods have been proposed within supervised and unsupervised learning techniques. Clustering is one of the primary and essential applications of KDN for various purposes, such as traffic classification and data storage. Table 10 summarizes the studies surveyed in this section.

(i) *Knowledge derived from supervised learning:* One of the problems in ML is class imbalance, where the class distributions are highly separated. This means that the total number of minority or scares classes (also known as positive ones) is far less than the majority class (represented as negative) for a two-class scenario. When we apply a traditional classifier in these scenarios, they are likely to predict everything as a majority or negative class. In [161], the authors used logistic regression for imbalanced problems to improve the performance of the learning procedure. The proposed method is called logistic regression for imbalanced learning based on clustering (LRILC). First, K-mean clustering was applied to the dataset to partition the majority class into small clusters. Logistic regression is then used to overcome the class-imbalance problem. The experimental results show a higher accuracy in clustering the dataset compared to state-of-the-art classification methods. The proposed method can be used in a centralized KDN to solve imbalanced problems with large datasets.

In [162], fraud calls are identified by investigating the user’s behavior. Their method uses the application of SVM alongside fuzzy clustering to identify fraudulent phone subscribers. Fuzzy clustering takes unlabeled input data and clusters the data ac-cording to their similarities. Furthermore, after a trained data algorithm obtains input data, it generates a value between 0 and 1. If the output value is closer to 1, it shows a higher degree of similarity. Their algorithm takes large datasets and

27

*S. Ashtari et al.*  *Array 14 (2022) 100136*

coordinate their transmissions while reducing energy consump-tion and traffic load. This study formulated the problem as a noncooperative game between clusters, where clusters seek to minimize the cost function to reduce energy consumption. Based on the information regarding the location of SBSs and the capability of handling users and data traffic, the cluster determines their transmission power and on/off situation. The simulation results show improved overall performance when using the cluster-based coordination method in small-cell net-works. The algorithm attempts to reduce the overhead on a centralized controller by allowing the SBS to decide based on their locally acquired information. Hence, this method can be fitted to the distributed architecture of the KDN.

As the demand for ubiquitous access to wireless data increases, Tabrizi et al. [167] proposed a clustering and spectrum assign-ment and resource allocation (CaSRA) to cluster nodes in hotspot densely populated areas. The K-mean clustering algorithm was adapted to maximize spectrum utilization and increase network performance. In their algorithm, a mobile device can act as a hotspot or slave. Once the mobile user is connected to the cellu-lar network, it serves as a hotspot and provides broadband access to nearby devices known as slaves. The problem formulation is first to identify mobile devices that act as hotspots and then to obtain the users associated with each hotspot. To solve this problem, a modified version of K-mean clustering is used to cluster users based on their location and organize each cluster based on their maximum and minimum number of users. Then, the user with the minimum distance to both the BS and the center of the cluster is selected as the hotspot in that cluster. Further, using a graph coloring approach, power and spectrum are allocated to each cluster and from the cluster to hotspot and slaves. CaSRA increases the total number of supported users in a network with lower complexity. The centralized CaSRA scheme uses the BS to cluster the nodes into groups based on their distance, making the algorithm suitable for a centralized KDN. In [168], they proposed a new multicarrier waveform clas-sification for 5G communication systems. This study utilizes PCA and CNN on signal amplitude to mitigate noisy channels and obtain a high classification accuracy. Their algorithm can detect and cluster three multicarrier waveforms: universal fil-tered multicarrier (UFMC), filterbank-based multicarrier-offset quadrature amplitude modulation (FBMC-OQAM), and orthogo-nal frequency-division multiplexing-quadrature amplitude mod-ulation (OFDM-QAM). Their method works even in a dense chan-nel environment, where the transmission and detection of these three signals were not possible before. Moreover, compared to other methods, such as CNN using in-phase and quadrature (I/Q) modulation, it has better performance and less complexity. This technique can be used in a distributed manner to classify signals at 5G BSs. Hence, the scheme provides a distributed knowledge of signals.

*6.2.3. User association*   
 Current wireless communication networks rely on the existence of a cellular architecture. Cellular communication requires BSs for users to request, receive, and upload information. Every BS in the network has a coverage area that supports a specific geographic area with limited users. To increase the capacity of BSs in the cellular network, small cells were introduced to enable service providers to offload users from an overloaded BS to sub-BSs, namely macro BSs, pico BSs, and femto BSs. Therefore, more cellular networks are shifting toward heterogeneous networks (HetNets), enabling flexibility and low-cost deployment of new infrastructure. To associate users with an appropriate cell, the KDN requires reliable methods. Hence, some practical ML-based techniques have been investigated, including studies on SL and RL. Table 11 summarizes the studies surveyed in this section.

28

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| *S. Ashtari et al.* | |  |  |  | *Array 14 (2022) 100136* |
| **Table 10**  Knowledge-based strategies for clustering. | |
| Article | Knowledge objective | Architecture | ML algorithm | Deliverable | Conclusion |
| [161] | Intelligent clustering based on trained  supervised learning dataset | Centralized | Logistic regression | Traffic classification and balancing load | Extensive comparison among various classification methods |
| [162] | Automatic detection and drop of fraud  activities | Centralized | Support vector  machine and fuzzy clustering | Detecting the fraud  communication activities | The proposed algorithm  performs better compared to normal SVM |
| [163] | Efficient low power clustering solution for IoT devices | Distributed | K-nearest neighbor | Extending the battery life | The proposed scheme can reduce the transmission payload and lowering the transmission power |
| [164] | Intelligent clustering mechanism for  ultra-dense  heterogeneous  networks | Centralized | Supervised deep  neural network | Minimizes network sum energy consumption and reduces the transmission delay | The proposed algorithm is time efficient and has a 90% accuracy of user clustering |
| [169] | Smart clustering for random access  networks | Centralized | K-mean clustering | Low convergence time | Machine type communication device association with  conflict-free resource  allocation |
| [166] | Dynamic switching on/off SBSs | Distributed | K-mean clustering | Increases the energy efficiency of SBSs | Clustering wireless small cell networks based on location and traffic load, but they did not consider the frequent  mode transition |
| [167] | Intelligent node  clustering in densely populated wireless  areas | Centralized | K-mean clustering | Efficient offload traffic and reuse spectrum | The proposed algorithms  jointly cluster the nodes and assigns spectrum and physical resources |
| [168] | Automatic  multicarrier  waveforms  classification in  future cellular  networks | Distributed | Principle component analysis and CNN | Achieve low complexity and high accuracy | Classifying three waveforms including OFDM-QAM,  FBMC-OQAM, and UFMC, but should be further tested in  real channel environments  with the existence of other noises |

a bias value, and the reward of the system is determined by the BS when calculating the number of outages. The proposed algorithm increases the throughput and decreases the number of disconnected UEs. This algorithm is a greedy-based cell ex-pansion by every UE, which can be used in a distributed KDN architecture.

In vehicular networks, continuous connectivity is key to net-work intelligence and automation. However, user association also plays an essential role in vehicular networks by providing load balancing between SBSs and MBSs. The authors of [173] proposed an online RL approach (ORLA) for network load bal-ancing in VANETs. Their learning algorithm is divided into two main phases: initial RL procedure and history-based RL procedure. In the initial phase, the user and BS association problem is formulated as a multi-armed bandit problem. The vehicle association decision for each BS is the agent’s action, and the reward is the network load balance. In vehicular net-works, there are regularities in the spatial–temporal dimension due to urban traffic flows. In the history-based RL phase, by considering the dynamic changes of the environment and the spatial–temporal regularities, the association patterns obtained in the initial phase enable simultaneous load balancing of BSs. The history-based RL generates an association matrix for each BS based on the similarities between historical patterns and the current environment. The proposed algorithm is compared with the distributed dual decomposition optimization and max-SINR scheme, wherein ORLA outperformed load balancing in multiple cells. The algorithm provides a centralized architecture for future KDN-based networks.

29

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| *S. Ashtari et al.* | |  |  |  | *Array 14 (2022) 100136* |
| **Table 11**  Knowledge-based strategies for user association. | |
| Article | Knowledge objective | Architecture | ML algorithm | Deliverable | Conclusion |
| [170] | Trained sample is  used for user  association of the  mmWave systems | Centralized | Supervised learning | Accurate user association | The proposed approach shows good performance even with few training samples and  without CSI, but it consumes significant time in training  process |
| [171] | Intelligent BSs  selection by UAV  users in 5G  networks | Distributed | Supervised learning with neural network | Maximizes channel quality of the UAV-BS link | The proposed policy allows  UAVs to select the suitable BS based on BS transmit power, distance and the location of the BS, can be extended with varying mobility rates |
| [172] | Efficient user  association in  heterogeneous  networks after UE learning procedure | Distributed | Q-learning | Improving the system  throughput and reducing the number of UEs’ outages | Optimizes user cell association (PBSs or MBSs) through  learning the cell range  expansion |
| [173] | Online training  approach and  near-optimal  association in  dynamic vehicular environment | Centralized | Reinforcement  learning | Efficient load balancing in the network | Adapting the user association policy via learning the  spatial–temporal dimension regularities |
| [174] | Autonomous cell  association in  ultra-dense small  cell networks by  enabling  user-devices to store learned values | Distributed | Fuzzy Q-learning | Improving the convergence time | The proposed scheme is  memory-based user-centric  backhaul-aware but may have an impact on user’s memory usage |
| [175] | Intelligent adaptive decision making to maximize UEs’ QoS in heterogeneous  networks | Distributed | Multi-agent deep  Q-learning network | Maximizes energy  efficiency while jointly associated users | Joint optimization of user association and power  management |
| [176] | Near-optimal  solution to improve QoE in MEC-enabled live video streaming systems | Distributed | Deep reinforcement learning | Maximizing users’ QoE | Efficient user association and resource allocation |
| [177] | Smart user  association in  symbiotic radio  networks (SRNs) | Hybrid | Deep reinforcement learning | Achieving optimal user  association policy without full real-time channel  information | The proposed algorithm uses both centralized and  distributed DRL approaches to make a decision for IoT  devices, where centralized  converges with few dataset  and distributed is scalable |

appropriate BS to create communication links and determine the transmission power. The transmission power of the user is the action of the agent throughout the learning procedure. The reward function is the sum energy efficiency of all the UEs. The objective of the learning algorithm is to maximize the expected accumulated reward under QoS constraints. The convergence of multi-agent DQN was analyzed in the simulation results, and it proved to be superior to traditional RL-based techniques. Moreover, the algorithm maximizes the long-term overall network performance and demonstrates efficient energy consumption. The distributed method above shows a solid con-nection to the distributed KDN architecture. Chou et al. used DRL to jointly solve user association and resource management problems in mobile edge computing (MEC) to improve the QoE for online video streaming in 5G networks [176]. The problem is formulated based on the Markov decision process (MDP) and analyzed by a deep deterministic policy gradient (DDPG) algo-rithm based on the supply demand interpretation of the Lagrange dual problems. First, they used the traditional optimization La-grangian approach, where the source of the performance loss in this algorithm was identified as the Lagrangian multiplier update

30

*S. Ashtari et al.*  *Array 14 (2022) 100136*

To overcome this problem, two DRL algorithms were utilized to guarantee optimal user association. One of the algorithms is centralized, which makes decisions for IoT devices based on globally available information. At the same time, the other is distributed and makes decisions based on locally available information. The DRL algorithm can have two states based on the proposed DRL algorithm. The first state is the action space in a centralized DRL-based user association, which is a matrix of cellular users with the associated IoT device. The second state is a distributed DRL-based user association scheme with one IoT device. The immediate reward in both schemes is the sum rate of all IoT devices. The proposed scheme shows optimal user association with high scalability, even when IoT devices increase in the network. This algorithm is suitable for hybrid KDN architectures because it takes advantage of both centralized and distributed algorithms.

*6.2.4. Traffic and data aggregation*   
 As networks are increasing in complexity and become more difficult to manage, embedding intelligence into devices will ease optimization, recommendation, organization, and management. Most studies in the networking area are distributed in nature, which makes it difficult to include ML-based algorithms for controlling the devices. KDN function-ality provides an opportunity to bring intelligence and knowledge to the network. The KDN can collect global information to improve network performance. Traffic classification is a crucial activity in network man-agement, and massive growth in Internet users has brought network traffic classification into attention. Table 12 summarizes the studies surveyed in this section.

(i) *Knowledge derived from supervised learning*: Raikar et al. inte-grated SDN architecture with SL techniques, specifically SVM, Naïve Bayas (NB), and the nearest centroid is used to classify the network data traffic [181]. First, in the learning phase of their algorithm, the training data are fed to the system to map the network traffic into defined classes. Later, real-time data were captured and mapped based on the trained SL for network traffic classification. In their method, the SDN controller utilizes three different SL algorithms to classify the data into HTTP, mail, and streaming. Their proposed solution was able to obtain high accuracy in all three learning algorithms with the highest accuracy for NB, followed by SVM and the nearest centroid.

Their algorithm provides centralized data classification, which offers an opportunity to add intelligence to network devices.

(ii) *Knowledge derived from transfer learning*: To avoid training data from scratch, researchers in [182] method to address multi-class traffic classification problems, and they utilized Maxent as the base classifier in their approach. A new classification task in TrAdaBoost was used to extract labeled data from several network traffic sources. Next, the Maxent model is used to classify and convey traffic knowledge from the source domain to the target domain. The proposed scheme was trained and transferred as prior knowledge for different environments to reveal its performance. They tested their method with two tra-ditional ML algorithms based on Maxnet, known as NoTL and NoTL advance, where TrAdaBoost achieves better performance compared to the other two methods. Their learning algorithm can achieve high accuracy in classifying network data traffic and provide a promising solution for centralized KDN architecture.

*6.3. Mobility management*

Another critical aspect of service delivery in wireless networks is mobility management. Mobility management helps to develop user mo-bility and handover predictions. In future wireless networks, mobility management is crucial for network automation. Knowing the location

31

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| *S. Ashtari et al.* | | | |  |  | *Array 14 (2022) 100136* |
| **Table 12**  Knowledge-based strategies for traffic and data aggregation. | | | |
| Article | Knowledge objective | | Architecture | ML algorithm | Deliverable | Conclusion |
| [181] | Fully automated  management in  SDN-based networks | Centralized | | Supervised learning methods | More than 90% accuracy in data traffic classification | Provides security monitoring, fault detection and traffic  engineering |
| [182] | Real knowledge  transformation from one domain to the target domain for  efficient traffic  classification | Centralized | | Transfer learning | Achieving high  classification accuracy | In contrast to other studies training data and test data have different distributions |
| **Table 13**  Knowledge-based strategies for mobility prediction. | | | |  |  |  |
| Article | Knowledge objective | | Architecture | ML algorithm | Deliverable | Conclusion |
| [183] | Collected user  activity can be used to predict both  location and the  individual’s activity simultaneously | Centralized | | Supervised learning | Robust performance and smooth prediction upgrade | Predicts the users’ next  location by modeling users’activity patterns from GPS real-life trajectory data |
| [184] | Self-organizing  networks for  handover  management in LTE femtocell BS  networks | Centralized | | Unsupervised  learning | 70% reduction in  unnecessary handovers | Handover management based on the user indoor location |
| [185] | Allowing BSs in LTE heterogeneous  networks to  automatically and  autonomously  discover the RF  conditions in their  cell edge | | Centralized | Unsupervised  shapelets | Clustering result with an average accuracy of 95% | Classifies the users’ trajectories and optimizes handover parameters |
| [186] | Smart mobility  prediction for  efficient service  migration in the  mobile service  provision problems | Centralized | | Deep reinforcement learning | Offloading traffic and reducing latency in the system | Optimizes the service  provision problems in mobile edge computing |

and execution phases to perform a centralized knowledge-based concept. Moreover, to allow BSs to autonomously discover the RF conditions at the cell edge and their impact on the handover parameters, unsupervised-shapelets and data mining techniques were proposed to recognize patterns in the RSRP measurement reports from users [185]. Their method makes position estima-tion once a handover is triggered. Based on the positioning, the BS discovers new patterns while the network characteristics change and calculates the number of clusters in the network. The simulation results illustrate that even without prior knowledge, the algorithm provides 95% accuracy in clustering the nodes and predicting the user’s location. The proposed algorithm uses cluster heads to identify user’s movements and trajectories. This method is suitable for a centralized KDN for user-trajectory prediction.

(iii) *Knowledge derived from reinforcement learning*: In [186], a mobil-ity prediction model based on DRL at the edge of the network was proposed for mobile users. They designed a DRL framework to offload traffic by training a DQN for mobility prediction.

Their method comprises a glimpse mobility prediction model that gathers users’ mobility patterns and trains them in the DRL. The algorithm first assumes that the controller can select the best edge server and apply the DRL. Then, the controller predicts the users’ future locations based on historical data and past user mobility using the DRL agent. The authors used the actual human trajectory and user latency to obtain the perfor-mance of their algorithm. The experimental results show that

32

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| *S. Ashtari et al.* | | | |  |  | *Array 14 (2022) 100136* |
| **Table 14**  Knowledge-based strategies for handover management. | | | |
| Article | Knowledge objective | | Architecture | ML algorithm | Deliverable | Conclusion |
| [192] | Smart handover management  solution for LTE networks | Centralized | | Supervised learning with neural network | Improves users’ QoE | The handover algorithm  provides better QoE by  choosing the appropriate cell |
| [193] | Intelligent  preparation decision for better handover in 5G mmWave  networks | Centralized | | Supervised learning with deep neural  network | Reduces the signaling  overhead while improving the success rate | The proposed prediction-based conditional handover  demonstrates 98.8% and  above accuracy of prediction |
| [194] | Constructs a  centralized  knowledge-based server for future wireless networks | Centralized | | Supervised learning with recurrent  neural network | 90% traffic forecasting accuracy | The proposed model utilizes an artificial intelligence into the mobile network for  adaptive handover  optimization |
| [195] | Efficient handover management to  support intelligent vehicular networks | Centralized | | Supervised learning with recurrent  neural network | Accurate prediction of  handover trigger decision | The proposed algorithm initiates early handover registration to mitigate  communication disruption |
| [196] | Dynamic mobility management in  small cell 5G  networks | Centralized | | Fuzzy Q-learning | Reduces the unnecessary handovers while keeping an acceptable call dropped ratio | Optimization of handover parameters for different UE speeds to maximize network efficiency |
| [197] | Smart user-centered handover decision making for an  open-access  femtocell networks | Centralized | | Q-learning | Increases network capacity with less number of  overhead | The proposed scheme observes the historical data and  predicts the future user  association, but the model  only deploys one MBS |
| [198] | Adaptive user  selection and cell association in  open-access  femtocell networks | Centralized | | Fuzzy Q-learning | Fewer number of  handovers with  opportunistic channel capacity | Adapting the learning  algorithm from the past and future experiences |
| [199] | Optimal handover management in  large-scale wireless mobile networks  including IoT | Centralized | | Deep reinforcement learning and  supervised learning | Reduces the handover and ensures the system  throughput | SL uses the traditional  handover parameters to  initialize RL, then RL is used to obtain the optimal  controller for each UE |
| [200] | Robust tuning of  handover  parameters and load balancing in LTE  networks | | Centralized | Fuzzy Q-learning | Traffic engineering | Self-organizing mechanism for joint optimization of load  balancing and handover  management |
| [201] | Intelligent handover management for  LTE-advance  networks | Centralized | | Q-learning | Minimizes the handover failure rate and handover ping-pong effect | The proposed algorithm  selects the most efficient eNB using reference signal received power, reference signal  received quality, and other  criteria |
| [202] | Simultaneous  monitoring of signal strength and  optimal hand over control in 5G  networks | Centralized | | Reinforcement  learning | Maximizes the network throughput | A centralized RL agent uses the radio measurement reports from the UEs for handover  optimization |
| [203] | Optimal  decision-making for predicting handover in mmWave  networks | Centralized | | Q-learning | Maximizes the throughput | The proposed scheme finds the optimal policy based on the pedestrian information including the location and velocity |
| [204] | Intelligent handover learning in  mmWave networks | Centralized | | Double deep  reinforcement  learning | Improves the QoS | The proposed framework provides the optimal BS selection policy |
| [205] | Self-tuning of  handover  parameters and  power allocation in heterogeneous  networks | Centralized | | Deep reinforcement learning | Maximizes the throughput and reducing the handover frequency | The multi-agent algorithm  train decentralized policies for each UE to achieve better  performance |

33

*S. Ashtari et al.*  *Array 14 (2022) 100136*

studies have used ML techniques for self-organizing networks to improve network performance. Ali et al. presented an SL-based handover management scheme to improve the QoE for LTE users [192]. They utilized historical data to learn how the QoE of users changed when the handover decision was made. In particular, eNB gathers measurement parameters from the users, including the user’s radio link condition of the current eNB, the user’s neighboring eNBs, and the user QoE resulting from past experiences once handover was made. This information is fed to the two-level NN model, where the first level determines the QoE in terms of complete download or incomplete download, and the second level is trained to approximate the file download time. Based on the handover algorithm, the current serving eNB triggers the handover to the next eNB with uninterrupted service to the user. Their algorithm assigns data for users to download and measures the amount of information lost due to handover. The simulation results show that almost 96% of the data were downloaded even when handover occurred. BSs can use the proposed method to create a centralized KP and perform handover management with a low data loss ratio.

Owing to the vulnerability of mmWaves in 5G, which can cause sudden changes in the received signal strength, handover can often be misleading and lead to bad decision making. One of the contributions of 5G networks is the conditional handover (CHO). To enhance the performance of CHO, Lee et al. proposed a novel prediction-based CHO (PCHO) scheme based on a DNN [193]. Their method uses the former blockage information to predict the best gNB. Here, the signaling patterns from the BSs are collected and used in the algorithm. In this study, the authors focused on the preparation phase of handover because they claim that the preparation phase is the most vulnerable period in handover. The preparation phase occurs when the signal quality experienced by the UEs is low, and the interference from other BSs is severe. In fact, it is most likely that the UE experiences plenty of handover message delivery failures in the prepara-tion phase. Therefore, to achieve accurate handovers, a DNN is trained offline via a training dataset, allowing the current gNB to make a real-time decision on the preparation period to predict the next gNB. PCHO outperforms most of the CHO schemes in terms of the preparation success rate and prediction accuracy of the new BS for handover. The proposed algorithm also decreases the signaling overhead compared to the current studies on CHO-based mechanisms. The recommended scheme is suitable for centralized BS to manage handovers. In [194], an RNN was used to optimize the handover in 5G networks. In this work, mobile users increase the transmission of useful information on wireless channels and reduce service information. This is because each device is equipped with an NN that helps to process sensitive data locally. The result of the NN is always sent back to the knowledge-based server for further processing or storage. Their technique reduces the amount of service traffic transmitted by users through the communication channel and self-organizes the handovers in heterogeneous networks. The proposed method is simple, reduces system complexity, and provides sufficient knowledge for a centralized KP.

Intelligent vehicular networks (IVNs) have attracted many re-searchers because of their real-time road safety services and other essential applications for vehicles. However, the develop-ment of efficient and robust wireless communication in vehicular systems is still challenging for content delivery. This is mainly because of the high mobility rate of vehicles that disrupt con-nectivity. The authors of [195] proposed a two-tier ML-based scheme for intelligent handover management in an IVN. In the first tier, an RNN model is used to predict the receiving signal strength of APs to make a handover trigger decision. A stochastic Markov model is utilized in the second tier to select the next

34

*S. Ashtari et al.*  *Array 14 (2022) 100136*

parameters for self-organizing mechanisms in RANs. The authors of [200] proposed a fuzzy logic and RL technique to jointly im-prove the handover and load balancing for LTE users. The fuzzy logic tunes the handover parameters at the cell adjacency level. The Q-learning algorithm was used to optimize the fuzzy system to improve the network performance by forcing the scheme to select the most appropriate handover action. The RL action chooses the BS while handover occurs, and the reward function is based on jointly maximizing the handover management policy and load balancing function. The proposed algorithm proves that the network performance increases as both entities (handover and network load) run together to optimize the performance compared to schemes where one entity is running on SON. 3GPP introduced LTE-A/LTE-advanced to improve network coverage, capacity, and data rate, which provided low latency and a high data rate. The authors of [201] utilized the AHP-TOPSIS method with Q-learning algorithms to make intelligent handover man-agement. The proposed method selects the optimal eNB within an appropriate trigger time. To choose the most suitable eNB for handover, the algorithm uses RSRP, the reference signal received quality (RSRQ), uplink signal-to-interference plus noise ratio, location, moving direction of the UE, and current load on the eNB. Based on the running application on the network, such as speed-sensitive or delay-sensitive, AHP-TOPSIS is associated with UEs. According to the obtained ranks, UEs are allocated to a new eNB, where Q-learning evaluates the optimal triggering point for handover. The proposed algorithm is evaluated by numerical and simulation results, proving that their technique minimizes the handover failure rate and handover Ping-Pong effect compared to the fuzzy multiple-criteria cell selection (FM-CCS) scheme and other prior methods. Both algorithms can be deployed in a centralized KDN for future cellular networks. Handover is more challenging in 5G and beyond networks owing to the high attenuation rate of mmWave signals. Non-line-of-sight (NLOS) signals caused by random obstacles, rain, etc., create a limited coverage range in mmWave networks. Overcom-ing this challenge requires deploying several small cells, which will cause frequent switching of user connections between the BSs. The authors of [202]used a centralized RL-based algorithm for handover optimization in 5G cellular networks. Their method chooses an appropriate BS based on radio measurements ob-tained from the user. Handover management is formulated as a sub-class of RL contextual multi-arm bandit (CMAB) problems. The BS collects the RSRP reports from the UE and forwards them to a centralized CMAB where the handover decision takes place, which is the action of the learning agent. The agent’s reward is a function of the link beam of the RSRP, which is proportional to the network throughput, and the aim is to maximize the throughput. The proposed mechanism performs optimally in many simulated environments, demonstrating su-periority compared to the results obtained with state-of-the-art algorithms. Similarly, the authors of [203] used RL techniques to predict the optimal handover decision-making action based on the information acquired by pedestrian movements. They utilized the location and mobility rate of pedestrians to learn the optimal handover policy to maximize the throughput of the network. Their network consists of a single station, mmWave APs, an access controller, and a human tracking module. The human tracking module uses an RGB camera to collect the position and velocity of the users. Based on the collected data, the access controller performs a Q-learning algorithm and makes a handover decision. The action of the learning algorithm in each iteration involves choosing one of the deployed APs. The reward function performs the data rate of the communication link once a handover is made. The aim is to maximize the reward in

35

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| *S. Ashtari et al.* | | | |  |  | *Array 14 (2022) 100136* |
| **Table 15**  Knowledge-based strategies for indoor localization. | | | |
| Article | Knowledge objective | | Architecture | ML algorithm | Deliverable | Conclusion |
| [209] | Adaptive tuning to maximize indoor  localization | Distributed | | K-nearest neighbor | Achieves location error as low at 1.7 m | The proposed algorithm uses RSS-level-based feature scaling model to improve the  accuracy of indoor localization |
| [210] | Efficient near  real-time  localization system for navigation and monitoring in  mobile devices | Distributed | | Support vector  machine | Significant improvement in convergence time and  prediction accuracy | The trained offline scheme can decrease the online  learning error by 0.8 m |
| [211] | Robust and efficient wireless localization method | Hybrid | | Relevance vector  machine | Reduces the computational complexity while achieving acceptable localization  accuracy | The proposed algorithm uses a classifier to identify NLOS  signals |
| [212] | Optimal indoor localization for UWB networks | Distributed | | Support vector  machine | Significant performance improvement in various practical scenarios | The proposed algorithm  estimate the indoor  localization ranging error to improve the ranging error mitigation |
| [213] | Optimal K-nearest  neighbor positioning algorithm for WLAN | | Distributed | K-nearest neighbor | Maximizes the accuracy with optimal number of reference points | Adaptive indoor positioning system |
| [214] | Introducing a novel signal fingerprint  for geographical  localization in LTE networks | Distributed | | Neural network | Reduces the calculation time | By only using one LTE eNB, the algorithm measures the location with a maximum error margin of 6 meters |
| [215] | The trained 3D  model is used for accurate indoor  localization | Distributed | | Deep autoencoder network | High-performance 3D  localization in large indoor spaces | The proposed method extracts RSS measurements to increase the indoor positioning  accuracy |
| [216] | Online real-time  positioning for  mobile devices  based on the offline trained data | Distributed | | Kernel principal  component analysis (KPCA) | achieving 2 m positioning error while reducing the size of the radio map | The algorithm uses spatial division technique based on Random Forest technique to increase the accuracy of  localization |
| [217] | Blind indoor  localization using minimum  information about the surrounding | Distributed | | Principal component analysis | Accurate trajectory  estimation | Ability to distinguish floors in a building and trajectory  learning |
| [218] | High-performance localization system with smart  adaptation  mechanism | Distributed | | Autoencoder | Achieves high accuracy without utilizing radio path loss model | The proposed scheme  outperforms maximum  likelihood estimation,  fingerprinting methods and generalized regression NN |
| [219] | Suboptimal solution for indoor  location-based  services | Distributed | | Deep learning | Achieves high positioning accuracy | The algorithm utilizes CSI  phase information for  fingerprinting and outperforms three benchmark schemes  based on RSS or CSI |
| [220] | Optimal localization algorithm for  mobile devices | Distributed | | Deep autoencoder network | Maximizes the localization accuracy | The algorithm ensures  acceptable error ranging by using CSI and estimated AoAs |

(i) *Knowledge derived from supervised learning*: In [209], the authors presented a feature-scaling-based KNN (FS-KNN) to improve the localization accuracy. The algorithm depends on the measured RSS reported by the MS, which accounts for the actual relation-ship between the signal differences and geometrical distances. To obtain the parameters of the weight function, iterative train-ing was established to tune the parameters. After training the model, the algorithm finds the optimal values corresponding to the actual distance between a newly received RSS vector and each fingerprint. Then, the user location with an average error as low as 1.70 m is determined by solving a weight mean of locations based on K nearest reference points. This algorithm has two phases, which train the system in the offline phase and use it for online location estimation. The offline trained algorithm can

36

*S. Ashtari et al.*  *Array 14 (2022) 100136*

authors of [210] proposed an online independent support vector machine (OISVM) for indoor localization that avoids training from scratch. Their model uses the RSS of WiFi signals to make online predictions and facilitate mobile devices. The proposed model includes two phases: offline and online. The algorithms learn through pre-collected RSS with reference point (RP) labels appended to the corresponding RSS during the offline phase. The offline phase also incorporates kernel parameter selection and data sampling to deal with the imbalanced dispersion of the data samples. In the online phase, new RSS samples are collected by a centralized local AP for online learning and to estimate the location. Compared to traditional SVM methods, their method can balance the accuracy of localization and model size. From the simulation results, the location estimation error decreased by 0.8 m, while the training phase time and period were re-duced considerably compared to the traditional techniques. The proposed technique can be used via a distributed architecture of KDN.

Due to the complexity of the indoor environment and the exis-tence of various obstacles, high-accuracy localization prediction is rarely achievable. This is due to the NLOS radio blockage and insufficient information from the nodes. To this end, it is helpful to identify and categorize the NLOS signals. Nguyen et al. developed a robust relevance vector machine (RVM) for ultra-wideband (UWB) signals using time-of-arrival (TOA) [211]. They introduced a hybrid two-step iterative (TSI)-based lo-calization algorithm. In particular, they utilized cooperative localization, meaning that both centralized and distributed ap-proaches for positioning contributed to estimating the location. In centralized localization, a central processor gathers the node’s information and builds a map. This map is later sent to nodes, and the nodes perform location estimation in a distributed manner. Additionally, they designed a classifier and regressor to mitigate the limitations of SVM. The RVM classifier can classify the line-of-sight (LOS) and NLOS signals received from the unknown node positions. At the same time, an RVM regressor was adapted to predict the ranging error of the input data. First, the RVM regressor is trained using a feature vector as input data, which consists of maximum amplitude, received energy, mean excess delay, and so on. The benefits of the RVM classifier include the use of a smaller number of relevance vectors than the support vectors in the SVM and mitigating the range estimation error. This method proved to reduce the communication overhead and computational complexity while providing high localization accuracy. In contrast to this work, the authors of [212] presented a novel approach to mitigate the ranging error by directly mitigating the bias occurring in both LOS and NLOS. Therefore, explicit signal identification between LOS and NLOS is not necessary. Specifically, an SVM-based two-class non-parametric regressors method is used to learn and map the features extracted from the received signal to the ranging errors. The simulation results demonstrate that the proposed regressors greatly improve the performance in various envi-ronmental scenarios compared to conventional methods. Both studies presented above use UWB anchor nodes to estimate the location of objects, where UWB nodes collect the surrounding information and process it in a distributed manner to compute the location.

In addition to SVM and RVM, some researchers have used the KNN to achieve acceptable indoor localization. For instance, Xu et al. [213] proposed an optimal KNN positioning algorithm based on theoretical accuracy criteria (TAC) in WLAN indoor environments. In this method, the optimal number of nearest RPs that can locate the user is theoretically analyzed. The KNN-based localization algorithm demonstrated that even with k=1 and

37

*S. Ashtari et al.*  *Array 14 (2022) 100136*

mapping function. In the denoising section, all the nodes’ RSS errors are improved by using a multi-layer denoising archi-tecture. Finally, from the location section, the corresponding square grid labels were detected to estimate the location. They have a two-stage training procedure, where the first stage uses an UL algorithm for pre-training each layer and a fine-tuning stage to minimize the error of the entire network. The proposed algorithm shows higher location accuracy than the maximum likelihood estimation (MLE), generalized regression neural net-work (GRNN), and fingerprinting methods. The proposed works in [217,218] suggest a distributed KDN architecture.

Indoor fingerprinting localization systems that use RSS-based WiFi measurements are easy to implement and have low hard-ware requirements. However, RSS-based methods have two main problems. First, the RSS values are highly random and can change according to path loss, shadowing, etc. Second, the RSS formation is the average of all the obtained received signal am-plitudes. To tackle this problem, the authors of [219] proposed a novel deep-learning-based indoor fingerprinting system via CSI. The CSI amplitude responses are fed to the input of the DNN with a greedy behavior to reduce the complexity of the training phase. The weights in the DNN are stored as fingerprints to aid the lo-calization procedure in the online phase. The online process uses probabilistic data fusion based on the radial basis function (RBF) to estimate the user locations. Their algorithm was examined in varying propagation environments and showed promising results with high accuracy. The same group of authors of [220] considered the calibrated CSI data of the 5-GHz OFDM channel information for indoor fingerprinting. Similar to other methods, their algorithm consists of two phases: online and offline. In the offline phase, a deep autoencoder network is utilized for each user position to rebuild the CSI phase information, where the weights are stored as the fingerprint. In the online phase, once a new user location needs to be obtained, the algorithm uses a probabilistic method that uses a weighted average of all the reference locations in the fingerprint to estimate the position. The proposed algorithm was validated with experimental results, which indicated that the proposed method was superior to the traditional CSI and RSS-based methods. Proposed studies above establish localization in mobile devices in a distributed manner.

**7. Terms and conditions associated with machine learning algo-rithms for generating knowledge**

In this section, the terms and conditions associated with KDN prob-lems are generalized for researchers to consider before using knowledge in the network. These conditions are categorized to identify the type of problem KDN is required to solve: implementation complexity, time consumption, training data, and the differences between ML techniques in the same KDN problem. After checking each condition and meeting the requirements, a final decision can be made on whether to adapt the KDN algorithm and which ML algorithm is more cost-efficient.

*7.1. Identifying suitable ML algorithm*

The first and most crucial step is to discover the type of ML algo-rithm that is most suitable for any particular wireless communication problem. The majority of wireless communication problems are solved within a few different ML algorithms, categorized as regression prob-lems, classification problems, clustering problems, and MDP problems. In regression problems, the ML algorithm is required to predict a continuous value output given an input. In classification problems, the ML algorithm needs to predict a discrete value output, usually answered by a yes or no, and zero or one to essentially identify the class to which the input belongs. The clustering problems are ML techniques, where the data are grouped based on their type or value. Finally,

38

*S. Ashtari et al.*  *Array 14 (2022) 100136*

changes in the characteristics of the communication environment may affect the performance of the ML algorithm. However, the training time in NN-based approaches can be reduced using GPUs, and using TL helps to decrease the training time and accelerate the learning procedure, as illustrated in [227,228].

The response time of the trained learning algorithm is more im-portant than the training time. Most applications in wireless networks require a quick response, on a timescale of milliseconds, such as decision making in resource management. Let us consider two differ-ent approaches of ML applied to these wireless network applications, namely NN-based approaches and alternative approaches. The time cost can be discussed as follows:

(1) *NN-based approaches:* In some applications, we can achieve an acceptable response time without GPU usage. For example, as shown in Table 1 in [89], on average, it is possible to achieve 0.0149 ms for power control decision making in a network with 30 users. Therefore, making resource management decisions within an acceptable time frame is feasible for power allocation problems using a trained DNN [89,229,230]. However, as the network size increased, both the response time and training time exponentially increased. One solution proposed by [231] is to use GPU-based parallel computing to enable NNs to predict within a tolerable range (milliseconds). Furthermore, there is a deep Q network in the DRL, and the deep Q network depends on the output of Q-values from the NN. Hence, the response time mostly depends on the NN process time. However, one promising solution that KDN naturally provides is using the pre-trained data as knowledge, so the response time can be near-optimal and suitable for future wireless network applications.

(2) *Alternative approaches:* In resource allocation problems, in both RL-based Q-learning and joint utility and strategy estimation-based approaches, the aim is to find a policy or strategy that suites the dynamic nature of the environment. In RL-based meth-ods, after the learning algorithm converges, the policy or strat-egy from the trained algorithm becomes fixed. In Q-learning, the strategy is represented by a set of Q-values, where each set cor-responds to a system state and an associated action. Therefore, a well-trained RL algorithm can respond in milliseconds. Addition-ally, the joint utility and strategy estimation-based learning goal is to choose a probability value that indicates an action. Here, the agent only needs to generate a random number between 0–1 to select the appropriate action. Consequently, a well-trained al-gorithm can accelerate the decision-making process and achieve a millisecond response.

*7.4. Suitable dataset availability for training and generating knowledge*

Different training data are collected or generated for supervised, unsupervised, and reinforcement learning, depending on the character-istics of the problem. For instance, in spectrum allocation problems, the authors of [82,83] used non-cooperative game methods where the data were trained based on the collected information from the primary and secondary users. In [86] the same non-cooperative spectrum allocation is modeled with the difference where the data is collected using an RNN model at each BS; here, BSs continuously interact with each other to collect training data. Other authors of [84] used joint utility and strategy estimation-based learning to collect D2D information and gen-erate training data. In power allocation, when the SL method is used, such as in [89,229,230], the authors aim to adapt a neural network to approximate power allocation in a complex environment, including the genetic and WMMSE algorithms. Using these algorithms enables these studies to generate training data under different network scenarios that can later be used in KPs as knowledge. Obtaining appropriate raw data is important for researchers on the same topics. For example, in cache problems, the authors used different datasets for cache management to

39

*S. Ashtari et al.*  *Array 14 (2022) 100136*

high accuracy, although the learning procedure can be time-consuming owing to the optimization of various parameters.

However, with sufficient training datasets and powerful GPUs, DNNs are more recommended than other learning machines.

Furthermore, other neural networks, such as CNN and RNN, can reduce the training time and system overhead. Both algorithms have their own advantages for solving different problems. For instance, CNN is suitable for learning spatial features, including the channel gain matrix, while RNN is good at processing time series problems for feature extraction.

(2) *Techniques for deploy clustering:* ML models applied to clustering mainly include K-mean clustering. However, there are other supervised and NN studies deployed to clustering problems. K-mean clustering is one of the most popular and simplest methods for clustering data. Generally, the K-mean is used to differentiate between groups with similar data points and patterns. Neural networks are also used to organize the input data. For instance, given a set of images, the NN can organize and provide images with similar content. This process does not provide clusters, but it creates meaningful representations of the dataset, which can be used for clustering. The main difference between these two algorithms is the complexity of implementation. Moreover, in K-mean clustering, the number of cluster centroids (K value) is essential; however, in NN, the design of the hidden layers and other factors must be considered.

(3) *Techniques to deploy decision making:* ML-based decision-making algorithms are Q-learning, joint utility and strategy estima-tion, actor–critic, and DRL. These algorithms are applied to dynamic environments to learn patterns and decide accordingly.

Q-learning is an off-policy RL that seeks to take the best action in any particular state. An actor–critic learning strategy is helpful in non-Markov environments, where the algorithm can learn an explicit stochastic policy. Compared with the Q-learning value-based learning algorithm, policy knowledge transfer is easier in actor–critic learning because policy and value functions are up-dated independently. Joint utility and strategy estimation-based learning are more suitable for multi-agent scenarios, providing a stable system when one agent diverges from its mixed strategy.

The agent in DRL is capable of learning from a high-dimensional input state, which is the advantage of DRL over Q-learning and actor–critic learning [234–237]. Another advantage of DRL is its ability to make good action decisions even in unfamiliar situations [238]. Moreover, both Q-learning and actor–critic learning require a storage repository for each state and action procedure. Therefore, they are not suitable for multidimensional datasets. Furthermore, training a DRL has a high computational complexity.

**8. Motivations for applying knowledge-defined networking**

After summarizing the terms and conditions associated with ML algorithms, we need to investigate the motivations for applying the appropriate ML algorithm to KDN-based networks. It is essential to ex-amine the reasons and motivations for adapting KDN-based approaches to wireless networks. The subsections below provide the reasons for applying ML and knowledge to the network based on the literature surveyed throughout this study.

*8.1. Lack of network knowledge and intelligence*

Although there exist centralized algorithms and techniques for opti-mization that can achieve the objectives of various performance crite-ria, the lack of global network knowledge and intelligence has been recognized by researchers [11,239]. For instance, the baseline tech-nique to achieve load balancing and backhaul management in [120–

40

*S. Ashtari et al.*  *Array 14 (2022) 100136*

*8.4. Learning patterns and predicting the future*

By utilizing neural networks, hidden patterns in a system can be learned and used to estimate future values or predict the future, which is an advantage in KDN-based networks. In this context, the authors used this NN functionality to improve system performance. In [92], a multi-agent DRL technique was used to observe the spatial features based on the collected CSI and QoS to make a wiser power allocation decision when the network experiences dynamic changes. Moreover, in traditional indoor localization, the system’s performance can be easily affected, resulting in inaccurate positioning owing to the complexity of the environment. As a result, many researchers are now motivated to use NNs to increase the positioning accuracy by learning and updating user patterns [214,219]. Consequently, the motivation of surveyed works utilizing NN can be observed in [12,86,89,92,97,111–113,116, 117,141,164,168,193–195].

*8.5. Ability to handle complex problems and provide low-complexity solu-tions*

One of the important reasons and motivations for using ML al-gorithms is their ability to handle complex problems and datasets. The authors of [86] trained a multi-agent RL-based with ESN for efficient spectrum allocation and load balancing in LTE-U networks. Other researchers in [168] used CNN to achieve a low-complexity and high-accuracy clustering algorithm to classify three different types of waveforms in wireless communication systems. Moreover, the ability to handle high-complexity problems is the main reason why authors use RL. For example, for the on/off sleep mode control of small cells, the authors of [103] used distributed Q-learning to decide on the sleep mode of each BS, which led to a low-complexity sleep-mode control algorithm. Overall, the motivation for providing a low-complexity so-lution for the KDN paradigm can be seen in the literature [12,79,118, 152,167,194,211,216,219].

*8.6. Avoiding wrong decisions*

Some traditional and heuristic approaches for optimization and estimation based on a fixed set of rules are often unable to avoid faulty and unsatisfactory results that have occurred previously. This means that these approaches are unable to learn from their mistakes and correct their decisions. Such problems can be seen in OSPF-based routing strategies, as stated in [141], where OSPF routing will result in some congestion in certain situations at some routers even though it may know the congested router. Hence, in these situations, the OSPF repeats the same action (wrong decision), leading to congestion again. In handover strategies based on RSSI values, a similar problem can accrue, as shown in [244]. Moreover, other similar problems can be observed in the literature for user association based on max-SINR [173] and in the BS on/off sleep mode control strategy [106]. To prevent wrong decision-making in traditional algorithms, RL and deep learning are adapted to learn through historical and new data to prevent any previously incorrect decisions. For instance, deep learning in routing strategies enables the algorithm to avoid mistakes, such as congestion in the network under different traffic patterns. RL can also be utilized in other approaches to overcome the same type of problem [245,246]. In summary, decision making with poor performance outcomes can be avoided using ML strategies in the KDN framework, which can be further observed in the surveyed literature of [63,89,95,98,114,115, 118,133,134,141,150,247–251]

41

*S. Ashtari et al.*  *Array 14 (2022) 100136*

*9.3. Knowledge validation, uncertainty and compromises*

ML and intelligence have been envisioned by many researchers as the most important feature in 6G, as ML algorithms have been extensively used in complex scenarios. Therefore, it is evident that the KDN architecture can be used to address the challenges of 6G. One of the main challenges faced by all technologies is the validation of knowledge. If 6G targets an automatically configured cellular network, there must be a mechanism to verify the confidence and certainty of knowledge. As a result, a certainty mechanism must be deployed to acknowledge the level of certainty, whether knowledge is practical or compromised. The output of the ML algorithm must be checked by the expected results to evaluate the degree of uncertainty. A threshold barrier can be used to validate the usefulness of knowledge. If the knowledge is authorized to deploy in the network, the ML output has been successful, but if the compared strategy has revealed unauthorized knowledge, then the ML’s output cannot pass the threshold value. In this case, a new ML technique must compute the new knowledge and go through the same procedure, which causes a delay that affects the system performance. In the worst-case scenario, if the knowledge is re-jected again, then an extreme case must be considered. To mitigate the worst-case scenario, any proposed algorithm must undergo various ex-perimental analyses in a real testbed to find the possible ML substitutes in any given scenario. Therefore, ML algorithms must be tested in the same environment with similar characteristics to provide insights into different ML techniques. Consequently, those with similar performance will be selected to be each other’s replacement to mitigate the worst-case scenario. Moreover, related regulations and standardization must be established to guarantee the 6G service requirements.

*9.4. Lapse, or loss, or priority of knowledge*

There are two approaches to ML: reactive and proactive. In the reactive approach, the learning agent must address an issue that has already been defined. However, in a proactive manner, the learning agent must simultaneously adapt to possible future problems. Cur-rent network functions are mainly designed using reactive network protocols with human intervention for maintenance and upgrading. To enable intelligence and deploy KDN in future wireless networks, proactive learning ability is necessary [252]. In proactive learning, knowledge must be prioritized such that important metrics of the network, including channel allocation, traffic clustering, traffic pre-diction, computing offloading, radio resource scheduling, and network configuration, have a ranking of knowledge prioritization. Thus, once the network is fully congested with heterogeneous nodes, the KP must prioritize which tasks have the most pressing matter and which ones are crucial for the network functionality. Another reason for prioritization could be the lack of storage for knowledge, where it can cause lapse or, in a worst-case scenario, loss of knowledge. A lapse of knowledge is referred to as incomplete knowledge throughout prioritization, ML training, or technical issues. Moreover, loss of knowledge can accrue due to memory shortages and technical issues. In both cases, knowledge is useless and has a negative impact on the network performance. Therefore, a systematic monitoring management module must observe the KP plane. Specifically, to ensure that each network application is prioritized for knowledge extraction, the storage capacity is sufficient, and the knowledge is stable and informative.

*9.5. Effectiveness of an ML algorithm in KDN*

Although ML studies have been a point of discussion over the last few years [253], current research in the wireless communication area is still unripe. This is the main reason why ML has not been practically applied to existing wireless networks. To evolve ML-based algorithms to meet the requirements of future 6G systems, it is essen-tial to standardize AI-embedded communication. The performance of

42

*S. Ashtari et al.*  *Array 14 (2022) 100136*

the Application layer perspective, various indoor localization tech-niques are presented. Moreover, appropriate ML-based studies were thoroughly explored for each surveyed application, and the most suit-able KDN architecture was suggested. We achieved a comprehensive review of different parts of wireless networks and provided insights into how different algorithms perform, enabling future researchers to adapt the most appropriate ML-based study with the suitable architecture of KDN. Further, the conditions associated with ML-based strategies in the context of KDN were provided, followed by the motivation to apply KDN. Finally, we outlined several unsolved problems and challenges within the KDN paradigm.

**Declaration of competing interest**

The authors declare that they have no known competing finan-cial interests or personal relationships that could have appeared to influence the work reported in this paper.

**References**

[1] [Shafin R, Liu L, Chandrasekhar V, Chen H, Reed J, Zhang JC. Artificial intelligence-enabled cellular networks: A critical path to beyond-5G and 6G. IEEE Wirel Commun 2020;27(2):212–7.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb1)

[2] [Jiang](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb1)  [W,](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb1)  [Han](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb1)  [B,](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb1)  [Habibi](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb1)  [MA,](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb1)  [Scho](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb1)[tten](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb2)  [HD.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb2)  [The](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb2)  [road](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb2)  [towards](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb2)  [6G:](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb2)  [A](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb2)  [comprehensive survey. IEEE Open J Commun Soc 2021;2:334–66.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb2)

[3] [Clark DD, Partridge C, Ramming JC, Wroclawski JT. A knowledge](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb2) [plane for the internet. In: Proceedings of the 2003 conference on applications, technologies, architectures, and protocols for computer communications. ACM; 2003, p. 3–10.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb3) [4] [Mestres A, Rodriguez-Natal A, Carner J, Barlet-Ros P, Alarcón E, Solé M,](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb3) [et al. Knowledge-defined networking. ACM SIGCOMM Comput Commun Rev 2017;47(3):2–10.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb4)

[5] [Bosshart P, Daly](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb4) [D, Gibb G, Izzard M, McKeown N, Rexford J, et al. P4: Programming protocol-independent packet processors. ACM SIGCOMM Comput Commun Rev 2014;44(3):87–95.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb5)

[6] [Hu F, Hao Q, Bao K. A survey on](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb5) [software-defined network and openflow: From](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb6)  [concept to implementation. IEEE Commun Surv Tutor 2014;16(4):2181–206.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb6) [7] [Niknam S, Roy A, Dhillon HS, Singh S, Banerji R, Reed JH, et al. Intellige](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb6)nt o-ran for beyond 5g and 6g wireless networks. 2020, arXiv preprint [arXiv:](http://arxiv.org/abs/2005.08374)  [2005.08374](http://arxiv.org/abs/2005.08374).

[8] [Pan C, Elkashlan M, Wang J, Yuan J, Hanzo L. User-centric C-RAN architecture for ultra-dense 5G networks: Challenges and methodologies. IEEE Commun Mag 2018;56(6):14–20.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb8)

[9] [Kazemifard N, Sh](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb8)[ah-Mansouri V. Minimum delay function placement and resource allocation for open RAN (O-RAN) 5G networks. Comput Netw 2021;188:107809.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb9)

[10] [Balasubramanian](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb9) [B, Daniels ES, Hiltunen M, Jana R, Joshi K, Sivaraj R, et al.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb10)

[RIC: A RAN intelligent controller platform for AI-enabled cellular networks. IEEE Internet Comput 2021;25(2):7–17.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb10)

[11] [Kato](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb10)  [N,](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb10)  [Mao](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb10)  [B,](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb10)  [Tang](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb10)  [F,](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb10)  [Kawamot](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb10)[o](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb11)  [Y,](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb11)  [Liu](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb11)  [J.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb11)  [Ten](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb11)  [challenges](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb11)  [in](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb11)  [ad-](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb11) [vancing machine learning technologies toward 6G. IEEE Wirel Commun](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb11)  [2020;27(3):96–103.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb11)

[12] [Pham TAQ, Hadjad](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb11)[j-Aoul Y, Outtagarts A. Deep reinforcement learning based QoS-aware routing in knowledge-defined networking. In: International con-ference on heterogeneous networking for quality, reliability, security and robustness. Springer; 2018, p. 14–26.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb12)

[13] [Lu W, Liang L, Kong B, Li B, Zhu Z](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb12)[. AI-assisted knowledge-defined network orchestration for energy-efficient data center networks. IEEE Commun Mag 2020;58(1):86–92.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb13)

[14] [Careglio D, Spadar](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb13)[o S, Cabellos A, Lazaro J, Perelló J, Barlet P, et al. ALLIANCE project: Architecting a knowledge-defined 5G-enabled network infrastructure.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb14)

[In: 2018 20th International conference on transparent optical networks. IEEE; 2018, p. 1–6.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb14)

[15] [Hyun J, Van](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb14) [Tu N, Hong JW-K. Towards knowledge-defined networking using in-band network telemetry. In: NOMS 2018-2018 IEEE/IFIP network operations and management symposium. IEEE; 2018, p. 1–7.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb15)

[16] [Liu G, Jiang D. 5G: Vision and requirements for m](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb15)[obile communication system](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb16)  [towards year 2020. Chin J Eng 2016;2016(2016):8.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb16)

[17] [Bhalla MR, Bhalla AV. Generations of mobile wireles](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb16)[s technology: A survey. Int](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb17)  [J Comput Appl 2010;5(4):26–32.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb17)

[18] [Marshall C. By 2019, 80% of the world’s internet traffic will be video. Tubular](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb18)  [Insights 2015;3:2017, Retrieved on January.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb18)

[19] [Cass S. The age of the zettabyte cisco: the](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb18) [future of internet traffic is video](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb19)  [[Dataflow]. IEEE Spectr 2014;51(3):68.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb19)

[20] [Abrol A, Jha RK, Jain S, Kumar P. Joi](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb19)[nt power allocation and relay selection strategy for 5G network: a step towards green communication. Telecommun Syst 2018;68(2):201–15.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb20)

43

*S. Ashtari et al.*  *Array 14 (2022) 100136*

[54] [Vapnik VN. An overview of statistical learning theory. IEEE Trans Neural Netw](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb54)  [1999;10(5):988–99.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb54)

[55] [Mohammed M, Kh](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb54)[an MB, Bashier EBM. Machine learning: algorithms and](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb55)  [applications. Crc Press; 2016.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb55)

[56] [Celebi ME, Aydin K. Unsuper](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb55)[vised learning algorithms. Springer; 2016.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb56) [57] [Kim K. Face recognition using principle component analysis. In: Intern](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb56)[ational](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb57)  [conference on computer vision and pattern recognition, Vol. 586. 1996, p. 591.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb57) [58] [Madsen RE, Hansen LK, Winther O. Singular value decomposition and principal](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb57)  [component analysis. Neural Netw 2004;1:1–5.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb58)

[59] [Sutton RS, Barto AG. Reinforcement learning:](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb58) [An introduction. MIT Press; 2018.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb59) [60] [Sutton RS, Barto AG, et al. Introduction to reinforcement learning, Vol. 2, no. 4. MIT press Cambridge; 1998.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb60)

[61] [Glorennec PY. Fuzzy Q-learnin](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb60)[g and dynamical fuzzy Q-learning. In: Proceed-ings of 1994 IEEE 3rd international fuzzy systems conference. IEEE; 1994, p. 474–9.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb61)

[62] Goodfellow I, Bengio Y, Courville A. Deep learning, book in preparation for MIT press. 2016, [Online]. Available: <http://www.deeplearningbook.org/>.

[63] [Ye H, Li GY. Deep reinforcement learn](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb63)[ing based distributed resource allo](http://www.deeplearningbook.org/)[cation for V2V broadcasting. In: 2018 14th International wireless communications & mobile computing conference. IEEE; 2018, p. 440–5.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb63)

[64] [Ye H, Li GY, Juang B-H. Power of deep learn](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb63)[ing for channel estima-tion and signal detection in OFDM systems. IEEE Wireless Commun Lett 2017;7(1):114–7.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb64)

[65] [Mnih V, Kavukcu](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb64)oglu K, Silver D, Graves A, Antonoglou I, Wierstra D, et al.

Playing atari with deep reinforcement learning. 2013, arXiv preprint [arXiv: 1312.5602](http://arxiv.org/abs/1312.5602).

[66] [Silver D, Huang A, Maddison CJ, Guez A, Sifre L, Van Den Driessche G, et al.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb66)

[Mastering the game of go with deep neural networks and tree search. Nature 2016;529(7587):484.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb66)

[67] [Li Y. Deep reinforce](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb66)ment learning: An overview. 2017, arXiv preprint [arXiv:](http://arxiv.org/abs/1701.07274)  [1701.07274](http://arxiv.org/abs/1701.07274).

[68] Karagiannakos S. The idea behind actor-critics and how A2C and A3C improve them. 2018, [Online]. Available: <https://sergioskar.github.io/Actor-critics/>.

[69] [Singh](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb69)  [S,](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb69)  [Jaakkola](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb69)  [T,](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb69)  Littman [ML,](https://sergioskar.github.io/Actor-critics/)  [Szepesvári](https://sergioskar.github.io/Actor-critics/)  [C.](https://sergioskar.github.io/Actor-critics/)  [Convergence](https://sergioskar.github.io/Actor-critics/)  [re](https://sergioskar.github.io/Actor-critics/)[sults](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb69)  [for](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb69)  [single-step](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb69)  [on-policy](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb69)  [reinforcement-learning](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb69)  [algorithms.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb69)  [Mach](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb69)  [Learn](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb69)  [2000;38(3):287–308.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb69)

[70] [Alpcan T, Boche H,](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb69) [Honig ML, Poor HV. Mechanisms and games for dynamic](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb70)  [spectrum allocation. Cambridge University Press; 2013.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb70)

[71] [Perlaza SM, Tembine H, Lasaulce S. How can ignoran](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb70)[t but patient cognitive terminals learn their strategy and utility? In: 2010 IEEE 11th international workshop on signal processing advances in wireless communications. IEEE; 2010, p. 1–5.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb71)

[72] [Anthony M,](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb71) [Bartlett PL. Neural network learning: Theoretical foundations.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb72)  [cambridge University Press; 2009.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb72)

[73] [Ruder S. An overview of gradient](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb72) descent optimization algorithms. 2016, arXiv preprint [arXiv:1609.04747](http://arxiv.org/abs/1609.04747).

[74] [Niepert](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb74) [M, Ahmed M, Kut](http://arxiv.org/abs/1609.04747)[zkov K. Learning convolutional neural networks for](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb74)  [graphs. In: International conference on machine learning. 2016, p. 2014–23.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb74) [75] [Yamashita R, Nishio M, Do RKG, Togashi K. Convolutional neural networks: a](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb74)[n](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb75)  [overview and application in radiology. Insights Imaging 2018;9(4):611–29.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb75)

[76] [Islam M, Chen G, Jin S. An overview of neural network. Am J Neural N](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb75)[etw](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb76)  [Appl 2019;5(1):7–11.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb76)

[77] [Taylor ME, Stone P.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb76) [Transfer learning for reinforcement learning domains: A](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb77)  [survey. J Mach Learn Res 2009;10(Jul):1633–85.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb77)

[78] [Kazmi SA, Tran NH, Saad W, Le LB, Ho T](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb77)[M, Hong CS. Optimized re-source management in heterogeneous wireless networks. IEEE Commun Lett 2016;20(7):1397–400.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb78)

[79] [Kaur A, Kumar K. A r](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb78)[einforcement learning based evolutionary multi-objective optimization algorithm for spectrum allocation in cognitive radio networks. Phys Commun 2020;43:101196.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb79)

[80] [Naparstek O, Cohen K. Deep m](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb79)[ulti-user reinforcement learning for distributed](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb80)  [dynamic spectrum access. IEEE Trans Wireless Commun 2018;18(1):310–23.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb80) [81] [Srinivasan M, Kotagi VJ, Murthy CSR. A Q-learning framework for user QoE e](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb80)[n-](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb81) [hanced self-organizing spectrally efficient network using a novel inter-operator](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb81)  [proximal spectrum sharing. IEEE J Sel Areas Commun 2016;34(11):2887–901.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb81) [82] [Fan C, Li B, Zhao C, Guo W, Liang Y-C. Learning-based spectrum sharing](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb82)  [and spatial reuse in mm-wave ultradense networks. IEEE Trans Veh Technol](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb82)  [2017;67(6):4954–68.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb82)

[83] [Zhang Y, Tay WP,](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb82) [Li KH, Esseghir M, Gaïti D. Learning temporal–spatial](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb83)  [spectrum reuse.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb83)  [IEEE Trans Commun 2016;64(7):3092–103.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb83)

[84] [Sun Y, Peng M, Poor HV. A distributed approach to i](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb83)[mproving spectral efficiency in uplink device-to-device-enabled cloud radio access networks. IEEE Trans Commun 2018;66(12):6511–26.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb84)

[85] [Alnwaimi G, Vahid S, Moessner K. Dy](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb84)[namic heterogeneous learning games for opportunistic access in LTE-based macro/femtocell deployments. IEEE Trans Wireless Commun 2014;14(4):2294–308.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb85)

[86] [Chen M, Saad W, Yin C. Echo state](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb85) [networks for self-organizing resource allocation in LTE-U with uplink–downlink decoupling. IEEE Trans Wireless Commun 2016;16(1):3–16.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb86)

44

*S. Ashtari et al.*  *Array 14 (2022) 100136*

[117] [He Y, Zhang Z, Yu FR, Zhao N, Yin H, Leung VC, et al. Deep-reinforcement-](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb117) [learning-based](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb117)  [optimization](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb117)  [for](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb117)  [cache-enabled](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb117)  [opportunistic](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb117)  [interference](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb117)  [alignment wireless networks. IEEE Trans Veh Technol 2017;66(11):10433–45.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb117) [118] [Zhong C, Gursoy MC, Velipasalar S. A deep reinforcement learning-based frame](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb117)[-](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb118) [work for content caching. In: 2018 52nd Annual conference on information](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb118)  [sciences and systems. IEEE; 2018, p. 1–6.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb118)

[119] [He Y, Zhao N, Yin H. Integrated netw](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb118)[orking, caching, and computing for connected vehicles: A deep reinforcement learning approach. IEEE Trans Veh Technol 2017;67(1):44–55.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb119)

[120] [Baştuğ E, Bennis M, Debba](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb119)[h M. A transfer learning approach for cache-enabled wireless networks. In: 2015 13th International symposium on modeling and optimization in mobile, Ad Hoc, and wireless networks. IEEE; 2015, p. 161–6.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb120) [121] [Bastug E, Bennis M, Debbah M. Anticipatory caching in small cell networks: A](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb120) [transfer learning approach. In: 1st KuVS workshop on anticipatory networks. 2014.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb121)

[122] [Bharath B, Nagananda KG, Poor HV. A learning-based approach to caching in](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb122)  [heterogenous small cell networks.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb122)  [IEEE Trans Commun 2016;64(4):1674–86.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb122) [123] [Venables WN, Ripley BD. Modern applied statistics with S-PLUS. Springe](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb122)[r](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb123)  [Science & Business Media; 2013.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb123)

[124] [Hyvärinen A, Karhunen J, Oja](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb123) [E. Independent component analysis, vol. 46.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb124)  [Hoboken. NJ, USA: Wiley; 2004.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb124)

[125] [Abrams M, Standridge CR, Abd](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb124)[ulla G, Williams S, Fox EA. Caching proxies:](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb125)  [Limitations and potentials. 1995.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb125)

[126] [Arlitt M, Cherkasova L, Dilley](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb125) [J, Friedrich R, Jin T. Evaluating content management techniques for web proxy caches. ACM SIGMETRICS Perform Eval Rev 2000;27(4):3–11.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb126)

[127] [Psounis K, Prabhaka](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb126)[r B. A randomized web-cache replacement scheme. In: Proceedings IEEE INFOCOM 2001. Conference on computer communications.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb127)

[Twentieth annual joint conference of the IEEE computer and communications society (Cat. No. 01CH37213), Vol. 3. IEEE; 2001, p. 1407–15.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb127)

[128] [Dulac-Arnold G, Evans R, van Hasselt H, Sunehag P, Lillicrap](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb127) T, Hunt J, et al. Deep reinforcement learning in large discrete action spaces. 2015, arXiv preprint [arXiv:1512.07679](http://arxiv.org/abs/1512.07679).

[129] [Lee D, C](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb129)[hoi J, Kim J-H,](http://arxiv.org/abs/1512.07679) [Noh SH, Min SL, Cho Y, et al. On the existence of a spectrum of policies that subsumes the least recently used (LRU) and least frequently used (LFU) policies. In: Proceedings of the 1999 ACM SIGMETRICS international conference on measurement and modeling of computer systems. 1999, p. 134–43.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb129)

[130] [Bastug E, Bennis](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb129) [M, Debbah M. Living on the edge: The role of proactive caching in 5G wireless networks. IEEE Commun Mag 2014;52(8):82–9.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb130)

[131] [Blaszczyszyn B, Giovanidis A. Optimal geographic caching in cellular ne](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb130)[tworks.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb131)

[In: 2015 IEEE international conference on communications. IEEE; 2015, p. 3358–63.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb131)

[132] [Liu A, Lau VK. Exploiting base station caching in MIMO cellular networks: Opportunistic cooperation for video streaming. IEEE Trans Signal Process 2014;63(1):57–69.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb132)

[133] [Hamidouche K, Sa](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb132)[ad W, Debbah M, Song JB, Hong CS. The 5G cellular backhaul management dilemma: To cache or to serve. IEEE Trans Wireless Commun 2017;16(8):4866–79.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb133)

[134] [Jaber M, Imran M,](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb133) [Tafazolli R, Tukmanov A. An adaptive backhaul-aware cell range extension approach. In: 2015 IEEE international conference on communication workshop. IEEE; 2015, p. 74–9.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb134)

[135] [Xu Y, Yin R, Yu G. Adaptive biasing scheme f](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb134)[or load balancing in backhaul](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb135)  [constrained small cell networks. IET Commun 2015;9(7):999–1005.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb135)

[136] [Fadlullah ZM, Tang F, Mao B, Kato N, Akashi O, Inoue T, et al. S](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb135)[tate-of-the-art deep learning: Evolving machine intelligence toward tomorrow’s intelligent network traffic control systems. IEEE Commun Surv Tutor 2017;19(4):2432–55.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb136) [137] [Yanjun L, Xiaobo L, Osamu Y. Traffic engineering framework with machine](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb136) [learning based meta-layer in software-defined networks. In: 2014 4th IEEE international conference on network infrastructure and digital content. IEEE; 2014, p. 121–5.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb137)

[138] [Chen Y-S, Hsu](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb137) [C-S, Siao Y-G. Linear regression-based delay-bounded routing](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb138)  [protocols for VANETs. Wireless Commun Mob Comput 2014;14(2):186–99.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb138)

[139] [Wang Y, Martonosi M, Peh L-S. A supervised learning approach for rou](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb138)[ting optimizations in wireless sensor networks. In: Proceedings of the 2nd interna-tional workshop on multi-hop ad hoc networks: From theory to reality. 2006, p. 79–86.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb139)

[140] [Sharma](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb139) [DK, Dhurandher SK, Woungang I, Srivastava RK, Mohananey A, Rodrigues JJ. A machine learning-based protocol for efficient routing in opportunistic networks. IEEE Syst J 2016;12(3):2207–13.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb140)

[141] [Tang F, Mao B, Fadlullah ZM, Kato N, Akashi O, Inoue](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb140) [T, et al. On removing routing protocol from future wireless networks: A real-time deep learning approach for intelligent traffic control. IEEE Wirel Commun 2017;25(1):154–60.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb141) [142] [Kato N, Fadlullah ZM, Mao B, Tang F, Akashi O, Inoue T, et al. The deep](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb141) [learning vision for heterogeneous network traffic control: Proposal, challenges, and future perspective. IEEE Wirel Commun 2016;24(3):146–53.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb142)

[143] [Yao H, Yuan X, Zhang P, Wang J, Jiang C, Guizani M. A ma](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb142)[chine learning approach of load balance routing to support next-generation wireless networks.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb143)

[In: 2019 15th International wireless communications & mobile computing conference. IEEE; 2019, p. 1317–22.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb143)

45

*S. Ashtari et al.*  *Array 14 (2022) 100136*

[172] [Kudo](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb172)  [T,](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb172)  [Ohtsuki](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb172)  [T.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb172)  [Cell](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb172)  [range](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb172)  [expansion](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb172)  [using](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb172)  [distributed](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb172)  [Q-learning](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb172)

[in](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb172)  [heterogeneous](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb172)  [networks.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb172)  [EURASIP](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb172)  [J](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb172)  [Wireless](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb172)  [Commun](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb172)  [Networking](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb172)

[2013;2013(1):61.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb172)

[173] [Li Z, Wang C,](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb172) [Jiang C-J. User association for load balancing in vehicular](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb173)

[networks: An online reinforcement learning approach. IEEE Trans Intell Transp](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb173)

[Syst 2017;18(8):2217–28.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb173)

[174] [Pervez F, Jaber M, Qadir](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb173) [J, Younis S, Imran MA. Fuzzy Q-learning-based user-](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb174)

[centric backhaul-aware user cell association scheme. In: 2017 13th International](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb174)

[wireless communications and mobile computing conference. IEEE; 2017, p.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb174)

[1840–5.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb174)

[175] [Ding H, Zhao F, Tian J, Li D, Zhang H. A deep reinforcement learning for](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb175)

[user association and power control in heterogeneous networks. Ad Hoc Netw](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb175)

[2020;102:102069.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb175)

[176] [Chou P-Y, Chen](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb175) [W-Y, Wang C-Y, Hwang R-H, Chen W-T. Deep reinforcement](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb176)

[learning for MEC streaming with joint user association and resource manage-](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb176)

[ment. In: ICC 2020-2020 IEEE international conference on communications.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb176)

[IEEE; 2020, p. 1–7.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb176)

[177] [Zhang Q, Liang Y-C, Poor HV. Intelligent user association for symbiotic radio](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb177)

[networks using deep reinforcement learning. IEEE Trans Wireless Commun](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb177)

[2020.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb177)

[178] [Zhang Q, Zhang L, Liang Y-C, Kam P-Y. Backscatter-NOMA: A symbiotic system](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb178)

[of cellular and internet-of-things networks. IEEE Access 2019;7:20000–13.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb178)

[179] [Long R, Guo H, Zhang L, Liang Y-C. Full-duplex backscatter communica](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb178)[tions](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb179)

[in symbiotic radio systems. IEEE Access 2019;7:21597–608.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb179)

[180] [Guo H, Liang Y-C, Long R, Xiao S, Zhang Q. Resource allocation for symbiotic](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb180)

[radio system with fading channels. IEEE Access 2019;7:34333–47.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb180)

[181] [Raikar MM, Meena S, Mulla MM, Shetti NS, Karanandi M.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb180) [Data traffic](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb181)

[classification in software defined networks (SDN) using supervised-learning.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb181)

[Procedia Comput Sci 2020;171:2750–9.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb181)

[182] [Sun G, Liang L, Chen T, Xiao F, Lang F. Network traffic classification based on](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb182)

[transfer learning. Comput Electr Eng 2018;69:920–7.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb182)

[183] [Yu C, Liu Y, Yao D, Yang LT, Jin H, Chen H, et al. Modeling user activity](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb183)

[patterns for next-place prediction. IEEE Syst J 2015;11(2):1060–71.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb183)

[184] [Sinclair N, Harle D, Glover IA, Irvine J, Atkinson RC. An adva](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb183)[nced SOM](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb184)

[algorithm applied to handover management within LTE. IEEE Trans Veh](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb184)

[Technol 2013;62(5):1883–94.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb184)

[185] [Castro-Hernandez D, Paranjap](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb184)[e R. Classification of user trajectories in LTE Het-](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb185)

[Nets using unsupervised shapelets and multiresolution wavelet decomposition.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb185)

[IEEE Trans Veh Technol 2017;66(9):7934–46.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb185)

[186] [Wu C-L, Chiu T-C, Wang C-Y, Pang A-C. Mobility-aware deep reinforcement](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb186)

[learning with glimpse mobility prediction in edge computing. In: ICC 2020-2020](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb186)

[IEEE international conference on communications. IEEE; 2020, p. 1–7.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb186)

[187] [Zhang H, Dai L. Mobility prediction: A survey on state-of-the-art schemes and](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb187)

[future applications. IEEE Access 2018;7:802–22.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb187)

[188] [Dang](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb187)  [F,](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb187)  [Yang](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb187)  [X,](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb187)  [Long](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb187)  [K.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb187)  [Spray](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb187)  [and](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb187)  [forw](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb187)[ard:](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb188)  [Efficient](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb188)  [routing](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb188)  [based](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb188)

[on the Markov location prediction model for DTNs. Sci China Inf Sci](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb188)

[2012;55(2):433–40.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb188)

[189] [Song C, Qu Z, Bl](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb188)[umm N, Barabási A-L. Limits of predictability in human](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb189)

[mobility. Science 2010;327(5968):1018–21.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb189)

[190] [Monreale A, Pinelli F, Trasarti R, Giannotti](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb189) [F. Wherenext: a location predictor](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb190)

[on trajectory pattern mining. In: Proceedings of the 15th ACM SIGKDD](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb190)

[international conference on knowledge discovery and data mining. 2009, p.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb190)

[637–46.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb190)

[191] [Cicioğlu M. Performance analysis of handover management in 5G small cells.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb191)

[Comput Stand Interfaces 2021;75:103502.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb191)

[192] [Ali Z, Baldo N, Mangues-Bafalluy J, Giupponi L. Machine learning based han-](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb192)

[dover management for improved QoE in LTE. In: NOMS 2016-2016 IEEE/IFIP](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb192)

[network operations and management symposium. IEEE; 2016, p. 794–8.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb192)

[193] [Lee C, Cho H, Song S, Chung J-M. Prediction-based conditional handover for](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb193)

[5G mm-wave networks: A deep-learning approach. IEEE Veh Technol Mag](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb193)

[2020;15(1):54–62.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb193)

[194] [Shubyn B, Lutsiv](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb193) [N, Syrotynskyi O, Kolodii R. Deep learning based adap-](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb194)

[tive handover optimization for ultra-dense 5g mobile networks. In: 2020](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb194)

[IEEE 15th international conference on advanced trends in radioelectronics,](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb194)

[telecommunications and computer engineering. IEEE; 2020, p. 869–72.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb194)

[195] [Aljeri](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb194)  [N,](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb194)  [Boukerche](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb194)  [A.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb194)  [A](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb194)  [two-tier](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb194)  [machine](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb194)  [learning-based](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb194)  [h](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb194)[andover](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb195)

[management](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb195)  [scheme](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb195)  [for](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb195)  [intelligent](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb195)  [vehicular](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb195)  [networks.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb195)  [Ad](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb195)  [Hoc](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb195)  [Netw](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb195)

[2019;94:101930.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb195)

[196] [Wu J, Liu J, H](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb195)[uang Z, Zheng S. Dynamic fuzzy Q-learning for handover](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb196)

[parameters optimization in 5g multi-tier networks. In: 2015 International](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb196)

[conference on wireless communications & signal processing. IEEE; 2015, p. 1–5.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb196)

[197] [Dhahri C, Ohtsuki T. Q-learning cell selection for femtocell networks: Single-and](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb197)

[multi-user case. In: 2012 IEEE global communications conference. IEEE; 2012,](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb197)

[p. 4975–80.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb197)

[198] [Dhahri C, O](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb197)[htsuki T. Adaptive Q-learning cell selection method for open-access](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb198)

[femtocell networks: Multi-user case. IEICE Trans Commun 2014;97(8):1679–88.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb198)

[199] [Wang Z, Li L, Xu Y, Tian H, Cui S. Handover control in wireless systems via](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb199)

[asynchronous multiuser deep reinforcement learning. IEEE Internet Things J](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb199)

[2018;5(6):4296–307.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb199)

46

|  |  |  |
| --- | --- | --- |
| *S. Ashtari et al.* |  | *Array 14 (2022) 100136* |
| [229] [Ahmed KI, Tabassum H, Hossain E. Deep learning for radio resource allocation](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb229)  [in multi-cell networks. IEEE Network 2019;33(6):188–95.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb229)  [230] [Lee W, Kim M, Cho D-H. Deep power control: Transmit p](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb229)[ower control scheme](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb230)  [based on convolutional neural network. IEEE Commun Lett 2018;22(6):1276–9.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb230)  [231] [Goodfellow I, Bengio Y, Courville A, Bengio Y. Deep learning, Vol. 1, no. 2.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb230)  [MIT press Cambridge; 2016.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb231)  [232] [Victor K. Video hosting s](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb231)ervice based in Beijing, China. 2003, [Online].  Available: <https://www.youku.com/>.  [233] [Derpanis KG, Lecce M, Daniilidis K, Wildes RP. Dynamic scene understanding:](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb233)  [The role of orientation features in space and time in scene classification. In:](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb233)  [2012 IEEE conference on computer vision and pattern recognition. IEEE; 2012,](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb233)  [p. 1306–13.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb233)  [234] [Sun Y, Peng](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb233) [M, Mao S. Deep reinforcement learning-based mode selection and](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb234)  [resource management for green fog radio access networks. IEEE Internet Things](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb234)  [J 2018;6(2):1960–71.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb234)  [235] [Feng M, Mao S. D](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb234)[ealing with limited backhaul capacity in millimeter-](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb235)  [wave systems: A deep reinforcement learning approach. IEEE Commun Mag](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb235)  [2019;57(3):50–5.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb235)  [236] [Xiao K, Mao S, T](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb235)[ugnait JK. TCP-drinc: Smart congestion control based on deep](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb236)  [reinforcement learning. IEEE Access 2019;7:11892–904.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb236)  [237] [Shi D, Ding J, Errapotu SM, Yue H, Xu W, Zhou X, et al.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb236) [Deep](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb237) *[𝑄](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb237)*[-network-based](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb237)  [route scheduling for TNC vehicles with passengers’ location differential privacy.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb237)  [IEEE Internet Things J 2019;6(5):7681–92.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb237)  [238] [Yu](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb237)  [Y,](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb237)  [Wang](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb237)  [T,](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb237)  [Liew](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb237)  [SC.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb237)  [Deep-rein](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb237)[forcement](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb238)  [learning](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb238)  [multiple](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb238)  [ac-](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb238)  [cess](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb238)  [for](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb238)  [heterogeneous](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb238)  [wireless](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb238)  [networks.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb238)  [IEEE](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb238)  [J](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb238)  [Sel](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb238)  [Areas](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb238)  [Commun](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb238)  [2019;37(6):1277–90.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb238)  [239] [Huang C, Hu S, Alexandropoulos GC, Zappone A, Yuen C, Zhang R, et al.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb239)  [Holographic MIMO surfaces for 6G wireless networks: Opportunities, challenges,](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb239)  [and trends. IEEE Wirel Commun 2020;27(5):118–25.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb239)  [240] [Tan Y, Yang J, Gopalakrishnan N. Self-learning, ad](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb239)[aptive approach for intel-](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb240)  [ligent analytics-assisted self-organizing-networks (SONs). 2019, US Patent 10,](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb240)  [382, 979.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb240)  [241] [Feng S, S](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb240)[eidel E. Self-organizing networks (SON) in 3GPP long term evolution.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb241)  [Nomor Res GmbH 2008;20.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb241)  [242] [Blasco P, Bennis M, Dohler](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb241) [M. Backhaul-aware self-organizing operator-shared](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb242)  [small cell networks. In: 2013 IEEE international conference on communications.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb242)  [IEEE; 2013, p. 2801–6.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb242)  [243] [Jaber M, Imran MA, Ta](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb242)[fazolli R, Tukmanov A. A multiple attribute user-centric](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb243)  [backhaul provisioning scheme using distributed SON. In: 2016 IEEE global](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb243)  [communications conference. IEEE; 2016, p. 1–6.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb243)  [244] [Cao G, Lu Z, Wen X, Lei T, Hu Z. AIF: An artifi](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb243)[cial intelligence framework for](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb244)  [smart wireless network management. IEEE Commun Lett 2017;22(2):400–3.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb244)  [245] [Amiri R, Almasi MA, Andrews JG, Mehrpouyan H. Reinforcement learning](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb244) [for](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb245)  [self organization and power control of two-tier heterogeneous networks. IEEE](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb245)  [Trans Wireless Commun 2019;18(8):3933–47.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb245)  [246] [Silva MAL, de Souza SR, Souza MJF, Bazzan ALC. A reinforcement learning-](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb246)  [based multi-agent framework applied for solving routing and scheduling](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb246)  [problems. Expert Syst Appl 2019;131:148–71.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb246)  [247] [Galindo-Serrano A, Giupponi L, Auer G. D](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb246)[istributed learning in multiuser](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb247)  [OFDMA](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb247)  [femtocell](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb247)  [networks.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb247)  [In:](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb247)  [2011](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb247)  [IEEE](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb247)  [73rd](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb247)  [vehicular](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb247)  [technology](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb247)  [conference. IEEE; 2011, p. 1–6.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb247)  [248] [Chen M, Saad W, Yin C. Virt](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb247)[ual reality over wireless networks: Quality-of-](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb248)  [service model and learning-based resource management.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb248)  [IEEE Trans Commun](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb248)  [2018;66(11):5621–35.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb248)  [249] [Samarakoon S, Bennis](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb248) [M, Saad W, Latva-aho M. Backhaul-aware interference](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb249)  [management in the uplink of wireless small cell networks. IEEE Trans Wireless](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb249)  [Commun 2013;12(11):5813–25.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb249)  [250] [Chen X, Zhang H, Wu C, Mao](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb249) [S, Ji Y, Bennis M. Optimized computation of-](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb250)  [floading performance in virtual edge computing systems via deep reinforcement](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb250)  [learning. IEEE Internet Things J 2018;6(3):4005–18.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb250)  [251] [Klein A, Kuruvatti NP, Schneider J, Schotten HD.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb250) [Fuzzy Q-learning for mo-](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb251)  [bility robustness optimization in wireless networks. In: 2013 IEEE globecom](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb251)  [workshops. IEEE; 2013, p. 76–81.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb251)  [252] [Ma B, Guo W, Zhang J. A survey](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb251) [of online data-driven proactive 5G network](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb252)  [optimisation using machine learning. IEEE Access 2020;8:35606–37.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb252)  [253] [Saruta K, Kato N, Abe M, Nemoto Y. High accuracy recognition of ETL9b using](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb253)  [exclusive learning neural network-< cd02d36. gif>(ELNET-< cd02d36. gif>).](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb253)  [IEICE Trans Inf Syst 1996;79(5):516–22.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb253)  [254] [Silver D, Schrittwieser J, Simonyan K](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb253)[, Antonoglou I, Huang A, Guez A,](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb254)  [et](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb254)  [al.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb254)  [Mastering](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb254)  [the](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb254)  [game](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb254)  [of](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb254)  [go](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb254)  [without](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb254)  [human](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb254)  [knowledge.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb254)  [Nature](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb254)  [2017;550(7676):354–9.](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb254)  [255] [Paullada A, Raji ID, Ben](http://refhub.elsevier.com/S2590-0056(22)00007-8/sb254)der EM, Denton E, Hanna A. Data and its (dis) contents:  A survey of dataset development and use in machine learning research. 2020,  arXiv preprint [arXiv:2012.05345](http://arxiv.org/abs/2012.05345).  [256] Kahn G, Abb[eel](http://arxiv.org/abs/2012.05345)  [P,](http://arxiv.org/abs/2012.05345)  [Levine](http://arxiv.org/abs/2012.05345)  S. Badgr: An autonomous self-supervised  learning-based navigation system. 2020, arXiv preprint [arXiv:2002.05700](http://arxiv.org/abs/2002.05700). | **Sepehr Ashtari** (Graduate Student Member, IEEE) received the B.S. degree in electrical and electronic engineering from Eastern Mediterranean University (EMU), North Cyprus, Turkey, in 2016 and the M.S. degree in telecommunica-tion engineering from the University of New South Wales (UNSW), Sydney, Australia, in 2019. He is currently pursu-ing the Ph.D. degree in telecommunication and information technology at the University of Technology (UTS), Sydney, Australia.  From 2015 to 2016, he was a Research Assistant with the Department of Electrical and Electronic Engineering, North Cyprus, Turkey. He has worked as a research fellow at the Commonwealth Scientific and Industrial Research Organisation (CSIRO) on several wireless communication projects. His research interest includes improvement of wireless cellular communication, routing protocols, software or knowledge-based networking, 5G resource allocation, and machine learning optimization in wireless networks.  Mr. Ashtari’s awards and honors include the Fellowship of Excellence from Department of Electrical and Electronic Engineering, the High Honors degree in B.S. as well as M.S., and Fellowship award from UTS Faculty of Engineering and IT. |
|  | **Ian Zhou** (Graduate Student Member, IEEE) received the B.S. degree in computer science from the University of Sydney, Australia, in 2016. He also received the MBA degree from the University of Technology Sydney in 2019. His research interest includes AI, IoT, cyber–physical system, and blockchain. Currently, he is pursuing a Ph.D. degree at the University of Technology Sydney researching on machine learning-based frost monitoring systems. |
|  | **Mehran Abolhasan** (Senior Member, IEEE) is currently an Associate Professor and the Deputy Head of the School of Electrical and Data Engineering, University of Technology Sydney. He has over 20 years of experience in research and development and serving in research leadership roles. Some of these previous roles include serving as the Director of research programs for the Faculty of Engineering and IT, and the Laboratory Director for the Telecommunication and IT Research Institute, University of Wollongong. He has authored over 160 international publications and has won over four million dollars in research funding. He won a number of major research project grants, including the ARC Discovery Project, ARC Linkage Project, and a number of CRC and other government and industry-based grants. He currently leads the Software-Defined Networks Lab at UTS and his current research interests include software defined networking, the IoT, wireless mesh, wireless body area networks, cooperative networks, 5G networks and beyond, and sensor networks. |
|  | **Negin Shariati** (Member, IEEE) is a Senior Lecturer in the School of Electrical and Data Engineering, Faculty of Engineering and IT, University of Technology Sydney (UTS), Australia. She established the state of the art RF and Com-munication Technologies (RFCT) research laboratory at UTS in 2018, where she is currently the Co-Director and leads research and development in RF-Electronics, Sustainable Sensing, Low-power Internet of Things, and Energy Harvest-ing. She leads the Sensing Innovations Constellation at Food Agility CRC (Corporative Research Centre), enabling new innovations in agriculture technologies by focusing on three key interrelated streams; Energy, Sensing and Connectivity.  Since 2018, she has held a joint appointment as a Senior Lecturer at Hokkaido University, externally engaging with research and teaching activities in Japan. She attracted over $850K worth of research funding across a number of CRC and industry projects, where she has taken the lead CI role and also contributed as a member of the CI team. Negin Shariati completed her Ph.D. in Electrical-Electronic and Communication Technologies at Royal Melbourne Institute of Technology (RMIT), Australia, in 2016. She worked in industry as an Electrical-Electronic Engineer from 2009–2012. Her research interests are in Microwave Circuits and |

47

|  |  |  |  |
| --- | --- | --- | --- |
| *S. Ashtari et al.* | Systems, RF Energy Harvesting, low-power IoT, Simulta-neous Wireless Information and Power Transfer, Wireless Networks, AgTech, and Renewable Energy Systems. |  | *Array 14 (2022) 100136* |
|  | **Wei Ni** (Senior Member, IEEE) received the B.E. and Ph.D. degrees in Electronic Engineering from Fudan University, Shanghai, China, in 2000 and 2005, respectively. Currently, he is a Group Leader and Principal Research Scientist at CSIRO, Sydney, Australia, and an Adjunct Professor at the University of Technology Sydney and Honorary Professor at Macquarie University. He was a Postdoctoral Research Fellow at Shanghai Jiaotong University from 2005–2008; Deputy Project Manager at the Bell Labs, Alcatel/Alcatel-Lucent from 2005 to 2008; and Senior Researcher at Devices R&D, Nokia from 2008 to 2009. His research interests include machine learning, stochastic optimization, online learning, as well as their applications to system efficiency and integrity.  Dr Ni is the Chair of IEEE Vehicular Technology Society (VTS) New South Wales (NSW) Chapter since 2020 and an Editor of IEEE Transactions on Wireless Communications since 2018. He served first the Secretary and then Vice-Chair of IEEE NSW VTS Chapter from 2015 to 2019, Track Chair for VTC-Spring 2017, Track Co-chair for IEEE VTC-Spring 2016, Publication Chair for BodyNet 2015, and Student Travel Grant Chair for WPMC 2014. |
| **Justin Lipman** (Senior Member, IEEE) received a Ph.D. in Telecommunications Engineering from University of Wol-longong, Australia in 2004. He is an Industry Associate Professor at the University of Technology Sydney (UTS) and a visiting Associate Professor at Hokkaido University’s Graduate School of Engineering. He is the Director of Research Translation in the Faculty of Engineering and IT and is Director of the RF Communications Technologies (RFCT) Lab — where he leads industry engagement in RF technologies, Internet of Things, Tactile Internet and Software Defined Communication. He serves as committee member in Standards Australia contributing to International IoT standards and Digital Twins. Prior to joining UTS, he was based in Shanghai, China and held a number of senior management and technical leadership roles at Intel and Alcatel driving research and innovation, product develop-ment, architecture and IP generation. He is an IEEE Senior Member. His research interests are in all ‘‘things’’ adaptive, connected, distributed and ubiquitous. |

48