Machine Learning Based Unmanned Aerial Vehicle Enabled Fog-Radio Access Network and Edge Computing

Mohammed SEID[1,2], Stephen ANOKYE[1,3], and SUN Guolin[1]

(1. University of Electronic Science and Technology of China, Chengdu, Sichuan 611731, China;
2. Dilla University, Dilla, Ethiopia;
3. University of Mines and Technology, Tarkwa 237, Ghana)

Abstract: The emerging unmanned aerial vehicle (UAV) technology and its applications have become part of the massive Internet of Things (mIoT) ecosystem for future cellular networks. Internet of things (IoT) devices have limited computation capacity and battery life and the cloud is not suitable for offloading IoT tasks due to the distance, latency and high energy consumption. Mobile edge computing (MEC) and fog radio access network (F-RAN) together with machine learning algorithms are an emerging approach to solving complex network problems as described above. In this paper, we suggest a new orientation with UAV enabled F-RAN architecture. This architecture adopts the decentralized deep reinforcement learning (DRL) algorithm for edge IoT devices which makes independent decisions to perform computation offloading, resource allocation, and association in the aerial to ground (A2G) network. Additionally, we summarized the works on machine learning approaches for UAV networks and MEC networks, which are related to the suggested architecture and discussed some technical challenges in the smart UAV-IoT, F-RAN 5G and Beyond 5G (6G).

Keywords: unmanned aerial vehicle; machine learning; F-RAN; edge computing

DOI: 10.12142/ZTECOM.201904006

http://kns.cnki.net/kcms/detail/34.1294.TN.20191206.1105.006.html, published online December 9, 2019

Manuscript received: 2019-10-27

1 Introduction

In the recent past, cellular technologies have become more dynamic and improved the network infrastructure to the satisfaction of end users. There are a number of ultra-dense heterogeneous devices from individuals and organizations, which are always generating and storing a huge amount of data via sensors (edge Internet of Things (IoT) devices) and applications [1]. When the massive Internet of Things (mIoT) devices emerge, the data generated by various sensors will increase exponentially. Due to the huge volume of the data produced and different forms of conventional databases (with structured and unstructured data), big data analysis has attracted much attention in recent years and many organizations have focused on the analysis of collected data to extract useful data for making appropriate decisions [2]. The data generated from billions of heterogeneous IoT sensors are sent to the cloud for processing computing tasks, with a high cost of processing delay and energy consumption. However, some IoT sensors data need to be processed faster than the current processing capability of clouds. To solve this problem, fog and edge computing (FEC) is proposed to enable computing tasks processed at the network edge of IoT [3]–[5]. Edge computing is a new emerging paradigm to solve IoT computation and resource allocation problem in localized manner [5]. Fog computing is decentralized computing paradigm, where a number of smart devices which have a computational capacity are utilized [6], [7]. In this paradigm, key issues were discussed about the requirement and deployment of fog connectivity environment due to the existence of ultra-dense heterogeneous devices. Several technical issues on fog computing such as deployment, simulation, resource management, fault tolerance and services have been studied in [6], [8] – [13]. Even though fog computing and edge computing both move the computation and storage to the edge of the network, closer to end-nodes, their paradigms are not identical [14]. The rapid development of diverse mIoT devices such as wireless sensors, smart machines, and mobile users' applications enable the users to enjoy high quality of experience (QoE) and high quality of service (QoS) [5], [15], [16]. However, most of these applications are delay sensitive or real-time applications, which need high computational capacity. The edge devices could not compute each task due to the limitation of battery and low computation capability, so it is difficult for them to implement these applications [17]. The FEC can compute tasks of IoT devices and interplay with the cloud server to provide better QoS and QoE to end users. Some works were done on computation offloading to mobile edge computing (MEC) servers and on resource allocation of the IoT devices to maximize network performance and optimize the problem in ultra-dense heterogeneous network [18] – [20]. For ultra-dense IoT network system, a game theory computation offloading framework was designed in [21] and [22], to minimize the overall computation overhead of the task on edge IoT devices.

Radio access network (RAN) provides connectivity to the wireless terminals through wireless access points (base stations) and may use one or more radio access technologies (RATs). The fog radio access network (F-RAN) is composed of F-RAN nodes connected through a single or multiple RATs. The F-RAN has a unique feature better than the cloud radio access network (CRAN) and heterogeneous cloud radio access network (H-CRAN), which helps maximize the use of edge IoT devices of the network and improve network management and optimization mechanisms [5], [23], [24].

Based on the report from Federal Aviation Administration (FAA) [25], the fleet of drones will be more than doubled from estimated 1.1 million vehicles in 2017 to 2.4 million units by 2022. Benefitting from connecting unmanned aerial vehicles (UAVs) to cellular networks for better control and communications, the growth of the UAV market is expected to bring new promising business opportunities for cellular operators. Millions of UAVs have been used to perform various services such as public protection, disaster relief operation, surveillance applications, traffic management, commercial services, extending the cellular-network coverage to remote areas, and acting as flying base stations [26], [27]. The Third Generation Partnership Project (3GPP) is exploring the challenges and opportunities for serving UAVs as a new type of User Equipment (UE), called aerial UE. UAVs can facilitate the development of IoT ecosystems for mIoT applications [28]. UAVs will be the future of IoT because UAVs, at the beginning, efficiently replace the connected sensors at rest with one device that is deployable to different locations, capable of carrying flexible payloads, re-programmable in mission, able to measure anything from anywhere, easily deployed, and cost effective. In recent years, a number of works have been done on either UAVs networks or their integration with cellular networks. Those works focused on computation offloading, maximization of energy efficiency, optimization of UAV trajectory and path planning, throughput maximization of UE in UAV network, and terrestrial heterogeneous devices.

The authors of [29] summarized the journey of machine learning in the last thirty years and the roles machine learning played in the next-generation wireless network (NGWN) as a road for achieving the ambitious goal of NGWN and as a tool for managing the network complexity. The authors of [30] emphasized the role of diverse machine learning algorithms in different key issues of networking across different network technologies. Machine learning techniques are applied for fundamental problems in networking, including routing and classification, traffic prediction, congestion control, QoS and QoE management, resource and fault management, and network security. In [31], the authors studied the advanced machine learning application in wireless communication for mobility management in the network layer, resource management in the MAC layer, and networking and localization in application layer. The paper [32] discussed the future cellular networks or wireless networks which support ultra-reliable and low-latency communications, as well as the intelligent management for mIoT devices in dynamic environment. Deep reinforcement learning (DRL) approaches for cellular networks, next generation wireless networks and self-organization cellular networks were reviewed in [29] – [34]. Recently, DRL has become one of the mostly popular machine learning algorithms for edge computing resource management and a suitable optimization technique for radio access networks. DRL has recently been used as an emerging tool for effectively solve various problems and challenges in modern networks that are more decentralized, ad-hoc, and autonomous in nature, such as heterogeneous networks (HetNets), IoT, vehicle to vehicle (V2V) system, machine to machine (M2M) system, vehicle to everything (V2I) system, self-organization cellular networks, and UAV networks [31].

Different non-deterministic polynomial-time hardness (NP-hard) problems of UAV networks and UAV connected cellular networks were optimized by adopting traditional optimization techniques [35] – [38], [40] – [43]. However, traditional optimization techniques are difficult to be applied for complex network infrastructure and not suitable for the current and future intelligent wireless networks. Recently, machine learning algorithms have been used to easily optimize different problems in UAV networks and UAV connected cellular networks [63] – [66], [68], [69]. However, there are still challenges to using machine learning algorithms for UAV networks which assist the mIoT, public safety communication (PSC), and edge computing.

The main contributions of this work are summarized as follows:

- We suggest a new orientation with UAV enabled F-RAN architecture. This architecture adopts the decentralized DRL algorithm for edge IoT devices, which enables decision independently made for offloading, resource allocation, and association in the A2G network.

• We summarize the works on machine learning approaches for UAV networks and MEC networks, which are related to the suggested architecture.

• We discuss some technical challenges in the smart UAV-IoT, F-RAN 5G, and B5G (6G).

The rest of the paper is organized as follows. We provide a brief overview of UAV in wireless cellular networks and the use of UAV for emergency situation and computation offloading in Section 2. In Section 3, we review machine learning and its classification. In Section 4, we present our orientation with UAV-enabled F-RAN in MEC, which adopts the machine learning algorithm. In Section 5, we present the works on computation offloading and resource allocation using DRL in MEC and UAV networks. In Section 6, we discuss technical challenges and future research directions of intelligent UAV enabled F-RAN at the edge level. We conclude the paper in Section 7.

2 UAV in Wireless Cellular Networks

Currently, the use of flying UAV platform is popular; this rapidly growing technology has attractive attributes such as mobility, flexibility, and adaptive attitude, and has key potential applications in wireless system. UAVs can be used as aerial base stations (ABS) to enhance coverage, capacity, reliability, and energy efficiency of wireless networks, as well as flying mobile terminals in cellular network infrastructure. UAV can be connected with cellular networks as new user equipment and help increase the revenues for network operators.

The authors of [35] summarized the current state of UAV in cellular communication system from different points of view. Different types and characteristics of UAVs are available. A number of industry-led initiatives depend on the standards of cellular communications which support low-altitude UAVs for enabling beyond Line of Sight (LoS) control and establishing a reliable communication. The deployment of flying UAV base stations is better than that of ground base stations for reducing cost and minimizing electronics equipment of base stations. The deployment of ABS faces different practical challenges such as placement and mobility, but UAV flying base stations can be easily deployed at optimum locations in 3D space; they can potentially provide much better performance in different parameters such as coverage, load balancing, spectral efficiency, and user experience, compared to existing terrestrial based solutions. UAV can act as flying base stations in the heterogeneous 5G environment and also support millimeter wave (mmW) communications; it is collectively viewed as the nexus of next-generation 5G cellular systems. UAV-enabled mmW communications is a proposing application of UAVs, which can establish LoS communication links to users [27]. UAVs can also assist various terrestrial network infrastructure such as mIoT, cellular, and vehicular networks (V2V, V2X, V2I) in different ways; for example, UAVs can improve the reliability of wireless connection and scalability, replace destroyed bases stations, compute different tasks of edge IoT devices, and relay the data or signal into central network controller. **Table 1** compares terrestrial networks with base stations and UAV networks with bases stations.

UAV at the edge level in cellular networks has a major impact on 5G and beyond. A single or multiple UAVs can compute the tasks of edge IoT devices. The UAV used as relaying and ABS which connect terrestrial smart mobile users with edge servers in MEC have been studied in [36]. To minimizing the average weighted energy consumption of the smart mobile devices and the UAV, the authors of [37] studied the multi-cell edge which is three adjacent cells served by three base stations; at the multi-cell edge, some of the users out of the radius of the base stations are connected with UAV. The problems are how to optimize the maximal sum rate of edge users by avoiding the interference and how to improve QoS and optimize UAV trajectory for the users who are out of network coverage and served by UAV.

The recent literature works on UAV network and UAV as-

▼Table 1. Comparison between UAV networks with base stations and terrestrial networks with base stations

Terrestrial Networks	UAV Networks
Insufficient spectrum	Insufficient spectrum
Well defined energy constraints and models	Elaborate and stringent energy constraints and models
Mainly static association	Varying cell association
No timing constraints, with BS being always there	Hover and flight time constraints
Terrestrial BS	UAV BS
Typical two-dimensional deployment	By nature, three-dimensional deployment
Mostly long-term and permanent deployments	Short-term and frequently changing deployments
Few and selected locations	Mostly unrestricted locations
Fixed and static	Mobility dimension
Not suitable for mobility tracking	Suitable for mobility tracking

BS: base station UAV: unmanned aerial vehicle sisted cellular user or IoT focused on computation offloading, resource allocation and path planning, and trajectory optimization of either a single UAV or multi-UAV network. In all cases, the UAV assists the terrestrial users or IoT devices in offloading tasks and in requesting resources such as power, computational resources and bandwidth. LIU et al. [38] designed UAV-Edge-Cloud computing hybrid computing architecture to jointly optimize the computation offloading and routing problem for swarms of multi-UAV which are connected in D2D forms. The architecture in [38] aims to minimize the transmission delay and increase the computing capability between UAVs and mobile users. TI et al. [39] designed UAV based Fog-Cloud-Computing (FCC) to minimize the computation and power consumption of all users, which can jointly optimize the computation offloading, user-cloud/cloudlet association, transition power allocation, and path planning of mobile users. The UAV acts as a small distributed cloud and the local BS as micro cloud server; both users and UAV are movable.

When the terrestrial network infrastructure encounters a natural disaster such as earthquake, volcano, landslide and avalanche, UAVs can act as a network life saver, especially for emergency situations. One of the popular communication technologies is PSC, which plays a critical role in saving lives, property, and national infrastructure during natural or man-made emergency [40]. This technology is developed for delivering critical real-time streams (video, voice) using predefined spectrums. The UAV base station (UAVBS) or ABS, with LTE-advanced capabilities, can be utilized for emergency restoration and temporary expansion of public safety for disaster recovery [41]. ZHAO et al. [42] proposed a UAV-assisted emergency network to replace the destroyed base station by establishing multi-hop D2D users in different cells and relay the signal for emergency vehicular communication. And it is a promising method for establishing emergency networks. The authors of [43] studied how to replace destroyed base stations by UAV base stations after creating multi-hop D2D communications. They also designed a UAV transceiver for managing UAV uplink and downlink, extending the wireless coverage and guaranteeing the QoS of UAV communications for IoT in disasters.

3 Machine Learning: Overview

Machine learning is an application of artificial intelligence (AI), which provides systems with the ability to automatically learn and improve themselves from experience without being explicitly programmed. It is essentially based on the premise that machines should be furnished with AI that enables them to learn from previous computations and adapt to their environment through experience [32], [44]. Machine learning began to flourish in the 1990s. Before 1990s, logic-and knowledge-based schemes, such as inductive logic programming and expert systems dominated the AI scene relying on high-level human-readable symbolic representations of tasks and logic. Researchers in 2000s gradually renewed their interest on deep learning (DL) with the aid of advanced hardware-based computational capacity and the machine learning paradigm became popular at that time, supporting a wide range of services and applications in different areas [32], [44], [45].

3.1 Various Types of Machine Learning

Machine learning algorithms can be classified into three groups based on training data: supervised learning, unsupervised learning, and reinforcement learning (RL)

The supervised learning algorithm enables machines to be trained using labeled data. When dealing with labeled data, both the input data and its desired output data are known to the system. Supervised learning is commonly used in applications that have enough historical data. The algorithm is used to infer a function that maps the input data to the output label relying on the training of sample data-label pairs. Practically, considering a set of N sample data label pairs in the form of $\left\{\left(x_1,y_1\right),\left(x_2,y_2\right),\ldots,\left(x_N,y_N\right)\right\}$, where x_n is the n-th sample input data and y_n represents its label. Let $X=\left\{x_1,x_2,\ldots,x_N\right\}$ denotes the input data set and $Y=\left\{y_1,y_2,\ldots,y_N\right\}$ denotes the output label set. The sample pairs are independent and identically distributed (i. i. d.). The learning algorithms aim for seeking a function $g(x)$ that yields the highest value of the score function $f(x,y)$, hence we have $g(x)=argmaxy f(x,y)$. Supervised learning algorithms can be widely used in the context of classification, regression and prediction.

The unsupervised learning algorithm enables machines to be trained without labeled data. Unsupervised learning is typically about finding structure hidden in collections of unlabeled data. By analyzing N input data $X=\left\{x_1,x_2,\ldots,x_N\right\}$, a pair of popular methods have been conceived for revealing the underlying unknown features of N input data, namely density estimation and feature extraction.

RL enables machines to learn what to and how to map situations to actions so as to maximize a numerical reward signal. It is different from the above two algorithms and is currently the most popular research topic in the field of machine learning. There are elements which are necessary for reinforcement learning such as agent, state, action in a given environment. At each episode, the environment is in some state S and the agent selects a legitimate action A. The system responds at the next episode by moving into a new state S' with a certain probability influenced both by the specific action chosen and by the inherent transitions of the system. Meanwhile, the agent receives a corresponding reward $r(S,A)$ from the system, as time evolves. RL, an important branch of machine learning, is an effective tool and widely used Markov Decision Process (MDP) method [46]. In RL process, an agent can learn its optimal policy through interaction with its environment. Q-learning is the most effective method and widely used algorithm for RL. One of the most popular and widely used learning techniques is deep learning which allows the computer to build complex concepts out of simpler concepts. It is a set of algorithms and techniques that attempt to find important features of data and to model its high-level abstractions [40]. However, the learning process of RL takes a lot of time to reach optimal policy or generate best policy by exploring and generating knowledge of an environment, and this circumstance is not suitable and inapplicable for complex large problems. An artificial neural network (ANN) is a computational nonlinear model based on the neural structure of the brain, which is able to learn to perform tasks such as classification, prediction, decision-making, and visualization. The basic model of a neuron is mathematically expressed as follows:

$$Z_n\left(w_n,b_n,x_n\right) = f\left(b_n + \sum_{i=k}^{J} x_{nk} \, . w_{nk}\right), \tag{1}$$

where x_{nk} is an input signal from a given neuron n to neuron i, $x_n = \left[x_{n1}, x_{n2}, x_{n3}, ..., x_{nJ}\right]$ is a vector of the input signal of neuron n, w_{nk} is the corresponding input weight value, $w_n = \left[w_{n1}, w_{n2}, w_{n3}, ..., w_{nJ}\right]$ is a vector of input weight of neuron n, Z_n is the output signal of neuron n, b_n is the bias of neuron n, and $f(\ )$ is a nonlinear activation function. A bias value can shift the activation function, which is critical for successful learning. The activation function in a neural network will represent the rate of action potential ring in the cell of a neuron. An ANN constructed using linear activation functions in (1) cannot reach a stable state after training, and this problem can be controlled by normalizing different activation functions such as sigmoid function, tanh function, and rectified linear unit (ReLU) function.

3.2 Deep Reinforcement Learning

Deep learning was recognized as the first among the top ten AI technology trends for 2018 [45] and is already the leading machine learning technique successfully used in many scientific fields such as image recognition, text recognition, speech recognition, audio and language processing, and robotics [32], [44], [45]. Deep learning models are based on an ANN. As we mentioned above, the application of RL is insufficient for the current complicated problems. The combination of RL and deep learning, known as deep reinforcement learning (DRL), can break the limitation of RL in different areas. The DRL takes the advantage of deep neural networks (DDN) to train the learning process, improving the learning efficiency and performance of RL algorithms.

Q-learning is one of the most common used RL algorithms. It is an attempt to learn the value $Q(s,a)$ of a specific action given to the agent in a particular state. Considering a table where the number of rows represents the number of states, the RL agent interacts with the environment to learn the Q-values, based on which the agent takes an action. The Q-value is defined as the discounted accumulative reward starting at a tuple of a state and an action. Once the Q-values are learned after a maximum episode, the agent can make a quick decision under the current state by taking the action with the largest Q-value and the number of columns represents the number of actions which is called a Q-table [45], [47]. A large amount of state and action space in the environment makes the Q-table unmanageable. In current real-world examples like cellular edge computing, the state space is infinitely large. In order to eliminate the shortcoming of Q-learning, a neural network is used to predict the Q-values. One popular DRL algorithm is deep Q-network (DQN), which uses DNN to approximate the values. DQN is much more capable of generalization compared to the Q-network. DQN inherits and promotes advantages of both reinforcement and deep learning techniques, and thus it has a wide range of applications in practice such as game development, transportation, and robotics [44], [45], [47]. The study of DQNs has let too many improvements; new architectures have been designed for better performance and stability, including double DQN (DDQN), dueling DQN, and another asynchronous DRL algorithms studied on this articles [47]–[49].

4 System Architecture of UAV Enabled F-RAN

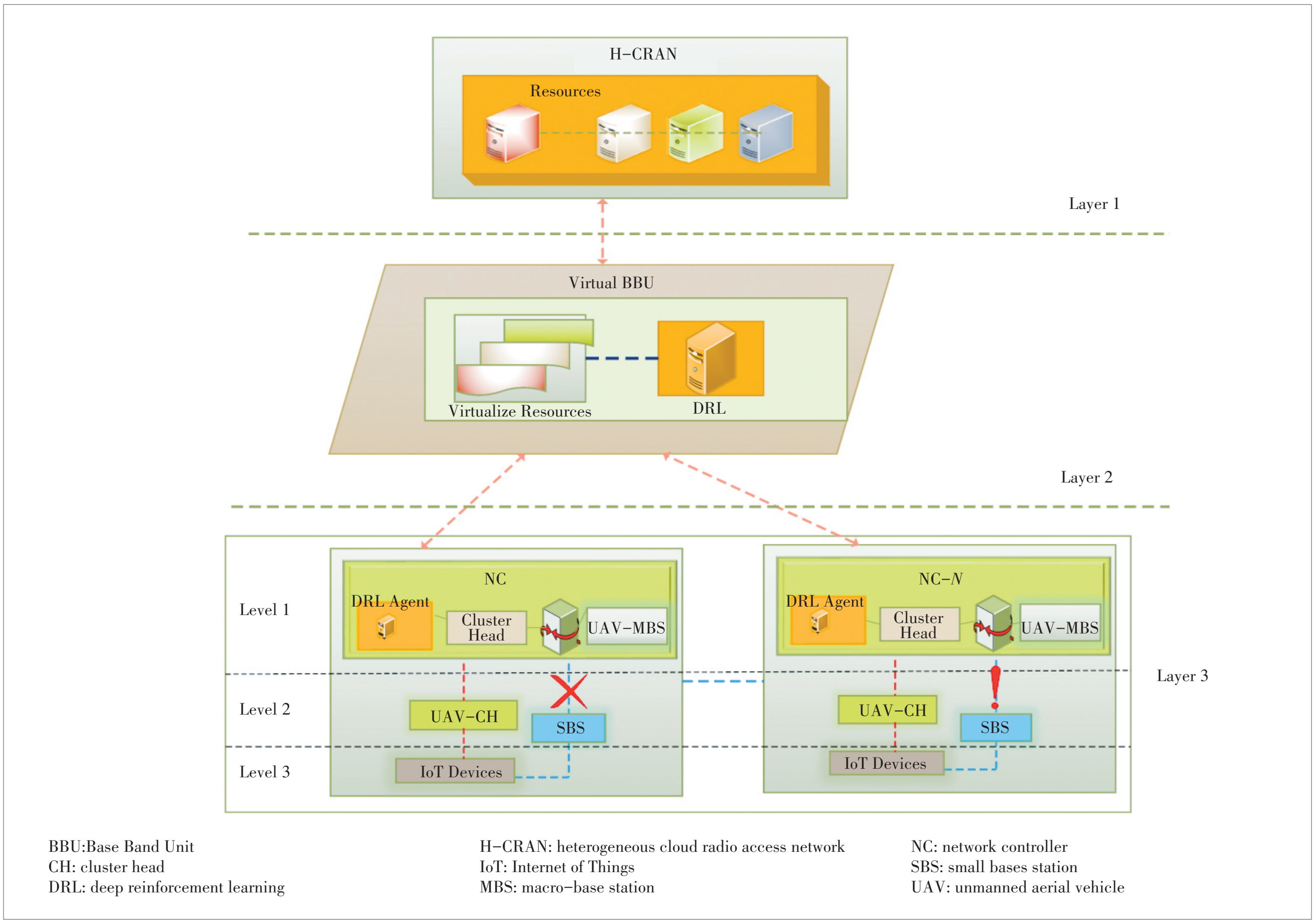
In the future, the ABS infrastructure will play a great role in 5G and beyond 5G communications. The ML algorithms applied in the current and future cellular technologies and aerial networks will be used to manage the dynamic network environment. **Figs. 1** and **2** depict the integration of UAV networks and terrestrial networks, where resources from cloud networks are accessed through the virtualized base band unit (VBBU). In VBBU, the network resources which are used for both aerial and terrestrial network infrastructures are virtualized in intelligent manner. The resources are allocated in the infrastructures, depending on the network demands. We categorize the architecture into three layers.

4.1 Layer 1

The H-CRAN that has cloud computing resources is delivered by server-based applications through digital networks or the public Internet itself. The resources which are available on cloud are far from edge IoT devices. Due to this, the edge IoT devices need localized computational nodes and resources to achieve features of 5G and B5G such as ultra-reliability, low-latency and massive (ubiquitous) connectivity.

4.2 Layer 2

The virtual BBU pool is located at the data center and multiple BBU nodes dynamically allocate resources to different network operators. The resources are allocated to aerial networks and terrestrial networks based on current network demands.



▲**Figure 1. UAV enabled fog radio access network (F-RAN) system architecture.**

On this layer, the resources are virtualized into N network slices which are found on cloud. The network virtualization allows network resources to be sliced and granted to multiple tenants. We assume the DRL is in decentralized manner and the fog-edge network can make decisions independently based on the local learning environment and inputs. The resulting decision will then be sent to the central controller.

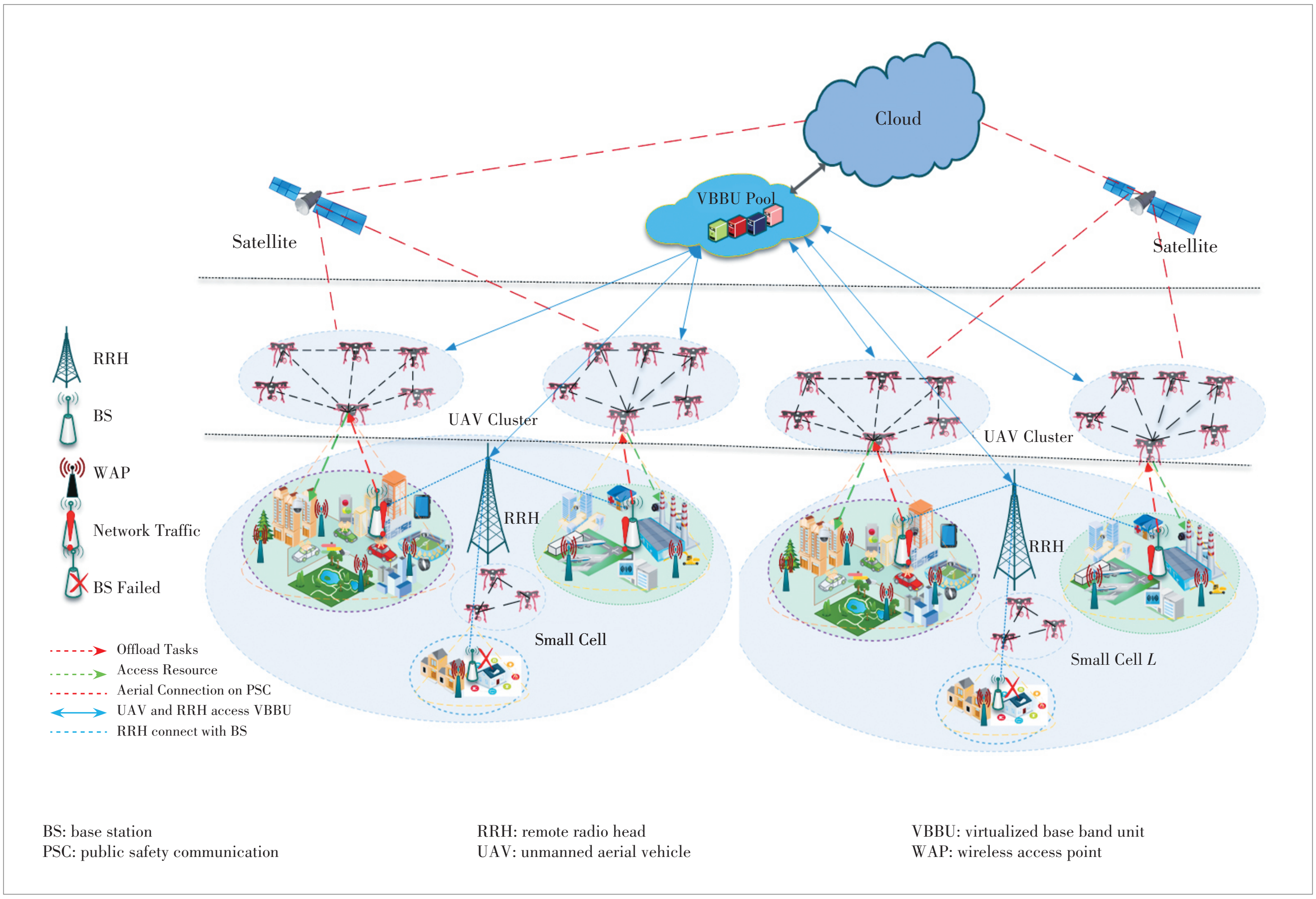
4.3 Layer 3

The main network operations such as DRL, resource management and computation offloading are performed at this layer. It has three levels which are network controller, UAV-small bases stations (SBS) and Edge IoT devices.

1) **Level 1**: The network controller (NC) is a central controller of the two network infrastructures and a communication platform where the aerial networks assist the terrestrial networks and DRL makes an intelligent coordination depending on network traffic, emergency and resource scarcity. A macro base station (MBS) with MEC server is used to manage resources which are allocated by the VBBU, allocate these resources to different network operators, and make a decision about the network condition for using DRL approach. To satisfy QoS and QoE of heterogeneous connected edge devices in each slice, the network will be assisted by UAV network in intelligent manners. Under MBS there are a number of SBSs with local servers in each small cell which are used to connect ultra-dense heterogeneous devices.

2) **Level 2:** UAV and SBS at this level are used to assist the communication in a given small cell mainly when the network is congested at specific time and in emergency situations; UAV acts as a flying base station to replace the destroyed BS and perform computational tasks and recharge of edge IoT devices. At this time the edge IoT devices are mainly wireless sensors, wearable devices and surveillance cameras, which offload the collected data into UAV for further analysis and decision making. Therefore, we consider UAV enabled F-RAN in which the UAV is considered as a flying remote radio head (RRH) or base station with computation capability to assist the edge IoT device. The UAV is part of cellular network; it recharges IoT sensor batteries and also sends collected data to MBSs.

3) **Level 3:** Edge IoT devices at this level are ultra-dense heterogeneous devices (mIoT devices), which are connected



▲**Figure 2. UAV enabled fog radio access network (F-RAN) and edge computing system model in public safety communication (PSC).**

with each other and SBSs. These devices share common resources, exchange information with the nearest devices, and have different interests. The MEC server may be crowed or even damaged when the devices request resources and need to offload their own tasks at the same time. The layer three has more network traffic than other layers and the cooperation of aerial network with terrestrial network is needed. The UAV assists the edge IoT network when either the network coverage is far from base station or some natural disaster has affected the network.

5 Machine Learning Algorithms in Edge Computing and UAV

In the current edge technology era, there is the sprite of direct communication between devices which are connected with the network infrastructures without travelling to base stations or core networks. D2D communication system is one of the most common networks and has been widely used in recent years; it is a milestone on the road towards self-organization and peer-to-peer (P2P) collaboration. Currently most of edge IoT devices need computing latency-sensitive support, which is not tolerable at the cloud level. In 2012, a group of researchers from Cisco proposed a new paradigm known as fog computing. Fog computing and edge computing appear similar since they both involve bringing intelligence and processing closer to UE. Most of the edge IoT devices have shortage of computational capacity and limitation of battery life. Due to this limitation, the edge IoT devices may fail to perform different operations properly. However, using the emerging MEC paradigm, the edge device can offload computation intensive tasks to the MEC server in different ways. The study of computation offloading and resource allocation in MEC and fog computing is complicated system analysis because of mobility patterns, radio access interfaces, strong couplings among mobile users with heterogeneities in application demands, QoS provisioning, and wireless resources. A machine learning approach specially using RL is a promising candidate to manage huge state space and optimization variables, especially by using different types of ANN.

DRL is an emerging tool for sophisticated problems in communication and networking in IoT, MEC, HetNet, and UAV networks. The network unities such as IoT devices, mobile users, and UAVs need to make local and autonomous decisions, like spectrum access, data rate selection, transmit power control, computation offloading decision, and base station association, to achieve the goals of different networks including throughput maximization, delay minimization, energy consumption minimization, and UAV deployment. The main problem is an uncertain and stochastic environment but the MDP model can solve the problem using dynamic programming, value iteration and RL [45]. LUONG et al. [31] studied the role of DRL in communication and networking. DRL minimizes the complexity of optimization and solves the problem in different perspectives. DRL allows network entities to learn and build knowledge about the communication and networking environment. By using DRL algorithms, mobile users can learn optimal policies for base station selection, channel selection, handover decision, caching and offloading decisions, UAV deployment, path planning, and trajectory optimization without knowing channel model and mobility pattern. In [31], different topics of the research works related to DRL were shown in percentages, for example, as the research in space communication is 13%, Ad-hoc 19%, cellular network 31%, IoT network 9%, and others 31%; the related issues to be solved were also presented in percentages, for example, the issues of wireless capacity is 19%, computation offloading 13%, rate control 8%, network access 13%, data collection 9%, resource scheduling 9%, connectivity preservation 8%, and network security 12%. Although there are a number of works on machine learning approaches for wireless communication networks [29], [31], [34], there is no research focus on machine learning based UAV enabled F-RAN infrastructures yet.

5.1 Machine Learning Based Computation Offloading in MEC

Edge IoT devices such as sensors and wearable devices have a limited computational capacity, short life time of battery, and storage. Due to this limitation, the IoT devices do not support advanced applications such as face recognition and online gaming (VR/AR). To tackle the problems in edge IoT devices and also in the network, an offloading mechanism is used to offload computational tasks and data to the nearest computational nodes (MEC server, UAV, or local servers). The offloading of data and computation tasks of the IoT devices can minimize the processing delay and energy consumption, and may enhance security. Under this circumstance, there are some critical challenges to computation offloading, such as choosing a computational node from multiple computational nodes and determining the offloading rate. Selection of an overloaded computational node also affects the computation time and energy consumption of IoT devices. The previous works on computation offloading and resources allocation used heuristic or iteration algorithms, but they have high complexity. Alternatively, machine learning is a promising tool used for solving the complex problem of computation offloading and resource allocation.

Recently, machine learning algorithms have been applied into fog edge computing to minimize the optimization problems. The authors of [50] proposed SDN NFV based DQN framework for caching and computation offloading to achieve energy efficiency in the network. The authors of [51] proposed a deep learning-based offloading framework to minimize the offloading cost for MEC networks. A deep supervised learning was also modeled to obtain the optimal offloading policy for mobile users. The authors of [52] tried to solve the resource allocation problems by joint optimization of caching, networking and computation for video content compressing and encoding, using feedforward neural network (FNN) based DQN. DDQN and dueling DQN approaches were proposed to improve the stability and performance of the DQN algorithm [53]. A DQN framework was also proposed for smart city applications, which is a dynamic orchestration of caching, bandwidth and computation to achieve QoS for different services [54].

The authors of [55] proposed offloading cellular traffic for WLAN by adopting the DQL algorithm and MDP model to minimize energy consumption and mobile user cost. The MEC server has a limitation of resources to allocate for all edge devices; due to this, the MEC server also minimize cost and energy. In a vehicular network, there is a huge action space and high complexity due to the vehicles' mobility and service delay. In [56], a multi-time scale DQN framework is proposed to minimize the system cost through jointly designing caching, communication and computing. The authors of [57] proposed DQN based joint optimization for computation offloading and resource allocation in MEC-enabled cellular networks. And the cost of delay and power consumption is accordingly minimized for all mobile users. In cellular networks, a DQL based optimal offloading policy was proposed to minimize the mobile users' cost and energy consumption [58]. In [59], a virtualized computation offloading framework using DRL was designed and a DDQN based DQL algorithm was proposed for an agent to learn the optimal offloading policy without prior knowledge of the network environment in a dynamic manner. This work also focused on the utility function by decomposing Q-function and combining with DDQN; a novel online SARSA-based DRL algorithm was proposed [59]. Besides, the computation offloading of multiple MEC servers have been considered [60] – [62]. The authors of [60] designed Q-learning and fast DQL offloading scheme to achieve optimal policy for IoT devices and energy harvesting capacity.

In [61], a two-layered DQL algorithm for offloading to maximize the utilization of cloud resources was studied; the first layer uses a convolutional neural network (CNN) -based DQL framework to estimate an optimal cluster for each computational task and the second layer uses Q-learning to determine the optimal serving physical machine in cluster. The authors of [62] proposed distributed deep learning-based offloading (DDLO) for multi-computing servers, users and tasks in MEC networks to minimize wireless device (WD) energy consumption by offloading WD tasks to the MEC server or cloud and allocating bandwidth. **Table 2** shows different machine learning algorithms in vehicular networks and cellular networks.

5.2 Machine Learning Based UAV Connected Cellular Networks

The application of machine learning for UAV is known as the drone system. Over the past years, many studies were conducted on either the integration of UAV networks with terrestrial networks or UAV networks in different application streams such as energy efficiency, computation offloading, resource allocation, and network coverage extension. However, Most of the previous works solved the existed problems using heuristic algorithm. The current research is focusing more on using machine learning algorithms to solve the aerial and terrestrial network integration for UAV assisted cellular networks, IoT, BSs and others to achieve a specific goal in the network. The cellular connected UAV will be a future hot research topic because it can integrate with future cellular networks and machine learning approaches to create a new intelligent aerial mobile user.

Many studies have been conducted on machine learning algorithms used in UAV or cellular connected UAV networks for optimizing UAV deployment, path planning, and trajectory as well as improving energy efficiency, UAV coverage, throughput, and resource allocation. GHANAVI et al. [63] proposed the optimal 3D UAV deployment to implement UAV-BSs which use RL to assist or serve the terrestrial network of mobility equipment for keeping the reliability of connection and increasing the QoS of users. The authors of [64] proposed an efficient 3D ABS positioning solution, in which DQN with DRL is used to assist the terrestrial BS in a small cell where the BS is overloaded and none of LoS exists for maximizing the spectra efficiency of the system. In [65], proposed a novel framework was proposed to deploy ABSs to assist overloaded or congested base stations in small cells. Researchers also adopted the machine learning approach to tackle the problem of predicting the traffic demand of each base station through previous histories, based on which ABSs are deployed for serving users in small cells and applying contract theory to jointly maximize the individual utility of each BS and UAV. In [66], an ANN based opportunistic computation offloading framework was proposed, the clustered UAV network assists a vehicular traffic network and the ground controller predicts the response time of each clustered UAV to offload intensive tasks. A clustered UAV network can compute intensive tasks by itself or borrow the resources from another cluster UAV network [66]. The authors of [67] studied the model free RL algorithm using Q-learning to optimize the trajectory of an UAV acting as a flying BS that serves multiple terrestrial network users. And the UAV also acts as an autonomous agent in the environment, learning the trajectory for maximize the sum rate of transmission during UAV flying time from one location to another location. CUI et al. [68] studied a multi-agent RL using Q-learning and stochastic game theory model for dynamic resource allocation in multi-UAV connected multi-users. Each UAV acts as an agent to make a decision independently for maximizing long-term rewards of each agent to provide reliable communications. Users, power levels and sub-channel selection strategies were also jointly studied in [68]. For cellular connected UAVs in beyond 5G system, a DRL algorithm was proposed based on the echo state network (ESN) for an interference aware path planning and management [69]. Each UAV acts as an agent that uses deep ESN to learn optimal path, transmission power level and cell association in each location of path and minimize sequence of time-dependent utility function. Authors of [69] also studied energy efficiency, the control of UAVs, and the fair covering of the active areas where the users are available and the UAVs are required to act as base stations by the DRL algorithm. In this work, the fairness index algorithm was applied to control UAV network coverage to minimize UAV energy consumption and improve UE QoS.

6 Challenges and Future Research Directions

According to the recent studies of various issues for future

▼Table 2. Machine learning algorithms for computation offloading and resource allocation in vehicular networks and cellular networks

Paper	Network	Agent	Model	Learning Algorithm
[50]	CRN	Base station	MDP	DQN using FNN
[52]	Vehicular Network	Service Provider	MDP	DQN using FNN
[53]	Vehicular Network	Service Provider	MDP	DQN using CNN
[55]	Cellular System	Base Station	MDP	DQN using CNN
[56]	Cellular System	Mobile User	MDP	DQN using CNN
[57]	Cellular System	Base Station	MDP	DQN using FNN
[58]	Cellular System	Mobile User	MDP	DQN using FNN
[59]	Cellular System	Mobile User	MDP	DDQN, SARSA
[60]	Cellular System	Mobile User	Game theory	DQN using CNN, Q-learning

CNN: convolutional neural network
CRN: cognitive radio network
DDQN: double deep Q-network
DQN: deep Q-network
FNN: feedforward neural network
MDP: Markov Decision Process
SARSA: state action reward state action

generation network infrastructures, we outline some challenges and future research directions for the integration of aerial networks and terrestrial networks with machine learning approaches in F-RAN, NFV and MEC paradigms.

6.1 Challenges

1) Machine learning used in virtualized UAV enabled F-RAN: RL (commonly DQN, Q-Learning and others) in virtualized MEC system has been used to tackle many issues at different layers of cellular networks. Deploying the machine learning algorithms at different layers of virtualized H-CRAN of UAV-enabled F-RAN will create the intelligence of the future network infrastructure of 5G and beyond. However, in this scenario there are a number of network infrastructure and concepts. Handling this multi-paradigm concept is complex in the current 5G technology and future 6G network.

2) Multi-agent in multi-layer UAV enabled F-RAN: Most of the current studies of cellular mobile networks or MEC system and UAV network focus on efficient resource allocation, energy efficiency, computation offloading, and caching to minimize delay and energy consumption or maximize revenue. The machine learning (commonly RL) algorithms have been used to tackle these issues, but most of them use a single agent at the base station or service providers. The recent years have witnessed the rapid evolution of network infrastructure and technologies from one generation to another generation every ten years. In the era of 5G, ultra-dense heterogeneous networks, which consist of different layers of IoT or fog network that supports ultra-low-latency (ULL) devices, are connected to each other at a given time step. In the future, beyond 5G or 6G (5G+AI) will support intelligent Personal Edge (IPE), genome database, autonomous health, sensors to AI fusion block-chain, etc. [70]-[72]. To perform complex multi-dimensional tasks in these networks, a multi-agent decentralized DRL approach needs to be adopted. Adopting this concept in the UAV-enabled F-RAN multi-agent at each layer is somehow complex and needs clear framework modeling.

3) Determination of the state of network traffic in different small cells: In 5G and beyond 5G era, there is ultra-dense heterogeneous network with massive IoT devices and smart mobile users which generate a huge amount of traffic in different circumstances. These ultra-dense devices will be assisted by UAV- cluster networks to satisfy the QoS and QoE rather than terrestrial base stations. In the UAV connected cellular network at lower layers such as fog or edge computing level, a single UAV or multi-UAVs are deployed and heuristic algorithms are used to identify network traffic in small cells, depending on the UAV capacity and coverage area. However, such application of machine learning in the dynamic network is unpredictable, has a large and continuous state space for making the determination of the network traffic state in different cells, and faces complex deployment of UAV-clusters.

4) Handover for transmitting data and task of mIoT devices for emergency situations: One of the attractive and promising paradigms of the UAV connected cellular network is acting as a flying base station to assist the emergency service. In this situation, the mIoT devices would send computational tasks and huge amount of request data traffic to the local base station at a specific time step. However, after the occurrence of a natural disaster, a good and intelligent handover framework is needed to manage the handovers in a terrestrial network environment in a disaster area. The application of machine learning algorithms in the handover process is much suitable.

6.2 Future Research Directions

1) Distributed machine learning based virtualized UAV enabled F-RAN: One of the popular machine learning algorithm frameworks in wireless communication and network is RL with deep neuron network, which requires large amount of training. Most of the time the large DNN is implemented at the central network controller which has sufficient resources such as computational capacity and is capable of training a large continuous state space and action space in the dynamic network environment. The central controller minimizes the burden of aerial mobile users and IoT devices by considering the limitation of capacities and capabilities. The main functionalities of UAV networks and terrestrial or cellular networks can be integrated with the central network controller. The virtualized DRL framework for UAV enabled F-RAN or UAV connected cellular system is an open issue. The network traffic exchanges from one layer to another and from aerial mobile users to terrestrial mobile users (mIoT devices) are efficient.

2) Dynamic deployment of multi-UAV cluster in F-RAN: In UAV networks, one of the open issues is UAV deployment in optimal 3D placement for different dynamic terrestrial network infrastructure. A number of previous works focused on UAV deployment with optimization of trajectory, path planning, and maximizing energy efficiency. Due to the dynamical network infrastructure in 5G and beyond 5G (6G), such as the rapid changes in coverage, the number of connected devices and network platforms, the DRL based approach for optimal 3D placement of UAV will be a necessity, with the integration of the cellular or IoT network. Under this consideration, there are other issues such as resource management (aerial mobile users and terrestrial network devices), optimal computation offloading, network coverage area, minimizing energy consumption of network, and cell association to maximize flight time.

3) Machine learning based resource management in UAV-Enabled F-RAN: A number of studies have been conducted on resource management at different layers in cellular networks, vehicular networks, and UAV networks to solve complex problems such as optimization, maximizing energy efficiency, resource allocation for UAV and bandwidth management. These studies aim to maximize the revenue or minimize the cost of delay and energy in the system. Other works that used heuristic algorithms to tackle the complex problems in cellular net-

works, vehicular networks, and fog and edge computing are now adopting machine learning, commonly RL (DQN, Q-learning, DDQN, DDGP, Actor- critics) for resource management and computation offloading. However, in the mixed network infrastructures such as UAV-enabled F-RAN, need to design a machine learning based joint resource management and computation offloading framework.

4) Machine learning for dynamic deployment of ABS in emergency (PSC): UAV plays a potential role in the future promising paradigm for emergency situations known as PSC. The current communication era heavily relies on the backbone networks. For the failure of base stations due to natural disaster or malevolent attacks, PSC is able to use machine learning to deploy a group of multi-UAVs in ultra-dense HetNet architecture as ABSs that can dynamically replace the destroyed or over-headed base stations in the terrestrial network. The UAVs are used to support the reliable connection for edge IoT devices, extend the network coverage, control the end user devices, etc. from the communication perspective. If a destroyed BS has the computational resource (local server), MEC server, and power source that cannot be accessed by edge IoT devices, the intelligent ABSs also replace the destroyed terrestrial BS to conduct computing task and allocate transmission power to satisfy the QoS and QoE of end users/IoT devices at the fog/edge level of RAN networks.

5) Machine learning based mobility control of multi-UAV connected cellular network/F-RAN: In a multi-UAV assisted cellular network/F-RAN, the UAV flies from one location to another location within the given time frame. At the time of UAV's flying over the terrestrial network, mobile users/IoT devices will wait for long time to get access to the UAV terminal. Due to this, the QoS and QoE of the network could be degraded. To tackle this issue, an intelligent machine learning based model is designed for multi-UAV mobility management, where the agents learn by themselves to adjust the mobility in the predicted location in the terrestrial network infrastructure. Besides, the model also considers the terrestrial network connected devices such as mobile users, vehicle, and other mobility environments. In this scenario, the management of resources (computational, bandwidth, and energy) is also considered in the mixed network infrastructures.

7 Conclusions

This paper presents a short review of the machine learning used to solve complex problems in modern network infrastructures and suggests the machine learning based multi UAV-enabled F-RAN. First, we introduce F-RAN and UAV for the current and future network technologies. Second, we discuss UAV in cellular networks and its replacement of base stations in terrestrial networks. Third, we review machine learning algorithms and RL and suggest the machine learning based UAV-enabled F-RAN framework architecture in H-CRAN network infrastructure for computation offloading and resource allocation. We also mention some previous works on edge computing and UAV using RL with DNN to solve different problems such as resource allocation, computation offloading and base station replacement in different networks. Finally, we outline the challenges and future research directions.

References

[1] ASSUNÇÃO M D, CALHEIROS R N, BIANCHI S, et al. Big Data Computing and Clouds: Trends and Future Directions [J]. Journal of Parallel and Distributed Computing, 2015, 79 – 80: 3 – 15. DOI: 10.1016/j.jpdc.2014.08.003

[2] ALHADDADIN F, LIU W, GUTIERREZ J A. A User Profile-Aware Policy-Based Management Framework for Greening the Cloud [C]//IEEE Fourth International Conference on Big Data and Cloud Computing, Sydney, Australia, 2014: 682 – 687. DOI: 10.1109/bdcloud.2014.116

[3] DASTJERDI A V, GUPTA H, CALHEIROS R N, et al. Fog Computing: Principles, Architectures, and Applications [M]//Internet of Thing Principles and Paradigms, Burlington, USA: Morgan Kaufmann, 2016. DOI: 10.1016/b978-0-12-805395-9.00004-6

[4] CHIANG M, ZHANG T. Fog and IoT: an Overview of Research Opportunities [J]. IEEE Internet of Things Journal, 2016, 3(6): 854 – 864. DOI: 10.1109/jiot.2016.2584538

[5] YU W, LIANG F, HE X, HATCHER W G, et al. A Survey on the Edge Computing for the Internet of Things [J]. Mobile Edge Computing, 2017, 6(88): 6900 – 6919. DOI: 10.1109/access.2017.2778504

[6] MAHMUD R, KOTAGIRI R, BUYYA R. Fog Computing: A Taxonomy, Survey and Future Directions [M]//Internet of Things. Singapore, Singapore: Springer, 2017: 103 – 130. DOI: 10.1007/978-981-10-5861-5_5

[7] GAO L X, LUAN T H, YU S, et al. FogRoute: DTN-Based Data Dissemination Model in Fog Computing [J]. IEEE Internet of Things Journal, 2016, 4(1):225 – 235. DOI: 10.1109/jiot.2016.2645559

[8] YI S, LI C, LI Q. A Survey of Fog Computing: Concepts, Applications and Issues [C]//Proc. 2015 Workshop on Mobile Big Data, New York, USA: Association for Computing Machinery, 2015: 37 – 42. DOI: 10.1145/2757384.2757397

[9] HU P F, DHELIM S, NING H S, et al. Survey on Fog Computing: Architecture, Key Technologies, Applications and Open Issues [J]. Journal of Network and Computer Applications, 2017, 98: 27 – 42. DOI: 10.1016/j.jnca.2017.09.002

[10] BACCARELLI E, NARANJO P G V, SCARPINITI M, et al. Fog of Everything: Energy-Efficient Networked Computing Architectures, Research Challenges, and a Case Study [J]. IEEE Access, 2017, 5: 9882 – 9910. DOI: 10.1109/access.2017.2702013

[11] VARSHNEY P, SIMMHAN Y. Demystifying Fog Computing: Characterizing Architectures, Applications and Abstractions [C]//IEEE 1st International Conference on Fog and Edge Computing (ICFEC), Madrid, Spain, 2017: 115 – 124. DOI: 10.1109/icfec.2017.20

[12] MOURADIAN C, NABOULSI D, YANGUI S M, et al. A Comprehensive Survey on Fog Computing: State-of-the-Art and Research Challenges [J]. IEEE Communications Surveys & Tutorials, 2018, 20(1): 416 – 464. DOI: 10.1109/comst.2017.2771153

[13] PERERA C, QIN Y R, ESTRELLA J C, et al. Fog Computing for Sustainable Smart Cities [J]. ACM Computing Surveys (CSUR), 2017, 50(3): 1 – 43. DOI: 10.1145/3057266

[14] YOUSEFPOUR A, FUNG C, NGUYEN T, et al. All One Needs to Know about Fog Computing and Related Edge Computing Paradigms: A Complete Survey [J]. Journal of Systems Architecture, 2019, 98: 289 – 330. DOI: 10.1016/j.sysarc.2019.02.009

[15] FRENCH A M, SHIM J P. The Digital Revolution: Internet of Things, 5G and Beyond [J]. Communications of the Association for Information Systems (CAIS), 206, 38(40): 40. DOI: 10.17705/1CAIS.03840

[16] MACH P, BECVAR Z. Mobile Edge Computing: A Survey on Architecture and Computation Offloading [J]. IEEE Communications Surveys & Tutorials, 2017, 19(3): 1628–1656. DOI: 10.1109/ comst.2017.2682318
[17] BUYYA R, SRIRAMA S N. Fog and Edge Computing: Principles and Paradigms [M]. Hoboken, USA: John Wiley & Sons, 2019. DOI: 10.1002/ 9781119525080
[18] ALNOMAN A, SHARMA S K, EJAZ W, et al. Emerging Edge Computing Technologies for Distributed IoT Systems [J]. IEEE Network, 2019, 33(6): 140–147. DOI: 10.1109/mnet.2019.1800543
[19] HUSAIN S, KUNZ A, PRASAD A, et al. Mobile Edge Computing with Network Resource Slicing for Internet-of-Things [C]//IEEE 4th World Forum on Internet of Things (WF-IoT). Singapore, Singapore, 2018: 1–6. DOI: 10.1109/wf-iot.2018.8355232
[20] BESHLEY H, BESHLEY M, MAKSYMYUK T, et al. Method of Centralized Resource Allocation in Virtualized Small Cells Network with IoT Overlay [C]//14th International Conference on Advanced Trends in Radioelecrtronics, Telecommunications and Computer Engineering (TCSET). Lviv-Slavske, Ukraine, 2018: 1147–1151. DOI: 10.1109/tcset.2018.8336397
[21] GUO H Z, LIU J J, ZHANG J, et al. Mobile-Edge Computation Offloading for Ultradense IoT Networks [J]. IEEE Internet of Things Journal, 2018, 5(6): 4977–4988. DOI: 10.1109/jiot.2018.2838584
[22] PARK Y, KIM S. Game-Based Data Offloading Scheme for IoT System Traffic Congestion Problems [J]. EURASIP Journal on Wireless Communications and Networking, 2015, 2015(1): 192. DOI: 10.1186/s13638-015-0418-y
[23] ZHANG H J, QIU Y, CHU X L, et al. Fog Radio Access Networks: Mobility Management, Interference Mitigation, and Resource Optimization [J]. IEEE Wireless Communications, 2017, 24(6): 120–127. DOI: 10.1109/ mwc.2017.1700007
[24] TRAN T X, HAJISAMI A, PANDEY P, et al. Collaborative Mobile Edge Computing in 5G Networks: New Paradigms, Scenarios, and Challenges [J]. IEEE Communications Magazine, 2017, 55(4): 54–61. DOI: 10.1109/ mcom.2017.1600863
[25] ZENG Y, WU Q Q, ZHANG R. Accessing from the Sky: A Tutorial on UAV Communications for 5G and Beyond [J]. Proceedings of the IEEE, 2019, 107 (12): 2327–2375. DOI:10.1109/jproc.2019.2952892
[26] HOSSEIN MOTLAGH N, TALEB T, AROUK O. Low-Altitude Unmanned Aerial Vehicles-Based Internet of Things Services: Comprehensive Survey and Future Perspectives [J]. IEEE Internet of Things Journal, 2016, 3(6): 899–922. DOI: 10.1109/jiot.2016.2612119
[27] MOZAFFARI M, SAAD W, BENNIS M, et al. A Tutorial on UAVs for Wireless Networks: Applications, Challenges, and Open Problems [J]. IEEE Communications Surveys & Tutorials, 2019, 21(3): 2334–2360. DOI: 10.1109/ comst.2019.2902862
[28] FENG W, WANG J C, CHEN Y F, et al. UAV-Aided MIMO Communications for 5G Internet of Things [J]. IEEE Internet of Things Journal, 2019, 6(2): 1731–1740. DOI: 10.1109/jiot.2018.2874531
[29] WANG J J, JIANG C X, ZHANG H J, et al. Thirty Years of Machine Learning: The Road to Pareto-Optimal Next-Generation Wireless Networks [EB/OL]. (2019-01-24). https://arxiv.org/abs/1902.01946
[30] BOUTABA R, SALAHUDDIN M A, LIMAM N, et al. A Comprehensive Survey on Machine Learning for Networking: Evolution, Applications and Research Opportunities [J]. Journal of Internet Services and Applications, 2018, 9(1): 16. DOI: 10.1186/s13174-018-0087-2
[31] LUONG N C, HOANG D T, GONG S M, et al. Applications of Deep Reinforcement Learning in Communications and Networking: A Survey [J]. IEEE Communications Surveys & Tutorials, 2019, 21(4): 3133–3174. DOI: 10.1109/ comst.2019.2916583
[32] CHEN M Z, CHALLITA U, SAAD W, et al. Artificial Neural Networks-Based Machine Learning for Wireless Networks: A Tutorial [J]. IEEE Communications Surveys & Tutorials, 2019, 21(4): 3039–3071. DOI: 10.1109/ comst.2019.2926625
[33] KLAINE P V, IMRAN M A, ONIRETI O, et al. A Survey of Machine Learning Techniques Applied to Self-Organizing Cellular Networks [J]. IEEE Communications Surveys & Tutorials, 2017, 19(4): 2392–2431. DOI: 10.1109/ comst.2017.2727878
[34] SUN Y H, PENG M G, ZHOU Y C, et al. Application of Machine Learning in Wireless Networks: Key Techniques and Open Issues [J]. IEEE Communications Surveys & Tutorials, 2019, 21(4): 3072–3108. DOI: 10.1109/ comst.2019.2924243
[35] FOTOUHI A, QIANG H R, DING M, et al. Survey on UAV Cellular Communications: Practical Aspects, Standardization Advancements, Regulation, and Security Challenges [J]. IEEE Communications Surveys & Tutorials, 2019, 21(4): 3417–3442. DOI: 10.1109/comst.2019.2906228
[36] ZHANG J, ZHOU L, TANG Q, et al. Stochastic Computation Offloading and Trajectory Scheduling for UAV-Assisted Mobile Edge Computing [J]. IEEE Internet of Things Journal, 2018, 6(2): 3688–3699. DOI: 10.1109/ jiot.2018.2890133
[37] CHENG F, ZHANG S, LI Z, et al. UAV Trajectory Optimization for Data Offloading at the Edge of Multiple Cells [J]. IEEE Transactions on Vehicular Technology, 2018, 67(7): 6732–6736. DOI: 10.1109/tvt.2018.2811942
[38] LIU B C, HUANG H W, GUO S, et al. Joint Computation Offloading and Routing Optimization for UAV-Edge-Cloud Computing Environments [C]//IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ ATC/CBDCom/IOP/SCI), Guangzhou, China, 2018: 1745–1752. DOI: 10.1109/ smartworld.2018.00295
[39] TI N T, LE L B. Joint Resource Allocation, OffloadingComputation, and Path Planning for UAV Based Hierarchical Fog-Cloud Mobile Systems[C]//IEEE Seventh International Conference on Communications and Electronics (ICCE), Hue, Vietnam, 2018: 373–378. DOI: 10.1109/cce.2018.8465572
[40] MERWADAY A, GUVENC I. UAV Assisted Heterogeneous Networks for Public Safety Communications[C]//IEEE Wireless Communications and Networking Conference Workshops (WCNCW), New Orleans, USA, 2015: 329–334. DOI: 10.1109/wcncw.2015.7122576
[41] HAINZL M. Integrating UAVs into Public Safety LTE Networks [R],Tech. Rep., 2014.
[42] ZHAO N, LU W D, SHENG M, et al. UAV-Assisted Emergency Networks in Disasters [J]. IEEE Wireless Communications, 2019, 26(1): 45–51. DOI: 10.1109/mwc.2018.1800160
[43] LIU X N, LI Z, ZHAO N, et al. Transceiver Design and Multi-Hop D2D for UAV IoT Coverage in Disasters [J]. IEEE Internet of Things Journal, 2018, 6 (2): 1803–1815. DOI: 10.1109/jiot.2018.2877504
[44] MOHRI M, ROSTAMIZADEH A, TALWALKAR A. Foundations of Machine Learning [M]. Cambridge, USA: MIT press, 2012
[45] SUTTON R S, BARTO A G. Reinforcement Learning: An Introduction [M]. Cambridge, USA: MIT press, 1998. DOI: 10.1109/tnn.1998.712192
[46] HOWARD R A. Dynamic Programming and Markov Processes [M]. New York, USA: Technology Press of the Massachusetts Institute of Technology, 1960. DOI: 10.2307/2312519
[47] MNIH V, KAVUKCUOGLU K, SILVER D, et al. Human-Level Control through Deep Reinforcement Learning [J]. Nature, 2015, 518(7540): 529–533. DOI: 10.1038/nature14236
[48] MNIH V, BADIA A P, MIRZA M, et al. Asynchronous Methods for Deep Reinforcement Learning [C]//International Conference on Machine Learning, New York, USA, 2016:1928–1937
[49] LILLICRAP T P, HUNT J J, PRITZEL A, et al. Continuous Control with Deep Reinforcement Learning [EB/OL]. (2015-09-09). https://arxiv. org/abs/ 1509.02971
[50] HE Y, ZHANG Z, ZHANG Y H. A Big Data Deep Reinforcement Learning Approach to Next Generation Green Wireless Networks [C]//IEEE Global Communications Conference, Singapore, Singapore, 2017: 1–6. DOI: 10.1109/glocom.2017.8254717
[51] YU S, WANG X, LANGAR R. Computation Offloading for Mobile Edge Computing: A Deep Learning Approach [C]//IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC). Montreal, Canada, 2017: 1–6. DOI: 10.1109/pimrc.2017.8292514
[52] HE Y, LIANG C C, ZHANG Z, et al. Resource Allocation in Software-Defined and Information-Centric Vehicular Networks with Mobile Edge Computing [C]// IEEE 86th Vehicular Technology Conference (VTC-Fall), Toronto, Canada, 2017: 1–5. DOI: 10.1109/vtcfall.2017.8288203
[53] HE Y, ZHAO N, YIN H X. Integrated Networking, Caching, and Computing for Connected Vehicles: A Deep Reinforcement Learning Approach [J]. IEEE Transactions on Vehicular Technology, 2018, 67(1): 44–55. DOI: 10.1109/

tvt.2017.2760281
[54] HE Y, YU F R, ZHAO N, et al. Software-Defined Networks with Mobile Edge Computing and Caching for Smart Cities: A Big Data Deep Reinforcement Learning Approach [J]. IEEE Communications Magazine, 2017, 55(12): 31 – 37. DOI: 10.1109/mcom.2017.1700246
[55] ZHANG C, LIU Z, GU B, et al. A Deep Reinforcement Learning Based Approach for Cost- and Energy-Aware Multi-Flow Mobile Data Offloading [J]. IEICE Transactions on Communications, 2018, E101. B(7: 1625 – 1634. DOI: 10.1587/transcom.2017cqp0014
[56] TAN L T, HU R Q. Mobility-Aware Edge Caching and Computing in Vehicle Networks: A Deep Reinforcement Learning [J]. IEEE Transactions on Vehicular Technology, 2018, 67(11): 10190 – 10203. DOI:10.1109/tvt.2018.2867191
[57] LI J, GAO H, LV T, et al. Deep Reinforcement Learning Based Computation Offloading and Resource Allocation for MEC [C]//2018 IEEE Wireless Communications and Networking Conference (WCNC). Barcelona, Spain, 2018: 1 – 6. DOI:10.1109/wcnc.2018.8377343
[58] CHEN X F, ZHANG H G, WU C, et al. Performance Optimization in Mobile-Edge Computing Via Deep Reinforcement Learning[C]//2018 IEEE 88th Vehicular Technology Conference (VTC-Fall). Chicago, USA, 2018. DOI: 10.1109/vtcfall.2018.8690980
[59] CHEN X F, ZHANG H G, WU C, et al. Optimized Computation Offloading Performance in Virtual Edge Computing Systems Via Deep Reinforcement Learning [J]. IEEE Internet of Things Journal, 2019, 6(3): 4005 – 4018. DOI:10.1109/jiot.2018.2876279
[60] MIN M H, XIAO L, CHEN Y, et al. Learning-Based Computation Offloading for IoT Devices with Energy Harvesting [J]. IEEE Transactions on Vehicular Technology, 2019, 68(2): 1930 – 1941. DOI: 10.1109/tvt.2018.2890685
[61] QUAN L, WANG Z L, REN F J. A Novel Two-Layered Reinforcement Learning for Task Offloading with Tradeoff Between Physical Machine Utilization Rate and Delay [J]. Future Internet, 2018, 10(7): 60. DOI: 10.3390/fi10070060
[62] HUANG L, FENG X, ZHANG L X, et al. Multi-Server Multi-User Multi-Task Computation Offloading for Mobile Edge Computing Networks [J]. Sensors, 2019, 19(6): 1446. DOI: 10.3390/s19061446
[63] GHANAVI R, KALANTARI E, SABBAGHIAN M, et al. Efficient 3D Aerial Base Station Placement Considering Users Mobility by Reinforcement Learning [C]//IEEE Wireless Communications and Networking Conference (WCNC), Barcelona, Spain, 2018: 1 – 6. DOI: 10.1109/wcnc.2018.8377340
[64] WU J, YU P, FENG L, et al. 3D Aerial Base Station Position Planning Based on Deep Q-network for Capacity Enhancement [C]//FIP/IEEE Symposium on Integrated Network and Service Management (IM), Arlington, USA, 2019: 482 – 487
[65] ZHANG Q Q, SAAD W, BENNIS M, et al. Predictive Deployment of UAV Base Stations in Wireless Networks: Machine Learning Meets Contract Theory [EB/OL]. (2018-11-03). https://arxiv.org/abs/1811.01149
[66] VALENTINO R, JUNG W S, KO Y B. A Design and Simulation of the Opportunistic Computation Offloading with Learning-Based-Prediction for Unmanned Aerial Vehicle (UAV) Clustering Networks [J]. Sensors, 2018, 18(11): 3751. DOI: 10.3390/s18113751
[67] BAYERLEIN H, DE KERRET P, GESBERT D. Trajectory Optimization for Autonomous Flying Base Station via Reinforcement Learning [C]//IEEE 19th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC), Kalamata, Greece, 2018: 1 – 5. DOI:10.1109/spawc.2018.8445768
[68] CUI J J, LIU Y W, NALLANATHAN A. Multi-Agent Reinforcement Learning Based Resource Allocation for UAV Networks [J]. IEEE Transactions on Wireless Communications, 2019: 1. DOI:10.1109/twc.2019.2935201
[69] LIU C H, CHEN Z Y, TANG J, et al. Energy-Efficient UAV Control for Effective and Fair Communication Coverage: A Deep Reinforcement Learning Approach [J]. IEEE Journal on Selected Areas in Communications, 2018, 36(9): 2059 – 2070. DOI:10.1109/jsac.2018.2864373
[70] NAWAZ S J, SHARMA S K, WYNE S, et al. Quantum Machine Learning for 6G Communication Networks: State-of-the-Art and Vision for the Future [J]. IEEE Access, 2019, 7: 46317 – 46350. DOI: 10.1109/access.2019.2909490
[71] LETAIEF K B, CHEN W, SHI Y M, et al. The Roadmap to 6G: AI Empowered Wireless Networks [J]. IEEE Communications Magazine, 2019, 57(8): 84 – 90. DOI: 10.1109/mcom.2019.1900271
[72] O'BRIEN C. Why 6G Research Is Starting Before We Have 5G [EB/OL]. (2019-7-1). https://venturebeat.com/2019/03/21/6g-research-starting-before-5g

Biographies

Mohammed SEID received his B. Sc. and M. Sc. degrees in computer science from Ambo University, Ethiopia in 2010 and Addis Ababa University, Ethiopia in 2015, respectively. He is currently pursuing his Ph. D. degree in computer science and technology at University of Electronic Science and Technology of China. From 2010 to 2016 he worked in Dilla University, Ethiopia as a graduate assistant and lecturer. His interests include mobile edge computing, fog computing, UAV networks, IoT, and 5G wireless networks.

Stephen ANOKYE received his M. Eng. degree in computer science from Hunan University, China in 2009 and his B. Sc. in computer science from Kwame Nkrumah University of Science and Technology, Ghana in 2004. He's currently a Ph. D. candidate at the University of Electronic Science and Technology of China. Between 2010 and 2012, he worked as a lecturer in the department of computer science at Garden City University College in Ghana. Since 2012, he has become a lecturer at the Department of Computer Science and Engineering, the University of Mines and Technology, Ghana. His research interests are security in wireless sensor networks, mobile and cloud networks with AI, UAV networks, IoT, and 5G wireless networks.

SUN Guolin (guolin. sun@uestc. edu. cn) received his B. S., M. S. and Ph. D. degrees all in communication and information system from University of Electronic Science and Technology of China (UESTC) in 2000, 2003 and 2005 respectively. After his Ph. D. graduation in 2005, he has got eight-year industrial work experience in wireless research and development for LTE, Wi-Fi, and IoT, Cognitive radio, Localization and navigation. Before he joined the UESTC as an associate professor in August 2012, he worked in Huawei Technologies Sweden. Dr. SUN has filed over 30 patents and published over 40 scientific conference and journal papers. He acted as TPC member of several conferences. Currently, he serves as a vice-chair of the 5G oriented cognitive radio SIG of the IEEE Technical Committee on Cognitive Networks (TCCN). His general research interests are software defined networks, network function virtualization, and radio resource management.