Machine Learning for Resource Management in Cellular and IoT Networks: Potentials, Current Solutions, and Open Challenges

Fatima Hussain, *Senior Member, IEEE*, Syed Ali Hassan , *Senior Member, IEEE*, Rasheed Hussain , *Senior Member, IEEE*, and Ekram Hossain , *Fellow, IEEE*

Abstract—Internet-of-Things (IoT) refers to a massively heterogeneous network formed through smart devices connected to the Internet. In the wake of disruptive IoT with a huge amount and variety of data, Machine Learning (ML) and Deep Learning (DL) mechanisms will play a pivotal role to bring intelligence to the IoT networks. Among other aspects, ML and DL can play an essential role in addressing the challenges of resource management in large-scale IoT networks. In this article, we conduct a systematic and in-depth survey of the ML- and DL-based resource management mechanisms in cellular wireless and IoT networks. We start with the challenges of resource management in cellular IoT and low-power IoT networks, review the traditional resource management mechanisms for IoT networks, and motivate the use of ML and DL techniques for resource management in these networks. Then, we provide a comprehensive survey of the existing ML- and DL-based resource management techniques in wireless IoT networks and the techniques specifically designed for HetNets, MIMO and D2D communications, and NOMA networks. To this end, we also identify the future research directions in using ML and DL for resource allocation and management in IoT networks.

Index Terms—Internet-of-Things (IoT), wireless IoT, machine learning, deep learning, resource allocation, resource management, D2D, MIMO, HetNets, NOMA.

I. Introduction

BY 2020, it is expected to have 4 billion connected people, 25 billion embedded smart systems generating 50 trillion GB of data [1]. In the wake of such statistics, Internet of Things (IoT), specifically wireless IoT, has a great potential for the future smart world; however, the deployment of

Fatima Hussain is with the API Operations and Delivery, Technology and Operations, Royal Bank of Canada, Toronto, ON M5J 0B8, Canada (e-mail: fatima.hussain@rbc.com).

Syed Ali Hassan is with the School of Electrical Engineering and Computer Science, National University of Sciences and Technology, Islamabad 44000, Pakistan (e-mail: ali.hassan@seecs.edu.pk).

Rasheed Hussain is with the Networks and Blockchain Laboratory, Institute of Information Security and Cyber-Physical Systems, Innopolis University, 420500 Innopolis, Russia (e-mail: r.hussain@innopolis.ru).

Ekram Hossain is with the Department of Electrical and Computer Engineering, University of Manitoba, Winnipeg, MB R3T 5V6, Canada (e-mail: ekram.hossain@umanitoba.ca).

IoT on a massive scale also brings along lots of challenges. Some of these challenges include design of the networking and storage architecture for smart devices, efficient data communication protocols, proactive identification and protection of IoT devices and applications from malicious attacks, standardization of technologies, and devices and application interfaces. Additionally there are challenges such as device management, data management, diversity, and inter-operability [2], [3].

For the wireless IoT, since radio resources are inherently scarce, resource management becomes very challenging in practical networks, specially when the number of contending users or smart devices is very large. The overall performance of any wireless network depends on how efficiently and dynamically hyper-dimensional resources such as time slots, frequency bands and orthogonal codes are managed, and also how well the variable traffic loads and wireless channel fluctuations are incorporated into the network design in order to provide connectivity among users with diverse Quality-of-Service (QoS) requirements. Development of new IoT applications advocates for an improved spectral efficiency, high data rates and throughput as well as consideration of context and personalization of the IoT services, and consequently, the problem of resource management becomes crucial [4], [5].

A. Motivation of the Work

1) Resource Management Issues in IoT Networks:

- *Massive channel access:* When a massive number of devices access the wireless channel simultaneously, it overloads the channel (e.g., Physical Random Access CHannel- PRACH in LTE networks). Also, a massive deployment of IoT devices bring challenges related to network congestion, networking and storage architecture for smart devices, and efficient data communication protocols [2], [3]. Solutions for combating congestion challenges in access channel for IoT devices are mostly based on using different access probabilities and prioritization among different classes. Resource management for cellular IoT networks needs to consider load balancing and access management techniques to support massive capacity and connectivity while utilizing the network resources efficiently.

- *Power allocation and interference management:* In a dense (and large-scale) IoT network, intra- and inter-cell

interference problems become crucial, and therefore, sophisticated power allocation and interference management techniques are required. Many of the IoT applications can be of heterogeneous nature with potentially varying network topology, traffic as well as channel characteristics. Therefore, it is very challenging to choose transmit power dynamically in response to the changing physical channel and network conditions.

- *User association (or cell selection) and hand-off management:* In cellular IoT applications, the devices have to associate with a Base Station (BS) or a gateway node to connect to the Internet. Again, the association may not be necessarily symmetric in both uplink and downlink scenarios (i.e., an IoT device may associate to different BSs at the uplink and the downlink). This association or cell selection affects the resource allocation. Also, for the mobile IoT devices, resource management methods need to consider the possibilities of hand-off in different cells (e.g., macrocells, microcells).

- *Harmonious coexistence of Human-to-Human (H2H) and IoT traffic:* For harmonious coexistence between the existing and new IoT traffic, intelligent resource management (e.g., for channel and power allocation/interference management, user/device association) is crucial so that both the types of communications can co-exist with their requirements satisfied.

- *Coverage extension:* IoT devices are essentially low-power devices with limited computational and storage capabilities, in contrast to the traditional devices such as smart phones and tablets. This phenomenon reduces the coverage range. Therefore, to extend the coverage for IoT services, innovative resource management solutions are required. IoT devices can exploit Device-to-Device (D2D) or relay-based communications techniques for their coverage extension. Dynamic resource allocation methods will be required for such communication scenarios.

- *Energy management:* Energy-efficient communication is very important since most of the sensors and actuators in IoT are battery-powered with limited battery capacity and charging facilities. Design of energy-efficient resource allocation and communication protocol design is essential to provide satisfactory QoS and support in dynamic network environments.

- *Real-time processing and Ultra-Reliable and Low Latency Communication (URLLC):* Mission-critical IoT applications such as remote surgery, public safety, Intelligent Transport System (ITS) require low-latency communication and ultra-reliable connectivity. This calls for reliable and delay-aware resource allocation techniques.

- *Heterogeneous QoS:* Since IoT networks are comprised of heterogeneous devices with different capabilities and characteristics as well as communications requirements, they will require specialized and customized resources specific to their requirements.

2) Limitations of Traditional Resource Management Techniques: Conventionally, the resource allocation problems are solved using optimization methods considering the instantaneous Channel State Information (CSI) and QoS requirements of the users. Generally the resource allocation problems are not convex and, therefore, the solutions obtained by using the traditional techniques are not globally optimal. Also, the solutions may not be obtained in real time. For practical implementation, more efficient (e.g., from computation as well as performance point of view) solutions are needed. In some scenarios, where the resource allocation problem may not be well-defined mathematically (e.g., due to the nature of the propagation environment, mobility patterns of the users), we may not be able to formulate the optimization problem. It motivates the exploration of new resource allocation schemes. In this context, the data-driven Machine Learning (ML)-based resource allocation solutions would be viable solutions in such scenarios and should be adaptable to the dynamic nature of IoT networks [22].

3) Why Machine Learning (ML) for Resource Management?: ML can be defined as the ability to infer knowledge from data and subsequently to use the knowledge to adapt the behavior of an ML agent based on the acquired knowledge. ML techniques have been used in tasks such as classification, regression and density estimation. IoT devices generate sheer amount of data and the data-driven ML techniques can exploit these data to develop automated solutions for the IoT services. When the available data is massive and multi-dimensional, ML or more specifically, Deep Learning (DL) can be used for feature extraction and useful classification.

Also, we can use ML techniques when the prior knowledge of system, network, users, and parameters is not available and needs to be predicted and the control decisions need to be made. One such technique is Reinforcement Learning (RL) in which the unknown parameters and system behavior are monitored over time with *trial and error* and optimal control actions are learned. Another such technique is DL, a subset of ML and it utilizes a hierarchical level of Artificial Neural Networks (ANN) to carry out the process of learning. These ANNs work similar to human brain, with neuron nodes connected together like a Web. Traditional ML performs data analysis in a linear way, while the hierarchical function of DL enables data processing with a nonlinear approach. For resource management problems, when there is a *model deficit* or *algorithm deficit*, ML is recommended [23]. Model deficit means that there is insufficient domain knowledge or non-existent mathematical models, while algorithm deficit means that a well-established mathematical model is present but the optimization of the existing algorithms using these models is very complex. In such a situation, lower complexity solutions using ML are preferred. Also, when the contextual information is very important to be incorporated into the decision making process, ML techniques are most suitable. Most of the IoT applications fall under the above characterization because of the unknown system or network states and parameter values, and large number of devices generating enormous amount of data.

B. Scope of the Survey

In the following, we succinctly summarize the published surveys and compare our survey paper with the existing

TABLE I
EXISTING SURVEYS ON IoT

Year	Paper	Topic(s) of the survey	Related content in our paper	Enhancements in this article
2015	[6]	Enabling technologies, protocols and applications of IoT	Sec. II	Enhanced list of papers and coverage of more state-of-the-art solutions with focus on ML and DL
2017	[7]	Communication standards in IoT	N/A	ML- and DL-based generic solutions in IoT
2016	[8]	Network types, technologies and their characteristics	Sec. II	Focus on state-of-the-art solutions in IoT and coverage of resource management, security, and privacy
2017	[9]	Health-care communications standards in IoT	N/A	Focus on generic IoT and in-depth coverage of the solutions
2018	[10]	Architecture, scheduling, network technologies, and power management of IoT operating systems	N/A	Focus on the current solutions for generic IoT applications and future research directions
2018	[11]	Blockchain and SDN solutions for IoT	N/A	Generic IoT and in-depth review of the ML- and DL-based solutions in multiple domains
2018	[12]	Resource scheduling techniques in IoT	Sec. II	Detailed investigation of resource management in IoT and state-of-the-art techniques
2016	[13]	Secure routing protocols in IoT	N/A	Focus on the applications and resource management with security and privacy
2017	[14]	Trust models for service management in IoT	N/A	Coverage of overall aspects of security with solutions based on ML and DL
2019	[15]	Data fusion in IoT applications through ML	N/A	In-depth coverage of ML and DL in different aspects of IoT
2018	[16]	IoT data analytics through DL	N/A	In-depth review of ML- and DL-based solutions in IoT
2019	[17]	ANN-based ML for Wireless Networking	Sec. II/V	Particularly focusing on the resource management of IoT and cellular networks
2019	[18]	ML in wireless networks, mobility management, localization	Sec. IV	ML/DL for resource management in broad classes of networks both short/long ranged and high/low powered
2019	[19]	ML for massive machine-type communications	Sec. V	Further additions of ML/DL-based machine-type and IoT/cellular resource management
2018	[20]	Reinforcement learning for content-centric networks	N/A	Holistic view of machine learning algorithms for broad class of IoT networks
2018	[21]	ML for networks, traffic prediction/routing/congestion	N/A	Particularly focus on RM related issues of both IoT and cellular networks

surveys in terms of novelty of the covered topics and depth of contents.

1) Existing Surveys: Table I lists the existing surveys related to IoT networks. Most of the published surveys do not focus on the ML and DL techniques in IoT networks. The current published surveys either focus on particular aspects of the IoT or do not cover the resource management in a holistic way. In Table I, we focus on the topics that are covered in the existing surveys and we reflect on the same topics (if applicable) in this article. Furthermore, we also categorically discuss the enhancements in this article. The surveys in Table I focus on enabling technologies, standards, architectures, scheduling, solutions, routing, data fusion, and analytics. Although Artificial Intelligence (AI)-based and ML techniques have been used in IoT to address different challenges that were not addressed through traditional techniques, the role of ML and DL has not been covered in a holistic way in the existing surveys.

To fill this gap, we investigate in depth, the role of ML and DL in IoT networks and discuss the ML- and DL-based resource management solutions for wireless IoT. We have covered the latest surveys (till 2019) in this article. Note that the role of ML and DL is multi-faceted in IoT, therefore covering the entire spectrum of applications of ML and DL techniques in not possible in one article. We only focus on the resource management aspect of wireless IoT networks.

2) Summary of the Contributions: We cover different aspects of resource allocation and management in IoT networks that leverage ML and DL techniques. The objective of this survey is to bridge the gap between the resource allocation requirements of the IoT networks and the services provided by the ML and DL techniques.

Resource allocation and management are essential for the QoS guarantees in IoT applications and therefore need close attention. After covering the resource management challenges in different IoT networks (e.g., cellular IoT, low-power IoT, cognitive IoT, and mobile IoT networks), we discuss the existing solutions including those for Device-to-Device (D2D), HetNet, (massive) Multiple Input Multiple Output (MIMO), and Non-Orthogonal Multiple Access (NOMA) systems. Then we motivate using ML and DL in addressing the resource allocation and management challenges in IoT that were not addressed by the traditional solutions. We then dive deeper into the ML and DL techniques that address the challenges of the traditional resource management and are leveraged for resource allocation in heterogeneous IoT networks. We also outline the future research opportunities and direction. The pictorial illustration of the scope of this paper is shown in Fig. 1. The contributions of this article are summarized as follows.

1) We discuss in detail the traditional techniques used for resource allocation in IoT network and their limitations that motivate the use of ML and DL techniques.
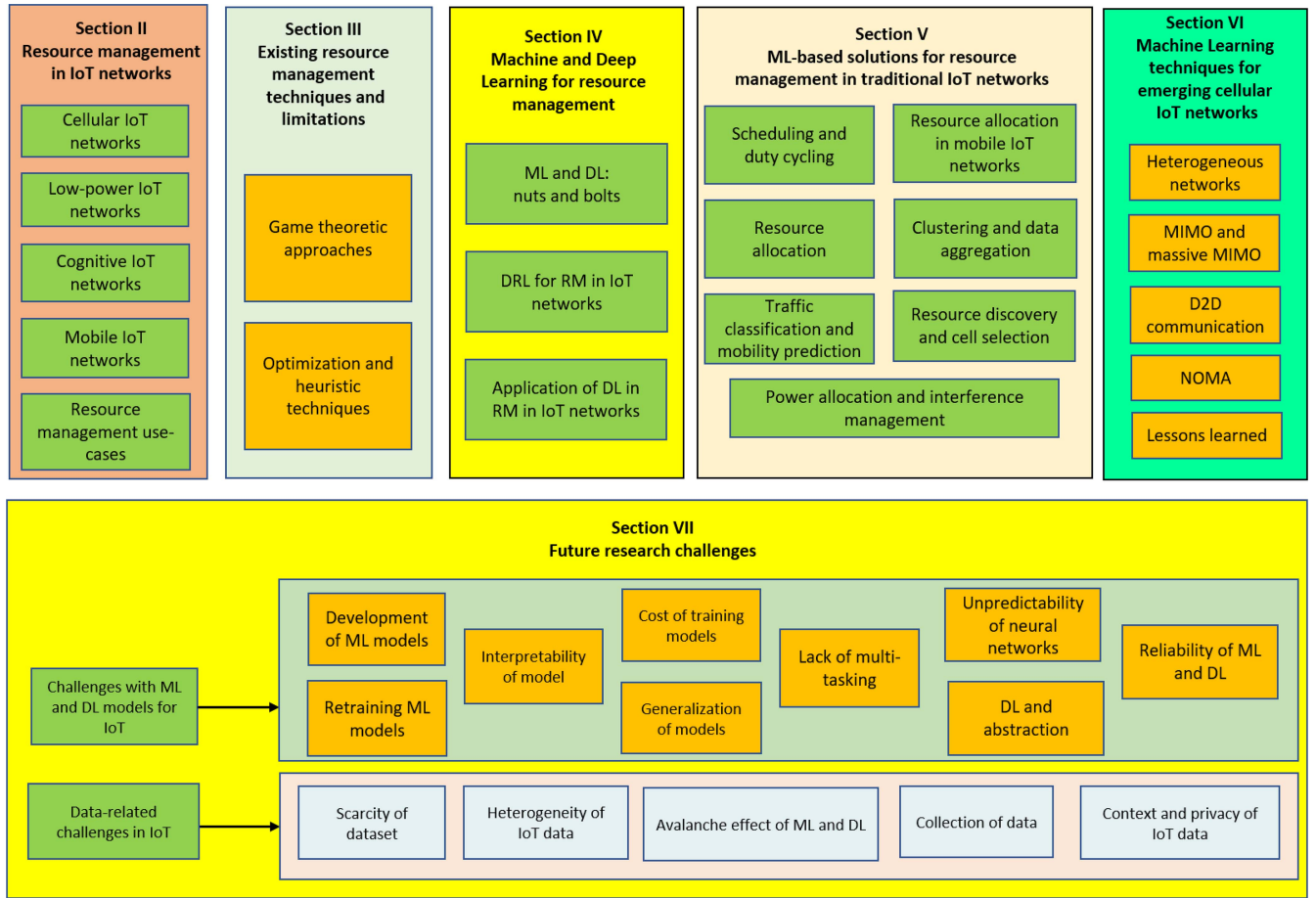
Fig. 1. Taxonomy of this survey. [ML: Machine Learning, DL: Deep Learning, RM: Resource Management, DRL: Deep Reinforcement Learning, NOMA: Non-Orthogonal Multiple Access].

2) We conduct an in-depth and comprehensive survey on the role of ML and DL techniques in the IoT networks.
3) We discuss the existing solutions in IoT networks that leverage ML and DL algorithms with an emphasis on the major aspects of the resource management in IoT networks.
4) Based on the conducted survey of the ML- and DL-based resource management in IoT networks supporting D2D, MIMO and massive MIMO, HetNets, and NOMA technologies, we identify the existing challenges and future research directions. This will stimulate further research in ML and DL for IoT networks.

The rest of the article is organized as follows: In Table II, we list all the acronyms used in the article. The challenges related to resource management in IoT networks are discussed in Section II. In Section III, we discuss the existing resource management techniques and their limitations followed by the role of ML and DL in resource management in Section IV. In Section V, we discuss the existing ML-based resource management solutions in IoT networks. Then in Section VI, we review the ML and DL-based solutions for resource management for the emerging cellular wireless technologies for IoT networks. In Section VII, we discuss the future research challenges in developing ML- and DL-based solutions for resource

management in IoT networks before we conclude the article in Section VIII.

II. RESOURCE MANAGEMENT IN IoT NETWORKS

In the following, we discuss the resource management challenges in four different breeds of the IoT networks, i.e., cellular IoT networks, low-power IoT networks, cognitive IoT networks, and mobile IoT networks.

A. Cellular IoT Networks

Cellular IoT networks can generally be characterized by the underlying technology they are using. For instance, the resource management techniques in different communication modes of a cellular IoT network (such as the D2D mode, HetNet mode, NOMA mode, or dynamic spectrum access-based mode) are different than those in a homogeneous cellular mode of operation. Generally the issues related to resource management in D2D systems include session setup and management [24], channel allocation and interference management in the underlay networks [25], [26]. HetNets also suffer from the similar problems of co-tier and cross-tier interference [27]. To date, several studies have been conducted on HetNet models and their resource optimization [28], [29]. Recently, NOMA has gained significant attention due to its potential

TABLE II
ACRONYMS AND THEIR DESCRIPTIONS

Acronym	Explanation	Acronym	Explanation
ML	Machine Learning	NOMA	Non-Orthogonal Multiple Access
WSN	Wireless Sensor Networks	DL	Deep Learning
RL	Reinforce Learning	SVM	Support Vector Machine
PRACH	Physical Random Access Channel	KNN	K-Nearest Neighbour
NB	Naive Bayes	NN	Neural Network
DNN	Deep Neural Network	CNN	Convolutional Neural Network
PCA	Principal Component Analysis	RNN	Recurrent Neural Network
MLP	Multi-Layer Perception	ELM	Extreme Learning Machine
ITS	Intelligent Transportation System	ANN	Artificial Neural Network
LSTM	Long-Short Term Memory	GAN	Generative Adversarial Network
V2V	Vehicle-to-Vehicle communication	DRL	Deep Reinforcement Learning
QoS	Quality of Service	CSI	Channel State Information
QoE	Quality of Experience	MIMO	Multiple-Input Multiple-Output
HetNet	Heterogeneous Network	D2D	Device-to-Device Communication
SDN	Software-Defined Network	CDSA	Control-Data plane Separation Architecture
RSSI	Received Signal Strength Indicator	URLLC	Ultra-Reliable and Low Latency Communication
DDoS	Distributed Denial of Service	LPWAN	Low-Power Wide Area Network
NB-IoT	Narrow Band Internet of Things	SINR	Signal-to-Interference-plus-Noise Ratio
BAN	Body Area Network	AML	Adversarial Machine Learning
NE	Nash Equilibrium	PSO	Particle Swarm Optimization
GPU	Graphics Processing Unit	MCS	Modulation and Coding Scheme
DOA	Direction-of-Arrival	D-SCMA	Deep learning aided Sparse Code Multiple Access
BER	Bit Error Rate	OFDM	Orthogonal Frequency Division Multiplexing
DQN	Deep Q Networks	SAX	Symbolic Aggregate approximation
MDP	Markov Decision Process	RBM	Restricted Boltzmann Machine
CSS	Cooperative Spectrum Sensing	CRN	Cognitive Radio Networks
TACL	Transfer Actor-Critic Learning	DPM	Dynamic Power Management
SMDP	Semi-Markov Decision Process	MWIS-DW	Maximum Weighted Independent Set problem with Dependent Weights
FIFO	First In First Out	RRM	Radio Resource Management
SIC	Successive Interference Cancellation	SDR	Software-Defined Radio

to provide high spectral efficiency. Some works have focused on the resource management aspect and optimization of the network parameters [30], [31], where the authors have used *K*-means for user clustering and ML-based interference management. The traditional methods for resource allocation in cellular IoT networks mainly rely on the optimization techniques. These methods face challenges when the number of users become large and/or when the multi-cell scenarios are considered. This is because the optimization space becomes overwhelmingly large to cater for the entire network, and therefore, computational complexity of finding solutions increases significantly.

B. Low-Power IoT Networks

Although the IoT networks offer both long-range and short-range communications through different technologies, there are many applications that require the energy portfolio of short-range technologies while demanding long-range communications. Therefore, low-power IoT networks especially those supporting long-range communications such as the Low-Power Wide Area Networks (LPWANs) become inevitable. LPWANs offer long-range communications by limiting the data rates and energy, and have been rigorously studied for various smart applications. Many technologies in the licensed bands (e.g., LTE-M, NB-IOT and EC-GSM) as well as those in un-licensed bands (e.g., LoRa, SigFox, Weightless) for low-power networks are under consideration which have to deal with different types of resource allocation problems [32]. The

authors in [32] presented emerging LPWA technologies and the standardization activities carried out by different standards development organizations (e.g., IEEE, IETF, 3GPP, ETSI) as well as the industrial consortia built around individual LPWA technologies (e.g., LoRa Alliance, Weightless-SIG, and Dash7 alliance). In most of the cases, the networks are characterized with a large number of nodes and the collisions among packets are avoided through scheduling and duty cycling [33]. The authors provided comprehensive and comparative study of leading LPWAN technologies such as, Sigfox, LoRa, and NB-IoT that compete for large-scale IoT deployment. They evaluated that Sigfox and LoRa are advantageous in terms of battery lifetime, capacity, and cost. Meanwhile, NB-IoT offers benefits in terms of latency and quality of service. Similarly, the problems of interference in these networks arising due to the underlying modulation schemes result in performance degradation of the system are discussed in [34]. The authors presented an analytical model of a single-cell LoRa system to account for the impact of interference. They modeled the interference field as Poisson point process under duty-cycled ALOHA and derived the SIR distributions under several interference conditions.

C. Cognitive IoT Networks

Many solutions have been proposed to solve different resource management problems in cognitive networks. One of the salient features of a cognitive IoT network is that the nodes constantly search for resources opportunistically that

are better suited for them to boost the overall performance of the network. The cognitive networks allocate resources of 'primary user' to so-called 'secondary user' when the primary user is absent or idle. However, as soon as the primary device is activated (the activity pattern of whom could be random), the secondary user has to vacate that channel. Therefore, resource allocation in these networks, which needs to consider primary users' activity and QoS requirements as well as secondary users' QoS requirements, wireless propagation, and other network parameters, is very challenging [35]. The channel sensing, detection, and acquisition are mainly done by static techniques that suffer from various imperfections such as collisions, increased outages, and decreased system throughput [36].

D. Mobile IoT Networks

Mobile IoT (MIoT) is the extension of traditional IoT with mobility [37]. More precisely, in MIoT, the IoT services and applications of IoT are mobile and can physically be transferred from one location to another. Another orthogonal variant of such IoT network is the Internet of Mobile Things (IoMT) where mobile *things* are employed to form an IoT network [38], [39]. Traditional IoT enables the communication among static *things* to realize different services and increase the connectivity among objects and the physical world, whereas in MIoT and IoMT, the communicating entities, i.e., smart things move and maintain their inter-connection and accessibility. Traditional IoT scenarios such as smart home mimic the static connectivity among objects, whereas mobile objects such as vehicles, mobile robots, wearable devices add the mobility dimension to the IoT. Traditional applications and services realized through IoMT and MIoT include, but not limited to, smart industry with mobile robots, smart transportation with vehicles, and so on.

As has already been mentioned, the defining feature of the IoMT and MIoT is the node mobility. Since mobility incurs extra control information to be exchanged among the network management entities, resource management in IoMT and MIoT is more challenging than the traditional static IoT networks. In addition to these challenges, application context is also an important parameter to consider for resource management.

E. Resource Management Use Cases

To discuss the importance and requirements of ML techniques in various IoT applications, we focus on resource management in cellular IoT networks and smart home environment:

- *Resource allocation in cellular IoT networks:* Cellular IoT networks (e.g., 5G cellular and beyond) should support extremely high data rates and require continuous connectivity to serve heterogeneous devices and users with varying QoS requirements. The transmitting devices are expected to autonomously connect to an Access Point (AP) or Base Station (BS) and access the frequency channels for achieving acceptable Signal-to-Interference Ratio (SIR). Furthermore, careful channel

and power allocation is required to support various IoT applications simultaneously. For this purpose, obtaining the network parameters such as CSI, traffic characteristics, and demands of the users/devices are challenging, especially in case of serving highly mobile users. Also, estimation of the communication parameters such as the Angle of Arrival (AoA) and Angle of Departure (AoD) are important for millimeter wave communications. To address the above challenges, ML can be a suitable solution from both the network and user perspectives. IoT networks (e.g., network entities such as eNodeB) are expected to learn the context and diverse characteristics for both human and machine users, respectively, in order to achieve optimal system configurations. Smart devices are expected to be capable of autonomously accessing the available spectrum with the aid of learning, and adaptively adjusting the transmission power for the sake of avoiding interference and for conserving the energy. For this purpose, Deep Reinforcement Learning (DRL) and linear regression can be used. In millimeter Wave (mmWave) communications, the values of CSI and AoA, which are not known a priori, can be estimated based on an RL technique, e.g., Q-learning. Location prediction of high-mobility users can be obtained using Support Vector Machine (SVM), and Recurrent Neural Networks (RNNs) [47], [48], and this location estimation can be used for resource allocation. Various correlated system parameters such as users' position and velocity, propagation environment, and interference in the system can be taken into account to develop ML-based resource allocation methods. Furthermore, techniques such as Principal Component Analysis (PCA) and K-means clustering can be used for clustering various types of devices so that the available resources can be allocated according to the QoS requirements of these groups or clusters [49]–[52]. It is worth noting that traditionally, the optimization techniques cannot incorporate the context and are unable to react according to the changing environments [53], [54]. Therefore, using these techniques may result in poor resource utilization. Also, the game theoretic approaches often do not incorporate heterogeneity of players as well as dynamic change in the network topology and number of players for resource allocation [54].

- *ML for resource management in smart home environment:* Smart home applications are among the most popular IoT applications that involve a combination of heterogeneous and ubiquitous devices such as security cameras, hand-held scanners, tablets, smart appliances, and wireless sensors. All of these devices randomly access the network resources and have varying access and QoS requirements. ML techniques such as Q-learning and multi-armed bandit can work well for scheduling and random access of resources in the smart home environment. It is because these reinforcement learning methods learn and can adapt dynamically to the variation in the network environment [55]–[57]. Also, *K*-means clustering and PCA can be used to group small sensors and aggregate the small payload data, respectively. Note that

TABLE III
EXISTING SURVEYS ON RESOURCE MANAGEMENT USING CONVENTIONAL METHODS IN CELLULAR NETWORKS

Topic(s) of the survey	Related content	Enhancements in this paper
Device-to-Device (D2D) Communications	[24], [25], [26], [40],	ML-based RM algorithms specially focusing Q-learning and bandit learning
Heterogeneous Networks	[27], [28], [29], [41],	Focus on deep learning techniques for various RATs in heterogeneous networks
Vehicular Networks	[42], [43]	Focus on reinforcement learning techniques for various RM related issues
Multiple Input Multiple Output (MIMO) and Massive MIMO	[44], [45], [46]	Focus on various ANN and KNN-based algorithms mainly for channel estimation and resource sharing
Non-Orthogonal Multiple Access (NOMA)	[30], [31]	Focus on K-means for user clustering and ML-based interference management

the traditional optimization and heuristics-based techniques can be computationally very expensive to run on small, inexpensive, and energy-constrained sensors. Furthermore, traditional game theoretical approaches may not be suitable due to heterogeneity of the devices as well as overhead due to the information updates and exchange.

III. EXISTING RESOURCE MANAGEMENT TECHNIQUES AND LIMITATIONS

In this section, we discuss the existing resource management tools and techniques in IoT networks and the enabling technologies.

A. Optimization and Heuristic Techniques

To date, many solutions have been proposed for efficient resource allocation in IoT networks using optimization and heuristics-based techniques. For instance, the authors in [58] used genetic algorithm for load balancing in fog IoT networks. However, the complexity of the algorithm increases with increasing the number of user requests and network size. In the same spirit, the authors in [59] proposed a resource allocation approach using genetic algorithm. They transformed the resource allocation problem into a degree-constrained minimum spanning tree problem and applied a genetic algorithm to solve the problem in a "near-optimal" fashion. However, the approach does not scale well in a dynamic IoT environment and may not be implementable in a practical scenario. Similarly, the authors in [59] used Particle Swarm Optimization (PSO) based meta-heuristic for the distribution of blocks of codes among the IoT devices. They developed multi-component applications capable of being deployed on various IoT nodes, edge, and cloud. However, it works well in static environment and lack its applicability in multi-objective and dynamic environments.

In general, the traditional optimization and heuristics-based techniques have the following characteristics.
- *Inaccuracy of the model-based resource allocation methods:* Numerous model-oriented methods have been developed in the literature for resource management (e.g., for power allocation and interference management). However, the mathematical models used for these methods, although analytically tractable, are not always accurate due to the practical channel environment and hardware imperfections. In this context, model-free and

data-driven methods are potentially more promising. Data-driven or model-free solutions correlate the system observed measurements with a state of a system and solve regression or classification problems, or assign a probability to a system state. This is in contrast to a model-driven approach, which attempts to capture knowledge and it derives decisions through explicit representation and rules.
- *Cost and complexity:* Heuristic techniques (e.g., branch and bound) are also computationally expensive and incur significant timing overhead. Therefore, it will not suit well in delay-sensitive and real-time IoT applications. Moreover, the computational complexity of these heuristics proportionally increases with the increase in the network size. Also, with evolutionary algorithms such as genetic algorithms, designing the objective functions and describing the parameters and the operators can be difficult.
- *Convergence to local optima:* For sufficiently complex problems (e.g., non-convex problems), the task of obtaining a global optimum solution in the presence of many local optimum points is very challenging.
- *Sub-optimal solutions in high dimensional problems:* With traditional heuristics-based algorithms, the quality of the solutions can degrade with the increase in dimensionality of the problem. Heterogeneity and dynamically changing characteristics of the IoT networks call for more intelligent, self-adaptive, and auto-configurable techniques.
- *Parameter sensitivity and lack of flexibility:* The solutions obtained through heuristics-based approaches are sensitive to the corresponding chosen system parameters. If the key decision variables and the constraints change (e.g., due to new devices added to the network), a hard- or pre-coded heuristic technique may no longer remain feasible and it may need to be reconfigured for producing more viable solutions. Also, the traditional heuristics algorithms do not utilize the contexts of the users/devices and the network, and therefore, are unable to react to a dynamically changing environment.

B. Game Theoretical Approaches

Game theory is generally used for distributed resource allocation in wireless and IoT networks in the presence of resource competition or cooperation among nodes (e.g.,

non-cooperative and cooperative game theory techniques, respectively). To date, many game-theoretic mechanisms have been employed to address the resource management issues in IoT networks [60]–[64]. In [63], the authors proposed a non-cooperative game-based resource allocation mechanism for LTE-based IoT networks. The authors focused on reducing the inter-cell and intra-cell interference and proposed an efficient power allocation algorithm. In [65], the authors proposed an enhanced mobile broadband IoT network capable of fog computing and solved the resource management problem when an added constraint of ultra reliability is considered. Specifically they used a game theoretical framework using an analytic hierarchy process to solve the optimization problem.

In another work, Zhang *et al.* [62], [64] proposed an optimized resource allocation framework for fog-based IoT networks. Specifically, the authors employed a Stackelberg game among data service operators, subscribers, and the fog nodes. The problem of computing resource allocation was formulated as a game between the service operators and subscribers. Similarly, Semasinghe *et al.* employed non-conventional game theoretic approaches for resource management in large-scale IoT networks [60]. The authors used evolutionary, mean field, minority games, and other similar approaches to solve the resource management problems in IoT networks. Kharrufa *et al.* [61] used a game theory-based approach to optimize resource allocation, where the mobile nodes compete for network resources and aim at reaching a Nash Equilibrium (NE) solution for resource management. In fact, most of the game-theoretic models aim at obtaining the NE solution for the resource allocation problem.

Despite their advantages, there are some limitations of the game-theoretic approaches for resource management. We discuss the limitations as follows.

- *Homogeneity of players and availability of network information:* Generally, NE solutions are based on the assumption that all the players are homogeneous, have equal/similar capabilities, and have complete network information. However, heterogeneous IoT devices have different computational capabilities, use different communication technologies, and exhibit different behaviors. Therefore, the conventional NE solutions may not be appropriate for a practical IoT environment. Also, developing game models and obtaining the corresponding solutions for a system with incomplete and/or imperfect information can be very challenging.
- *Lack of scalability and slow convergence:* The number of players plays an important role in the applicability of game-theoretic approaches. When the number of players increases, with a traditional game model, obtaining the solution of the model becomes more complicated. Although specific game models such as the evolutionary games or mean-field games are suitable for modeling competition among large number of nodes (i.e., players) [60], due to the poor convergence properties, however, the distributed solutions obtained using these approaches may not be feasible for practical systems. Therefore, in an IoT network with large number devices, use of game-theoretic approaches for resource allocation can be limited.

- *Information exchange overhead:* Implementation of game-theoretic approaches may require a significant amount of information exchange (which will lead to delay), and is not practical due to the massive number of devices in the IoT networks. These devices have limited memory and energy, and cannot maintain the records of all the actions pertaining to the devices in the network.

IV. MACHINE LEARNING AND DEEP LEARNING FOR RESOURCE MANAGEMENT

In this section, we discuss the basics of ML, DL, and DRL[1] and their suitability for resource management in IoT networks.

A. Machine Learning (ML) and Deep Learning (DL) Basics

Broadly speaking, ML is divided into three categories: supervised, unsupervised and reinforcement learning techniques.

Supervised learning: These techniques use models and labels known a priori, and can estimate and predict unknown parameters. Support Vector Machine (SVM), Naive Bayes Classifier, Random Forest, and Decision Tree (DT) are most commonly used supervised learning algorithms used for classification and modelling of the existing data sets. This family of algorithms is mainly used for channel estimation, localization, spectrum sensing problems.

Unsupervised learning: These techniques are used with unlabeled data and utilize input data in an heuristic manner. Potential applications of these algorithms in IoT networks include cell clustering, user association and load balancing. *K*-Means Algorithm and Principal Component Analysis (PCA) are two well known unsupervised ML algorithms used for classification of devices based on attributes / parameters and data compression as well as dimensionality reduction, respectively.

Reinforcement Learning (RL): RL techniques do not require training data sets and an agent learns by interacting with the environment [111]. The objective is to learn the *optimal policy* (i.e., optimal mapping between state and action), such that the long-term reward is maximized for the agent. Q-learning is one of the most popular RL techniques. RL techniques can be utilized by the IoT devices and sensors for decision making and inferring under unknown and dynamically changing network conditions. For instance, RL can be utilized during spectrum sharing for channel access in cognitive radio networks (in the presence of unknown channel availability) and in small cell area networks (with unknown resource conditions). It can also be utilized for BS association, specially with unknown energy and load status of the BS. However, RL techniques may suffer from slow convergence rate (in terms of learning the optimal policy at a system state) and curse of dimensionality, which can create challenges for the RL techniques in dynamic IoT networks.

Deep Learning (DL): DL is a breed of the ML techniques which is based on Artificial Neural Networks (ANN) with multiple hidden layers between the input and output layers. A Deep Neural Network (DNN), when trained, approximates

---

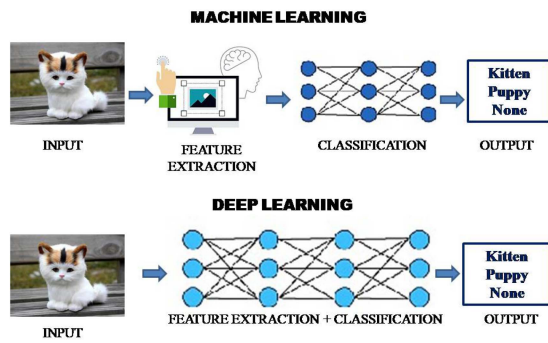[1]A more detailed description of the ML and DL techniques can be found in [66].

Fig. 2. Machine learning vs. deep learning.

the functional relationship, that is present in a data set. It is expected that the relationship will be generalized for data points not present in the training data set. Various neural networks are shown in Fig. 4. DL is recommended for solving various problems in prediction, classification, resource allocation, spectral state estimation, object detection, speech recognition and language translation. DL learns features at multiple abstraction levels and allows a system to learn complex functions, and maps the input to the output directly from data [112]. For instance, in Fig. 2, unlike ML, DL performs feature extraction and classification at various layers under the same process, in which these are two separate processes. In DL, the first layer learns primitive features such as color and shape in an image by finding frequently occurring combinations of digitized pixels. These identified features are fed to the next layer, which trains itself to recognize more complex (combination of already identified) features, such as edge, corner, nose, ears etc. of an image. These features are fed to the next layer, which performs identification of more complex features (e.g., human, animal). This recognition process is continued for successive layers until the object is identified. The available open source DL frameworks/platforms include Torch/PyTorch, Thenano, TensorFlow and MXNet. DL is considered better than the traditional ML techniques due to following reasons:

- DL is capable of handling large amounts of data and DL algorithms are scalable with increasing amount of data, which can be beneficial for the model training and may improve prediction accuracy.
- DL can automatically and hierarchically extract high-level features and corresponding correlations from the input data. Therefore, the learning model is not comprised of multiple non-linear layers designed by the humans. It simplifies and automates the entire process. This automatic feature extraction, which is not very common in traditional ML techniques, can be beneficial for IoT networks where data generated by heterogeneous sources can exhibit non-trivial spatial/temporal patterns.
- IoT applications mostly generate unlabeled or semi-labeled data. DL has the capability to exploit the unlabeled data to learn useful patterns in an unsupervised manner (e.g., using Restricted Boltzmann Machine (RBM) and Generative Adversarial Network (GAN)). More precisely, RBM is the restricted variant of

traditional Boltzmann machine, where the neurons in the two layers of the machine must form a bipartite graph and there must not be any intra-layer layer communication. RBMs are poised to be more efficient learning algorithms [113]. These RBMs are also used in deep belief networks where RBMs are stacked together. Furthermore, GAN is another alternative to traditional ML. Recently, GANs have been extensively used in the applications leveraging ML and DL methods [114]–[116]. In GAN, the two neural networks called generator and discriminator, compete with each other in an adversarial setting and enhance the learning process. The *generator* generates new data instances and the *discriminator* decides whether the instance belongs to real data set or not. The main idea for generator is to generate real instances, whereas for the discriminator is to learn to detect the non-real instances. This way, these two networks become stronger against each other till the point of their convergence. It is worth mentioning that data is not labeled at the input in these scenarios. On the other hand, the conventional ML algorithms work effectively only when sufficient labeled data is available.

- DL also reduces the computational and time complexity as single trained model can be used for various tasks and can achieve multiple objectives without model retraining. Graphics Processing Unit (GPU)-based parallel computing enables DL to make predictions (such as resource allocations) very fast. Also, when used along with technologies such as SDN and cloud computing, DL makes storage, analysis and computation of data easier and faster.

However, the major drawback of DL is that a large amount of data will be required for training and testing, which may not be available or could be hard to generate.

Deep Reinforcement Learning (DRL): DRL is a combination of DNN and RL where a DNN is used as an agent for RL. In other words, a DNN approximates the optimal policy function and it is trained using the data obtained through interaction with the environment. Therefore, no prior training data is required. For problems in which the number of system states and data dimensionality are very large and the environment is non-stationary, using traditional RL to learn the optimal policy may not be efficient (e.g., in terms of convergence rate). By combining DNN and RL, agents can learn by themselves and come up with a policy to maximize long-term rewards. In this approach, the experience obtained through RL can help the DNN to find the best action in a given state [117]–[121].

DRL is broadly divided into two categories: Deep Q Networks (DQN) and Policy Gradient. DQN is a value-based method while policy gradient is a policy-based method. In a value-based method, a Q value is calculated for every action in the action space and selection of the action depends on the Q value). In a policy-based method, the optimal policy is learned directly from a policy function (which gives the probability distribution of every action) without worrying about the value function (i.e., without calculating the Q-functions). In DQN, Q-learning is combined with a DNN and the state
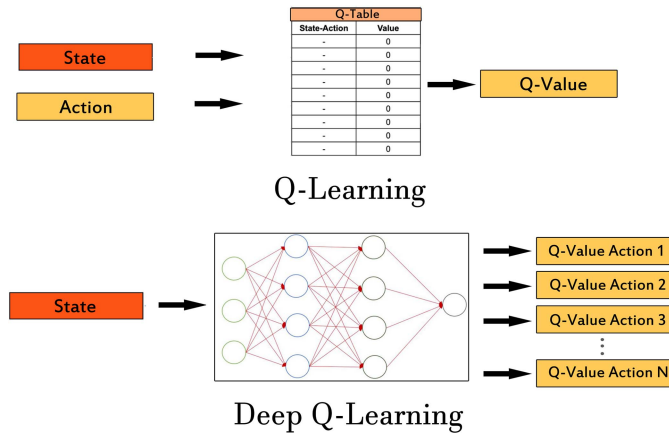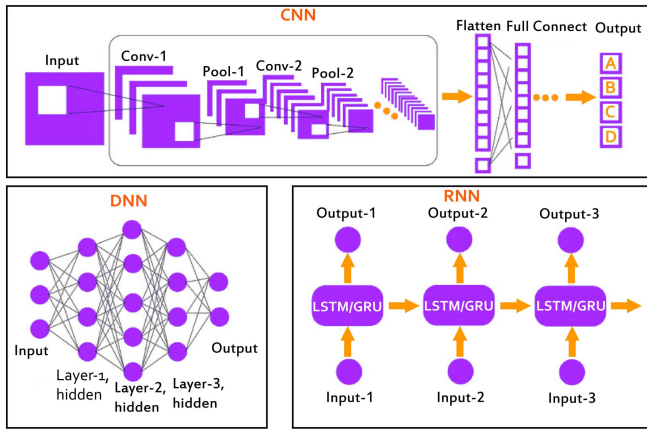
Fig. 3.   Deep Q-learning.

Fig. 4.   Various deep neural networks.

action value function (Q) is determined by the DNN. In a policy gradient method, the policy with respect to the expected future reward is directly learned by optimizing the deep policy networks using gradient descent.[2]

DRL is considered as a promising technique for future intelligent IoT networks, and can be applied to tasks with discrete and continuous state and action space, and also with multi-variable optimization problems. DRL has been proposed for IoT networks for predictive analysis and automation [52]. Detailed description of DL Methodologies and artificial neural networks can be found in [122]. Authors provided comprehensive literature survey on DL in wireless communication networks followed by various case-studies, in which DL has been useful. In the same spirit, authors presented and discussed various advance level Deep Q Learning (DQL) models in [123]. DQL (Fig. 3) [197] inherits the advantages of DL and Q-learning and find its use fullness in applications having large state and action spaces. To find the optimal policy and correlations between the Q-values and the target value (with varying data distribution), two DQN models; experience replay and target Q-network, are presented.

[2]https://medium.com/deep-math-machine-learning-ai/ch-13-deep-reinforcement-learning-deep-q-learning-and-policy-gradients-towards-agi-a2a0b611617e

B. Application of DL for Resource Management in IoT Networks

A flurry of recent research works has used DL techniques for solving the resource allocation problems in IoT networks. For instance, a Q-learning based algorithm that uses cognitive radio-based IoT was proposed in [74]. The proposed algorithm uses DL to optimize the transmission of different packets of variable sizes through multiple channels so that the overall network efficiency is maximized. It has also been used in cognitive IoT networks [99], [124] where a cooperative strategy between secondary users is used to detect multiple occupancy of bands in a primary user. In the same spirit, the authors in [125] proposed DRL-based spectrum sensing approaches using which multiple pairs of secondary users coexist with these primary users.

DL has also been used for advanced wireless transmission technologies such as MIMO and NOMA. For instance, [126] used DL for high-resolution channel estimation in MIMO systems. Reference [127] proposed a DL technique, referred to as DL-aided Sparse Code Multiple Access (D-SCMA) for code-based NOMA. An adaptive code book is constructed automatically which decreases the Bit Error Rate (BER) performance of a NOMA system. Using DNN-based encoder and decoder, a decoding strategy is learned and formulated. Simulation results for proposed scheme shows smaller BER and low computational time as compared to conventional schemes.

To this end, DL has also been used for load balancing and interference mitigation in IoT networks. For instance, the authors in [84] developed DL-based load prediction and load-balancing algorithms for IoT networks. Also, the authors in [87] presented a DL-based interference minimization method. Similarly, [128] used unsupervised DL in an autoencoder that optimizes the transmit parameters to provide diversity gains. In [129], the authors provided an DL mechanism to estimate the CSI in an Orthogonal Frequency Division Multiplexing (OFDM) system, where the estimator performs better for pilot training and channel distortions. It performs channel estimation and signal detection for wireless communications even in presence of complicated channel distortion and interference. The DL-based approach is more robust than conventional methods as fewer training pilots are used and the cyclic prefix (CP) is omitted.

C. A Sneak Peek of DRL for Resource Management in IoT Networks

DRL has been used for resource allocation problems to achieve high spectral efficiency. For example, the authors in [4] developed a deep Q-learning based spectrum sharing approach for primary and secondary users in which the primary users have fixed power control and the secondary users adjust their power after autonomously learning about the shared common spectrum. Similarly, the authors in [130], proposed a novel centralized DRL-based power allocation scheme for a multi-cell system. Specifically, the authors used a Deep Q-learning (DQL) approach to achieve near-optimal power allocation and also to maximize the overall network throughput.

Furthermore, in [131], the authors proposed a DRL-based throughput maximization scheme in small cell networks. The DRL algorithm is developed by using a specific type of deep Recurrent Neural Network (RNN), called a Long Short-Time Memory (LSTM) network for channel selection, carrier aggregation, and fractional spectrum access.

In [132], a multi-agent DRL method was used for resource allocation in vehicle-to-vehicle (V2V) communications. DRL is used to find the mapping between the local observations of each vehicle such as, local CSI and interference levels affecting the resource allocation decision. Each vehicle (agent) makes optimal decisions for transmission power by interacting with the environment. Ye and Li [132] also used a DRL-based approach for resource allocation in vehicular networks. To reduce the transmission overhead, this approach uses local information on the available resources rather than waiting for the global information. The authors in [133] proposed a DL-based approach to predict the resource allocation based on 24 hour data count in V2V communication. He *et al.* [134] proposed a DRL-based approach for optimized orchestration of computing resources in vehicular networks. Tan and Hu [135] used DRL in a multi-timescale framework for cache replacement and computing resource management. A detailed discussion about various DRL strategies can be found in [136].

V. ML-BASED SOLUTIONS FOR RESOURCE MANAGEMENT IN TRADITIONAL IoT NETWORKS

A summary of the ML-based resource management methods in IoT networks is provided in Table IV. A review of these methods is provided below.

A. Scheduling and Duty Cycling

For an energy-efficient operation, the nodes in an IoT network would be generally sending data intermittently and in short packets. This is to conserve the energy of the over-all network entities for network lifetime longevity. In addition to traditional ways of controlling this duty-cycling, ML has been recently used to optimize these networks and system level parameters. In [81], the authors propose an ML-based up-link duty cycle period optimization algorithm for a broad class of IoT networks such as Long Range Wide Area Networks (LoRaWANs). In LoRaWANS, hundreds of IoT devices transmit packets to a central aggregator using a scheduling mechanism to avoid collision. Therefore, the authors in [81] optimize the up-link transmission period of LoRa devices using a Bayes' classifier of ML that improves the network energy efficiency as well as reducing the collision rate. A scheduling method for an IoT network for a smart home application is studied in [75] where the authors compared various Q-learning algorithms for scheduling of home appliances. It has been noticed that various algorithms demands different configuration of Q-learning method to optimize the scheduling resources in home networks.

Traffic scheduling in IoT systems need to consider various kinds of traffic such as delay-tolerant traffic and video and/or voice traffic. In [77], an RL-based traffic scheduling method

was proposed that dynamically adapts to various traffic variations in a real-life scenario. More specifically, the authors applied RL algorithms on a 4-week real data traffic from a particular city and showed that the network performance almost doubles as compared to traditional scheduling. On the other hand, for a cloud-assisted IoT network, an ML-based task scheduling algorithm was proposed in [79], which uses multiple criteria to optimize the network performance based on the capability of cloud entities and complexity handling.

B. Resource Allocation

ML can be used to solve the resource allocation problems in IoT networks, where a huge data set can be collected and used to train algorithms that produce very robust results for various resource allocation problems. In this context, [137] uses cloud computing and ML together to perform resource allocation in a general wireless network and applies their proposed strategy for beam allocation in a wireless system. A comprehensive analysis of cloud computing-based resource allocation algorithms was also discussed in the same paper. Similarly, [80] used different ML algorithms in a cloudlet-based network to efficiently optimize the network in terms of distributing computing and processing functions over the entire network entities. In cloudlet-based networks, the resources should be ideally spread over the entire available cloudlet nodes. Therefore, the authors proposed two heuristic resource allocation algorithms based on Steepest Descent (SD) and Simulated Annealing (SA) and observed that SD performs 4 times less reallocations than SA in the same network.

Note that, for the centralized wireless IoT networks such as cellular networks, resource allocation is generally performed by the BS based on the CSI of the users/devices and their Quality of Servie (QoS) requirements. However, for distributed IoT networks such as Machine-to-Machine (M2M) and ad-hoc networks, there is no centralized control and the users may not have the channel statistics available for the entire network. Channel allocation tasks in such scenarios can harness the benefit of state-of-the-art ML techniques.

In this context, ML-based resource management algorithms have also been studied for video applications in IoT networks. In [138], the authors used a mixture of supervised and unsupervised learning algorithms to optimize the Quality-of-Experience (QoE) of the end users when video streams are transmitted to them over a wireless channel. The learning algorithms extract the quality-rate characteristics of unknown video sequences and manage simultaneous transmission of the video over channels to guarantee a minimum QoE for users. Heterogeneous cloud radio access networks have also gained attention in the past few years for which the resource allocation and interference management are the key problems. To this extent, in [139], the authors proposed a centralized ML scheme to limit the interference below a certain threshold which increases the energy efficiency of the network. To boost the performance of their cloud-based network, they used a model-free online learning algorithm and consider priority in assigning resource blocks to avoid interference, thereby, achieving energy efficiency. A similar study on resource

TABLE IV
RESEARCH PROBLEMS IN IoT NETWORKS AND APPLIED MACHINE LEARNING TECHNIQUES

Research Objective/Problems	Machine Learning Techniques (Surveyed References)	Comparison of ML Techniques
Clustering and Data Aggregation	- K-Means Clustering [52], [51], [67], [68] - Fuzzy C-Means Clustering [69] - Supervised Learning [70], [71] - Deep Learning [72], [73] - Principal Component Analysis (PCA) [49], [50]	- K-Means clustering are easy to implement and computationally faster even with large number of variables. However, it requires value of K to be known a priory. Also, initial value of seeds have a strong impact on the final results. - Fuzzy C-Means (FCM) is a method of clustering which allows one piece of data to belong to two or more clusters and gives best result for overlapped data set. It is better suited for spectral data and allows class overlaps. - PCA is used to reduce the dimensionality of feature data and project all data points into the lower-dimensional subspace. Afterwards, any clustering method is used for grouping and classification.
Resource Allocation (Scheduling, Random Access)	- Q-Learning [52], [74], [75] - Multi-Armed Bandit [76] - Reinforcement Learning [77], [78] - DNN [79] - Simulated Annealing [80] - Bayes Classification [81]	- Multi-armed bandit learning is used when there are multiple options to select and for each set of selection, the user is given a reward. - Reinforcement learning, on the other hand, is about making decisions sequentially where a reward is given at each individual step.
Traffic Classification and Mobility Prediction	- Clustering [82] - Deep Learning [83], [84] - Recurrent Neural Network - Supervised Inductive Learning [85] - SVM [47], [48]	- Clustering techniques are used for anomaly detection in traffic flow and has capability to detect unusual patterns in traffic. - DL based CSI and channel detection is accurately performed despite the presence of complicated signal distortion and interference. - Random Forest and Decision Tree algorithms are the best classifiers for achieving accuracy and computational efficiency in traffic classification.
Power Allocation and Interference Management	- Deep Reinforcement Learning [86]–[88] - CART Decision Tree [89] - Linear Regression [90] K-Means Clustering [91] - Supervised Learning [92] - KNN [93] - Q-Learning [94]–[96] - Deep Learning [97]	K-Means clustering group the users in an area and allocate powers to each of them to minimize the overall interference of the network. - Supervised learning relies on a previously acquired data set of the number of users, the environment, the underlying network topologies to optimize power allocated to users on run-time as a test data. - Decision trees and Q learning are used to make a hierarchy of power allocations for users in different network settings.
Resource Discovery, Cell and Channel Selection	- Deep Learning [98], [99] - Supervised Learning [100], [101], [102] - PCA [103] - Bayesian Learning [104], [105] - KNN [103], [106] - Reinforcement Learning [52] [107], [108] - Support Vector Machine [103], [109], [110]	- KNN groups a particular set of channels, in a non-parametric way, which could be allocated to a set of users/devices for optimal network performance. An initial value of K is specified. - SVM discriminatively labels the data in various classes. The algorithm is used for channel or cell selection when a variety of APs or BSs are deployed for device connectivity.

allocation in wireless systems is presented in [78] where the authors used DRL algorithms to manage resources efficiently in the presence of interference and noise.

C. Power Allocation and Interference Management

Power allocation plays a vital role for enhancing the performance of any network as it reduces the interference to

other network entities. In [89], the authors used distributed Q-learning and Decision Tree algorithms for interference and power management in D2D communication networks. To solve the power allocation to various users, they used a multi-agent Q-learning algorithm where each user gets an optimal share of power resources. The authors were able to reduce the time complexity of the algorithm using binary trees and improved the system capacity and energy efficiency. In [90], a system was proposed for interference level detection and power adaptation according to the interference level in the radio channels. The authors utilized a linear regression algorithm which uses CSI to predict the transmit power levels in order to minimize the interference and power wastage. Similarly, the authors in [88] proposed a hierarchical framework for resource allocation and power management in a cloud computing environment. The authors used a DRL method to develop a Dynamic Power Management (DPM) policy for minimizing power consumption.

Collaborative distributed Q-learning was proposed in [94] to reduce congestion in the access channel for resource-constrained MTC devices in IoT networks. The proposed system finds unique Random Access (RA) slots and reduce the possible collisions. Similarly, in [93], the authors presented a model incorporating ML for low power transmission among IoT devices using LoRa.[3] Specifically, they introduced a CSI learning mechanism in LoRa edge devices which is then used to control the power of the entire network. Moreover, the authors in [92] developed a real-time burst-based interference detection and identification model for IoT networks. They developed an analytical model for interference classification by first extracting Spectral Features (SFs) by exploiting the physical properties of the target interference signals. Afterwards, they used supervised learning classifiers such as multi-class SVM and Random Forest tree for the isolation and identification of interference.

D. Clustering and Data Aggregation

Clustering is one of the techniques to avoid overload and congestion in an IoT network. Clustering can be performed on the basis of parameters such as Euclidean distance, Signal-to-Interference-plus-Noise Ratio (SINR), device types, and QoS parameters. The most commonly used technique is K-means clustering algorithm, which is an unsupervised ML algorithm. K-means clustering partitions the available data set into K clusters and every data point belongs to a cluster which has the least Euclidean distance from the centroid of that cluster. In [68], the authors propose a normalization algorithm to compute the distances iteratively, which increases the accuracy of clustering and also decreases the clustering time. In [69], the authors proposed and implemented a hybrid semi-supervised ML technique employing Active learning SVM and Fuzzy C-Means (FCM) clustering for deigning an efficient IDS. Authors tested their model on NSL KDD bench mark IDS data set and revealed promising results. Moreover, their model can update

---

[3]LoRa is a long-range low-power radio transmission technology which uses the license-free sub-GHz frequency bands.

itself with minimum computational overhead (when new training data is included) and act as an low cost semi-supervised IDS model that can perform well with little labelled data.

In [140], a Semi-supervised Multi-Layered Clustering (SMLC) model was presented for the detection and prevention of network intrusion. SMLC builds an ensemble model of multiple randomized layers using a K-Means Algorithm and it achieve comparable/better performance even with partially labeled data, when compared with supervised ML models such as RF, and AdaBoostM1 in terms of model accuracy, detection rate, training/testing time (on two benchmark network intrusion datasets, NSL and Kyoto 2006+). The local learning models of SMLC are learned from the resultant clusters at different layers. The final prediction of a test instance is obtained by choosing the classification with the most votes among all decisions from all layers. Authors used the concept of the weighted Euclidean distance measure and atomic clusters,thus reducing the classification time and update time of the proposed model.

The data streams generated in an IoT network can be huge and the data are generally unlabeled. The processing of raw data is computationally expensive and manually labelling the data is not possible. In this context, [71] uses the Symbolic Aggregate approximation (SAX) algorithm to reduce the dimension of data stream and data is labeled using density-based clustering. The density-based clustering extracts the number of classes in data and labels them using the found classes. In [50], the authors proposed INSTRUCT, an access control mechanism for IoT devices using clustering techniques. Proposed model administered access control at the intermediate switch/gateway level by extracting the valid patterns out of the previous communication history using the clustering. It used two separate algorithms for UDP and TCP traffics and achieved 100 percent success for TCP and 95 percent for UDP traffics, compared to the signature-based manual analysis.

E. Resource Discovery and Cell Selection

A multi-band spectrum sensing policy was devised in [98], where the authors used an RL-based algorithm to sense the free channels of primary users. In [100], the authors proposed a cooperative sensing scheme where multiple primary users with complex channel states cooperate using Extreme Machine Learning (ELM) algorithm to obtain accurate channel state information and classification. This enables the secondary users to successfully access the idle channel which is not in use by primary user. Spectrum sensing can be very challenging in large-scale heterogeneous IoT networks. The authors in [105] proposed a new spectrum sensing scheme using Bayesian machine learning in which heterogeneous spectrum states are inferred using the Bayesian inference carried during the learning process.

In [109], the authors used an SVM-based cooperative spectrum sensing technique. Similarly, in [108], the authors presented an intelligent spectrum mobility management model for cognitive radio networks which can be used by the secondary users to switch among the spectrum bands. The authors

used an ML scheme known as Transfer Actor-Critic Learning (TACL) which was used for spectrum mobility management. The TACL algorithm utilizes the transferred learning expertise in historical periods or neighboring regions and use the polices made in adjoining regions to learn the current environment. For spectrum mobility, two modes are investigated in [108]: spectrum hand-off and stay-and-wait. In the spectrum hand-off mode, the user switches to another channel whereas in the stay-and-wait mode, the user stops transmission until the channel quality improves. Also, long-term impact of the proposed scheme on network performance was analyzed in terms of throughput, delay, packet dropping rate, and packet error rate. TACL is also used in [141] where the authors studies base station switching in an area subject to dynamic traffic loads. Similarly, in [102], the authors proposed an online learning-based cell load approximation scheme in a wireless network.

F. Traffic Classification and Mobility Prediction

Traffic engineering with traffic classification as well as traffic (and hence mobility) prediction would be required to accommodate the time-varying IoT traffic [82]. The authors in [85] proposed a methodology for construction of the model for normal behaviour for ongoing network traffic. Authors proposed traffic prediction procedure by separating short term (non-periodical or temporal)and longer-term(hour or day)variations. A longitudinal model of network connections was developed which can be used to identify deviations in the network traffic patterns. In [142], the authors analyzed and extracted features from packet headers for various real time network application flows, using time window method. Authors used three kinds of ML algorithms; support vector machine (SVM), back propagation (BP) neural network and BP neural network (optimized using particle swarm optimization (PSO), for training and identification of network traffic,respectively. The experimental results showed superiority of SVM algorithm, as traffic identification based on SVM not only showed 98 percent accuracy but also efficiently generate classifier model. In [83], a deep learning-based multimodel integration framework was presented for traffic forecast, with large variations and uncertainties of traffic flows. The framework can dynamically choose an optimal model (from set of candidate models), to extract the implicit relationships hidden in the traffic flow data and perform traffic forecast. In [129], the authors used DL for channel estimation and signal detection in OFDM systems. Unlike existing OFDM receivers that estimate CSI explicitly and afterwards detect/recover the transmitted symbols (using the estimated CSI), proposed DL-based approach estimates CSI implicitly and recovers the transmitted symbols directly. This DL-based method is more robust than conventional methods as fewer training pilots are used and cyclic prefix (CP) is also omitted. Authors trained the DL model offline using simulated data (based on channel statistics) and then used for recovering the online transmitted data directly. This method performs channel estimation and signal detection despite the existence of complicated channel distortion and interference. In [48], the authors surveyed the

flow-based statistical features of P2P traffic. They employed an SVM model to develop a system for classification of the encrypted BitTorrent traffic from usual Internet traffic.

The authors in [143] proposed a learning-based network path planning model. They considered the sequence of nodes in a network path and determined implicit forwarding paths based on empirical data of the network traffic. The authors employed neural networks for capturing the data characteristics of the traffic forwarding and routing path. In [144], the authors compared the performance of six widely-used supervised ML algorithms (Na*i*ve Bayes, Bayes Net, Random Forest, Decision Tree, and Multilayer Perceptron) for classifying network traffic. The authors concluded that Random Forest and Decision Tree algorithms are the best classifiers for network traffic classifications, when classification accuracy and computational efficiency are required.

G. Resource Allocation in Mobile IoT Networks

Cellular networks, vehicular networks, mobile and ad hoc networks are examples of mobile IoT networks where the communicating devices are mobile. Various ML techniques have been proposed for modern cellular networks including D2D communication, massive MIMO and mmWave communications technologies as well as heterogeneous cellular networks (HetNets). *We will provide a survey of these techniques in the next section.* For Vehicular Ad-hoc NETworks (VANETs), the traditional approaches for resource allocation include dynamic programming [145], [146], Semi-Markov Decision Process (SMDP) [147], Maximum Weighted Independent Set problem with Dependent Weights (MWIS-DW) [148], and users' aggregate utility maximization [149]. Recently, ML algorithms have been used to address the resource optimization problem in vehicular networks. For instance, an ML algorithm was applied in [150] to reduce the network overhead in a delay-tolerant routing system where repeated copies of a message signal are discarded.

In the same context, over the last decade, vehicular networks have evolved to rich service and application space, for instance, vehicular clouds. In vehicular clouds, vehicles can form a cloud, can use the services of the cloud, or the combination of the two [151]. There are plethora of services realized through vehicular clouds such as traffic information as a service, cooperation as a service, storage as a service, and vehicle witness as a service [152]–[154]. In such scenarios, resource management is of paramount importance for efficient service provisioning. For this purpose, both ML and DL techniques have been proposed to allocate and manage resources in vehicular clouds. In [155], the authors explored the applications of ML and DL in managing the resources of vehicular networks. Arkian *et al.* [156] proposed a resource allocation mechanism for vehicular networks using cluster-based Q-learning approach. The authors aim at efficient cluster-head selection mechanism through employing fuzzy logic. The essence of the resource management scheme is to select the vehicle with enough resources to offer, in an efficient and QoS-aware way. In another work, Salahuddin *et al.* [157] proposed an RL-based technique for resource provisioning in vehicular clouds. The

specific feature of the proposed RL-based resource management is that it provisions the resources based on a long-term benefit, which on one hand increases the utility, and on the other hand decreases the computational overhead.

VI. MACHINE LEARNING TECHNIQUES FOR EMERGING CELLULAR IoT NETWORKS

In this section, we survey the existing ML- and DL-based resource management solutions for specific cellular IoT technologies including HetNets, D2D, MIMO (massive MIMO), and NOMA.

A. ML Techniques in Heterogeneous Networks (HetNets)

In a HetNet, cellular users coexist with small cell users using different Radio Access Technologies (RATs). The major issues related to resource management include mitigation of cross- and co-tier interference, mobility management, user association, RAT selection, and self organization. In addition to the conventional methods (e.g., based on optimization or heuristics), ML-based methods have been used for resource management in HetNets [158]. For instance, [159] presented a mobility management method for a HetNet using RL, where inter-cell coordination and handovers are learnt by the devices to efficiently manage network resources. Similarly, [160] reduced the energy consumption of the networks using a fuzzy game-theory approach where decision rules are learnt over time for better power consumption while maintaining a minimum QoS level for the users. The authors in [161] proposed a novel framework for cognitive RAT selection by the terminal devices in a HetNet using ML techniques.

Recently, ML techniques have also been used to provide self-organization (which includes self-configuration, self-organization, and self-healing) in the networks. In [162], the authors studied an ANN approach to deal with self-optimization in HetNets. In [163], the authors proposed an efficient resource allocation algorithm for QoS provisioning in terms of high data rates that uses an online learning algorithm and Q-value theory to optimize the resources. A similar work on resource allocation in heterogeneous cognitive radio networks can be found in [139].

B. ML Techniques in Device-to-Device Communications

In D2D networks, two devices in proximity connect to each other bypassing the central base station. By performing proximity communications, a D2D network offloads the traffic from the main BS and also improves the spectral efficiency as well as energy efficiency (EE) of the network [164]. A low transmit power between the radios ensures EE, whereas low path-loss ensures high spectral efficiency and sum rate. There is a considerable amount of literature addressing various issues of D2D networks including resource and power allocation, mode selection, proximity detection, interference avoidance [165], [166]. Recently, ML has been used to address different problems related to D2D communications such as caching [167], security, and privacy [168] and so on.

A major challenge in D2D networks is to efficiently utilize the limited resources to provide QoS requirements to all network entities, including cellular as well as D2D users. In [76], the authors devised a channel access strategy based on bandit learning in a distributed D2D system, where each pair selects an optimal channel for its communications. This does not only reduce the interference among users on the same channels, but also increases the rates of individual D2D pairs. Similarly, another work on channel selection using autonomous learning was presented in [169]. A method for resource allocation in D2D networks was presented in [170] where Q-learning is used to find the optimal resource selection strategy. Similarly, in [171], the authors used cooperative RL that improves the individual throughputs of the devices and the sum rates of the system using optimal resource allocation through a cooperative strategy planning. In [172], power allocations for various D2D pairs were optimally designed using ML to provide energy-efficient network solution.

C. ML in MIMO and Massive MIMO Communications

Mobile nodes connected in a cellular system enjoy very high data rates if the number of antenna elements at the BS is large. Recently, a relatively new concept known as massive MIMO has been an active area of research. Many algorithms have been studied for channel estimation, beamforming, mitigating pilot contamination, and improving energy-efficiency in massive MIMO systems. In this context, ML techniques have also been used to address different problems in massive MIMO systems. For instance, in [104], the authors addressed the channel estimation problem using sparse Bayesian learning algorithm. This algorithm needs good knowledge of the channel. Furthermore, the authors used the channel components in the beam domain as Gaussian Mixture (GM) distribution while Expectation-Maximization (EM) is used to learn the system parameters. The proposed channel estimator shows a good improvement as compared to the conventional methods in the presence of pilot contamination.

The positioning problem in MIMO systems was addressed in [173] using least square SVM. SVM finds the mapping between the sample space and coordinate space based on MIMO channel established by the multipath propagation. Simulations show improved location performance of the SVM as compared to ANN and *K*-Nearest Neighbor (KNN). Furthermore, in [103], the authors recommended a supervised ML technique, i.e., SVM-based technique for MIMO channel learning. Similarly, in [174], the authors proposed a supervised learning for pilot assignment to address the pilot contamination in MIMO systems, considering the output features as pilot assignment while the input features as user locations in the cells. This scheme is implemented using a DNN.

In [175], the authors presented ML algorithms for link adaptation in MIMO systems. Adaptive Modulation and Coding (AMC) is generally used for link adaptation, which is a challenging task in spatial and frequency selective channels. For the selection of suitable modulation scheme and coding rate in MIMO-OFDM systems, the past observations of the CSI and errors can be used. For the selection of transmission parameters, in [175], two techniques based on autoencoder-SVM and autoencoder-softmax are proposed. The autoencoder-SVM

model is used to construct a mapping between CSI and the selection of optimal modulation order to achieve link reliability. The autoencoder is a neural network model to find features from the CSI. On the other hand, the autoencoder-softmax approach uses logical regression models to choose joint modulation and channel coding parameters to optimize the link reliability. Simulation results show improved spectral efficiency compared to the conventional non-learning algorithms. In [176], the authors proposed a supervised learning method that uses the past information of the CSI and errors to predict suitable modulation scheme and coding rate without finding the relation between input and output of wireless transceiver. To increase the accuracy of link adaptation using ML, low dimensional feature set using ordered SNRs was proposed. Simulation results show that the proposed scheme improves the frame error rate and spectral efficiency.

In another work, Mauricio *et al.* [177] presented an ML approach to solve the grouping problem in multi-user massive MIMO systems. The mobile stations are classified into clusters using K-means clustering and then mobile stations are selected from the clusters to make an Space-Division Multiple Access (SDMA) group for capacity maximization. Similarly, in [178], the authors presented a Radio Resource Management (RRM) and hybrid beamforming method for downlink massive MIMO system using ML techniques. The computationally complex resource management in multi-user massive MIMO is replaced with neural network using supervised learning. The results showed that the proposed ML-based method gives the same sum rate but decreases the execution time as compared to the performance of CVX-based method. Furthermore, Wang *et al.* [137] proposed an ML-based beam allocation method for MIMO systems that outperforms the conventional methods.

D. ML Techniques for NOMA

For the next generation wireless networks, NOMA technique is a prime candidate to give spectral efficient performance and massive connectivity. The key idea behind NOMA is to handle multiple users in the same system resource unit [179]. These resource units are spreading code, time slot, sub-carrier, or space [180]. Given the complexity of resource allocation problems in NOMA (i.e., user clustering, power allocation in the presence of imperfect Successive Interference Cancellation (SIC) and jamming attacks), it is of interest to exploit the benefits of learning-based technique for NOMA systems [181].

Notably, various works have recently applied ML and DL techniques for different NOMA problems. For instance, in [182], the authors proposed a novel approach using DL in NOMA system where single BS serves multiple users deployed in a single cell. The authors used a DL-based LSTM network which is integrated with traditional NOMA system. This enables the system to observe and characterize the channel characteristics automatically. In [183], the authors proposed an online learning based novel detection method for uplink NOMA scheme in a single cell. An online adaptive filter is designed for large cluster size of users,

which provides better results than SIC-based detection method. In [184], the authors proposed a MIMO-based anti-jamming game for NOMA transmission. They used a Q-learning-based allocation of power coefficients for transmission in dynamic downlink NOMA to counter jamming attacks. This mechanism improves the NOMA system transmission efficiency in the presence of smart jammers.

In [185], the authors investigated the key problems of mmWave-based NOMA system, i.e., power allocation to the users and clustering of users. They derived an efficient user clustering for NOMA system. Considering continuous arrivals of users in the system, an online user clustering algorithm on the basis of K-means clustering was proposed which greatly reduces the computational complexity of the NOMA system. Similarly, in [186], the authors proposed a NOMA-based resource management approach in a multi-cell network. Furthermore, in [187], the authors investigated the resource sharing problem for downlink transmission by considering maximum power constraint in NOMA networks. Moreover, in [188], the authors dealt with the resource allocation problem in a NOMA-based heterogeneous IoT network and proposed a deep RNN-based algorithm for optimal solution. The optimization is performed with respect to number of users in the system and the number of sub-carriers allocated to each of them to improve the spectral efficiency. The results showed that the proposed DL-based scheme outperforms the traditional schemes in terms of connectivity scale and spectral efficiency.

A comprehensive summary of the ML-based techniques for various cellular IoT networks and the problems solved through ML is given in Table V.

E. Lessons Learned

Resource management in IoT networks is a fundamental problem that becomes increasingly complex with the consideration of advanced wireless technologies as well as system uncertainties (e.g., due to dynamic channel and traffic variations, device mobility and multi-dimensional QoS requirements) and large-scale nature of the system (e.g., with thousands of devices and network nodes). Resource management problems are generally posed as optimization problems with one or more objective functions and some constraints. However, in many cases the conventional optimization problems are not solvable with reasonable complexity, unless the problems are simplified and/or the constraints are relaxed. Also, when the problems are non-convex, they often provide only sub-optimal solutions. To reduce the complexity of obtaining resource allocation solutions, new methods will be required. Also, in many cases the resource allocation problems cannot be modeled analytically due to the lack of appropriate modeling tools. Therefore, we note that the conventional resource management solutions do not scale well in the massively heterogeneous networks such as IoT networks.

Another common approach for resource management in distributed IoT networks is the use of game-theoretic approaches. However, modeling and analysis of the game models for large-scale heterogeneous IoT networks under practical system

TABLE V
CELLULAR IoT NETWORK AND MACHINE LEARNING TECHNIQUES APPLIED

Type of Network	Machine Learning Techniques	Considered Problems
Device-to-Device (D2D) Communications	• Bandit Learning [76] • Autonomous Learning [169] • Q-Learning [170] • Cooperative RL [171], [172]	Channel access strategy Channel/resource selection Cooperative strategy planning Power allocations
Heterogeneous Networks	• RL [159], [161] • Fuzzy Game Theory [160] • ANN [162] • Q-Learning [163], [139]	Inter-cell interference coordination Handoff management Maximization of energy efficiency Self-organization QoS provisioning
Multiple Input Multiple Output (MIMO) and Massive MIMO	• Bayesian Learning [104] • SVM [103], [137], [178] • DNN [174] • Autoencoder-SVM [175] • K-means [177]	Pilot contamination BS positioning Channel learning Pilot assignment Link adaptation Multi-user grouping
Non-Orthogonal Multiple Access (NOMA)	• DL-based LSTM [182] • Q-Learning [183], [184] • K-means [185], [186] • RNN [187], [188]	Channel esstimation User clustering Anti-jamming mmWave-based NOMA

constraints (e.g., uncertainty in terms of the number of devices and network characteristics, lack of availability of information about other devices) is extremely challenging. Also, even if the equilibrium solutions can be characterized, obtaining those solutions instantly would be a very highly ambitious goal due to slow convergence rate.

In the above context, increasing interest in the interdisciplinary research has led to the use of ML, DL, RL, and related tools to make IoT devices and networks smarter and intelligent, and adaptive to their dynamic environments. Decrease in the computational costs, improved computing power, availability of unlimited solid-state storage, and integration of various technological breakthroughs have made this possible. Applications of ML and DL techniques enable IoT devices to perform complex sensing and recognition tasks and enable the realization of new applications and services considering the real-time interactions among humans, smart devices and the surrounding environment. DL, RL, and DRL algorithms can be leveraged for automated extraction of complex features from large amounts of high dimensional unsupervised data produced by IoT devices. Furthermore, the inception of future generation Internet technologies such as 5G and beyond, it further necessitates to embed the intelligence to networks. The futuristic real-time applications will be highly resource-hungry and will need viable, scalable, ultra-fast, and efficient resource management solutions. The current ML and DL mechanisms can be a good starting point to deal with network size, complexity, context-awareness, and heterogeneity from resource management standpoint.

ML techniques can support self-organizing operations by learning and processing statistical information from the environment. These learning techniques are inherently distributed and are scalable for dense and large-scale IoT networks. Data-driven ML techniques are used to detect possible patterns and similarities in large data sets and can make predictions. ML is leveraged to develop algorithms that are capable of applying statistical methods on input data, performing analysis and predicting an output value without being explicitly programmed. Therefore, ML techniques require datasets to learn from and then the learned model is applied to the real data. However, the learned model may not generalize well to the entire range of features and properties of the data. In this regard, DL techniques have been employed to address some of the limitations of the ML techniques. Nonetheless, ML and DL algorithms have inherent limitations and also the development of these techniques brings along new challenges that must be addressed. In particular, data-driven ML and DL techniques require "good quality" data which may not be available and also difficult to generate. For example, for a DL-based resource management method for channel and power allocation problem in a large-scale cellular IoT network, for a given network state (e.g., number of users, their association to the BS, channel gains to the respective BS), the optimal channel and power allocation data will be required to train the DNN. Generation of these data will be computationally too expensive even using techniques such as genetic algorithm(s). Therefore, the challenges related to data generation will need to be addressed to develop DL-based resource management schemes.

Furthermore, in case of cellular IoT, the different generations of cellular networks (specially 5G and beyond) are primarily based on AI and ML. Therefore, the role of ML in the resource management of cellular IoT is definitive. From our discussion on the ML and DL-based solutions for cellular IoT networks (both traditional and emerging technologies), we note that despite the ML and DL-based solutions for resource management in cellular IoT, concrete steps need to be taken to catch up with the advancements in the communication technologies and the futuristic real-time applications. It is also worth mentioning that the type of applications and their criticality towards the human lives render the current DL techniques (which are mostly black-boxes) insufficient.

In other words, most of the ML and DL techniques work in a way that it is extremely difficult to understand the inner working principles of these methods and we have to rely merely on their output. These phenomena might not be acceptable in critical applications where human lives could be in jeopardy, such as resource management for health-care IoT application or critical industrial IoT application. The certainty and accuracy of the ML and DL mechanisms used in the aforementioned application scenarios is of paramount importance. However, in essence, most of the ML and DL mechanism are, at least in part, black boxes and do not give much information about the inner working principles of the mechanisms. Therefore, they have to be blindly followed and trusted. This phenomenon, should something go wrong, will have catastrophic consequences to the critical applications in every aspect (including resource management). Therefore, we note that there are broadly two ways to address these challenges. First is to focus on the security and credibility of these DL mechanisms. Security of the underlying ML and DL mechanisms is essential in the presence of adversarial environments. Furthermore, the privacy of data must also be taken into account. Therefore, new breeds of ML and DL mechanisms that are privacy-friendly, must be employed such as distributed and federated learning. Second way is to make the ML and DL mechanism more transparent to understand their working principles. This opens door for further research in the direction of transparent DL or explainable ML and DL mechanisms. As a result of these lessons, we discuss future challenges in more details and identify research directions in the next section.

VII. FUTURE RESEARCH CHALLENGES

In this section, we will discuss various challenges for the utilization of ML/DL techniques for resource management in IoT networks. Here we will focus on the challenges related to the development, (re)training, interoperability, cost, generalization, and the data-related issues of the ML models.

A. Challenges Associated With MLand DL Models for IoT Networks

1) Development of ML Models for IoT Networks: It is challenging to develop a suitable model to process data coming from a range of diverse IoT devices. Similarly, effective input data labeling is also a cumbersome task. For instance, in case of health-care applications, the training model must be pin-point accurate to derive the right features from the data and learn effectively, because false positives could have dire consequences in such models. Reliability of the ML models will be essential for mission-critical IoT applications which will be vulnerable to the anomalies (e.g., misclassifications) created because of ML or DL algorithms. Additionally, RL suffers from instability and divergence when all possible action-reward pairs are not used in a given scenario [117]. In a nutshell, depending on the use-case, the underlying IoT technology and the characteristics unique to that use-case, ML models must be efficiently crafted to meet the requirements of the use-cases. For the deployment of ML and DL models in

the resource-constrained IoT devices, it is essential to reduce the processing and storage overhead [189].

Again, engineering of the ML models (e.g., optimizing the *hyperparameters* in a DNN such as the number of hidden layers, number of neurons per layer, optimizing the learning rate in a DRL model) are challenging.

2) Lack of Multitasking Capabilities: DL models are highly specialized and trained in a specific application domain. These models are only effective to solve one specific type of the problem for which it is trained. Whenever there is a change in the definition, state, or the nature of the problem (even if the new problem is very similar to the older one), re-assessment and retraining of the model is essential. Restructuring the architecture of the entire model is required when multitasking is done by the DL algorithms. For real-time IoT applications, it is difficult for the IoT devices to change and retain the DL models with frequently changing input information in a timely manner. Therefore, DL models capable of multitasking should be further investigated.

3) Retraining of ML and DL Models: When ML algorithms are applied to IoT applications and resource management in IoT networks, the models need to be updated according to the fresh collected data. This phenomenon gives rise to many important questions such as when and how ML models should be retrained with the changing environment and how to ensure that an appropriate amount of data is collected.

4) Interpretability of ML and DL Models: When applying ML models to make informed decisions in IoT networks, in addition to accurate predictions, it is also helpful for the domain experts to gain insights into the decision making process of the models. In other words, interpretability of the model is very important to generate deep insights into the ML model. Powerful models such as DL and NNs are difficult to understand and interpret. Further research is required to build smart models which are not only highly accurate but also interpretable. The need for explainability of the ML models drives the need for simple models such as decision trees. It will be of fundamental importance to strengthen the existing theoretical foundations of DL with repeated experimental data. In this way, the performances of DL models can be quantified based on certain parameters such as computational complexity, learning efficiency, and accuracy.

5) Cost of Training of ML and DL Models: Sufficient amount of data is required to train ML models and a considerable cost is associated with training these models. This training cost becomes more pressing, specifically in real-time IoT applications. Furthermore, the ML models are needed to be up-to-date in order to have accurate predictions over longer duration of time. Although model retraining can be done manually in the same way as the model was trained, it is a time consuming process. Another method is continuous learning in which an automated system is developed that continuously evaluates and retrains the existing models. Both of these methods require a well architected system model that should take into account not only the initial training procedure of a model but also a plan for keeping the models up-to-date.

6) Generalization of ML Models: Generalization is defined as the ability of an ML model to adapt to new and unseen data.

The data is drawn from the same distribution used for training the model. Accuracy of the model is determined by how well it can be generalized to data points outside the training data sets. The generalization error is the measure of accuracy and the validity of an ML model. The propagation channels in a wireless IoT network can have varying and diverse physical characteristics. Therefore, although trained offline, the ML/DL models for resource management in IoT networks should be generalized to various channel conditions and must have dynamic adaptive capabilities. More research and implementation efforts are required to build practical models for real system with real user workloads.

7) Neural Networks are Unpredictable: Both NNs and DNNs, due to their layered structure, are unpredictable towards achieving the end result. It is believed that these mechanisms mimic the human brain where it is extremely difficult to follow the route through which a conclusion is reached as a result of the unpredictable number of layers. Depending on the input, the interconnection of layers is also complex and dynamic. This uncertainty of layers and the overall process of decision making renders these mechanisms less suitable for critical applications such as health-care and other mission-critical IoT application. In the context of resource management in cellular IoT networks, it is imperative to use these methods keeping in mind the underlying application. Without loss of generality, in this regard, the accuracy of the NN and DNN is essential. There are potentially two methods to address the unpredictability challenge of NNs, i.e., develop ultra-reliable DL models or somehow understand the working principle of these models. The former is more challenging whereas the latter has been the subject of recent research in the form of *Explainable Artificial Intelligence (XAI)*. XAI is a branch of AI (and its descendant ML and DL methods) that break the black-box nature of the inner layers of these model and make it transparent, and thus more reliable (or at least understandable) [190], [191]. In this context, we think that there is already a growing interest in XAI-based solutions for critical cellular IoT applications.

8) DL and Abstraction (How Deep the DL Should Be?): One of the most promising features of DL mechanisms is the layered abstraction while learning. However, depending on the types of applications, the level of depth might be different and context plays a pivotal role in decision-making. In the case of IoT, particularly in resource management, the context of the application is very important for DL algorithms. More precisely, it is important to know which underlying communication technology is used by the cellular IoT networks and thus, respective tweaks should be made to the DL algorithms for resource management. Furthermore, distributed DL [192] could also play an important role to separate different parts of the IoT networks and perform learning on the local data.

9) Reliability of ML and DL Mechanisms: The intelligent and informed decisions made as a result of ML and DL algorithms incorporate intelligence to the applications; however, the reliability of the output by these algorithms is very critical. The ML and DL algorithms, by nature, are very sensitive to change in the data. Even a minor change in the input will have a huge impact on the output. Therefore, it is important to measure the reliability of these algorithms. Recently, the

ML has been used in an adversarial setting to fool the learning mechanism and make the model decide other than what it is trained for [193]. This phenomenon is also referred to as Adversarial Machine Learning (AML). One way to attack the ML and DL methods is to introduce noise (perturbation) to the input of the learning algorithm so that the classification as a result of this learning is affected [194]. The attackers target the learning algorithm through the features of ML. For instance, the attackers manipulate the features that are used in learning and the result is manipulated as well. Therefore, it is essential to make sure that the input and the features used in learning are reliable. In the context of resource management in cellular IoT networks, AML-based attacks could have dire consequences on the performance and functionality of the applications. For instance, in an industrial IoT application, if the resources are not properly managed (as a result of the output of ML and DL algorithms), the quality of service will be affected and thus-forth jeopardizing the whole IoT application. In this vein, we believe that the current ongoing work in explainable and verifiable ML and DL mechanisms will be cornerstones.

B. Data Challenges in ML and DL-Enabled IoT Networks

1) Scarcity of Datasets: Learning algorithms entirely depend on the amount and quality of data to build models. Often, these datasets for IoT networks are scarce. Also, the availability of such datasets is subject to laws through which the data should be collected, the environment, and the consent of the data subject. Generation of synthetic datasets that could train models is also challenging. Another important issue is the uniformity of the datasets that could be used in cross-platform environments.

2) Heterogeneity of IoT Data: Due to the massive and heterogeneous nature of IoT networks, the data generated by such networks has multiple dimensions and ML models must be developed to extract useful information and features from the data. In this regard, pre-treatment of the data (pre-processing, cleansing, and ordering) and fusion of multi-source data will play an important role to make it ready for the respective ML model.

3) Collection of Data: The data generated in IoT networks can be used for a variety of different purposes depending on the application domain. For instance, some data maybe privacy-critical where it must be anonymized before using it for training. The example of such data is the data generated by smart-home environment and by the sensors on the human body. These data must be sent to central server(s) for further processing. Therefore, the context is also essential in the collection of data from different domains. Also, the ML and DL algorithms need to cope with anomalies in the data (e.g., data imbalance due to class skewness and class overlapping) in an intelligent manner.

4) Avalanche Effect of ML and DL: The avalanche effect is a desired characteristic of the cryptographic algorithms where a small change in the input causes huge change in the output. However, this effect is the default property of the ML and DL algorithms which also makes them vulnerable to adversarial

attacks. In the context of IoT, this property of ML and DL mechanism will cause dire consequences to the applications. Therefore, the integrity of the data must be pre-checked before giving the data as input to the ML and DL algorithms. Recent perturbation attacks on NN [195] advocate that concrete steps are required to guarantee the integrity of input data to the ML and DL algorithms. One possible solution would be to take the context of the data into account and incorporate an effective access control mechanism to the input data.

5) Context and Privacy of IoT Data: The use of ML and DL mechanism will totally depend on the underlying use-case, for instance, resource management, however, it is important to look at the context and the scenarios in which resource management is required. For example, the resource management in health-care IoT application might be different from the resource management in smart-home IoT environment and deployed ML and DL mechanism must be adaptive to these scenarios. From another perspective, the data required for ML and DL algorithms might be private or the IoT nodes might not want to share it. In this case, to reap the benefits of ML and DL mechanisms, new advancements in ML and DL have been pursued where ML and DL algorithms are deployed in a distributed fashion where the data is not shared globally and rather processed locally (at the proximity of the network). Furthermore, Federated Learning (FL) can also be a potential candidate to address these challenges because the data is distributed and never leaves the local environment while it is used to train a centralized model [196]. We believe that, this direction of research will like take further pace in future.

VIII. CONCLUSION

The fundamental notion of connectivity among different smart devices with processing, storage, and communication capabilities, collectively constitute an IoT network. IoT has the potential to revolutionize many aspects of our lives ranging from our life-style, business, environment, infrastructure and so on. The heterogeneity introduced by the IoT is humongous from many aspects such as device types, network types, and data types. In the wake of such massive heterogeneous environment, IoT networks need scalable, adaptive, efficient, and fair network management mechanisms. Among other aspects of the network management, resource management has a pivotal role in the successful realization of massively heterogeneous IoT networks. In essence, IoT networks do not only face resource allocation challenges but also require contextual and situational custom resource management decisions. Unlike the traditional resource management methods including optimization and heuristics-based methods, game theoretical and cooperative approaches, the ML and DL models can derive actions from the run-time context information and can re-tune and re-train themselves with changes in the environment. ML and DL techniques are promising for automatic resource management and decision making process, specifically for large-scale, complex, distributed and dynamically changing IoT application environment.

In this article, we have discussed the limitations of the traditional resource management techniques for wireless IoT networks. We have carried out a comprehensive survey of the ML- and DL-based resource management techniques in the IoT networks with enabling technologies such as HetNets, D2D, MIMO/massive MIMO, and NOMA. Furthermore, we have also covered different aspects of the resource management such as scheduling and duty cycling, resource allocation (in static and mobile IoT networks), clustering and data aggregation, resource discovery through spectrum sensing and cell selection, traffic classification and prediction, power allocation, and interference management. To this end, we have identified the future research challenges in applying ML and DL in IoT resource management. Applications of machine learning techniques, specifically deep reinforcement learning techniques, to solve complex radio resource management problems in emerging IoT networks is a fertile area of research. For these networks, we will need to carefully craft the solutions according to the challenges such as model uncertainty, interpretability of the models, cost of model training, and generalization from test workloads to real application user workloads.

REFERENCES

[1] F. Hussain, *Internet of Things; Building Blocks and Business Models*. Cham, Switzerland: Springer, 2017.
[2] D. B. J. Sen, "Internet of Things—Applications and challenges in technology and standardization," *Wireless Pers. Commun.*, vol. 58, pp. 49–69, Apr. 2011.
[3] U. S. Shanthamallu, A. Spanias, C. Tepedelenlioglu, and M. Stanley, "A brief survey of machine learning methods and their sensor and IoT applications," in *Proc. IEEE Conf. Inf. Intell. Syst. Appl.*, Mar. 2018, pp. 1–8.
[4] Y. Li, Z. Gao, L. Huang, X. Du, and M. Guizani, *Resource Management for Future Mobile Networks: Architecture and Technologies*. Hoboken, NJ, USA: Elsevier, 2018, pp. 392–398.
[5] X. Li, J. Fang, W. Cheng, H. Duan, Z. Chen, and H. Li, "Intelligent power control for spectrum sharing in cognitive radios: A deep reinforcement learning approach," *IEEE Access*, vol. 6, pp. 25463–25473, 2018.
[6] A. Al-Fuqaha, M. Guizani, M. Mohammadi, M. Aledhari, and M. Ayyash, "Internet of Things: A survey on enabling technologies, protocols, and applications," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 4, pp. 2347–2376, 4th Quart., 2015.
[7] V. Gazis, "A survey of standards for machine-to-machine and the Internet of Things," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 1, pp. 482–511, 1st Quart., 2017.
[8] E. Ahmed, I. Yaqoob, A. Gani, M. Imran, and M. Guizani, "Internet-of-Things-based smart environments: State of the art, taxonomy, and open research challenges," *IEEE Wireless Commun.*, vol. 23, no. 5, pp. 10–16, Oct. 2016.
[9] S. B. Baker, W. Xiang, and I. Atkinson, "Internet of Things for smart healthcare: Technologies, challenges, and opportunities," *IEEE Access*, vol. 5, pp. 26521–26544, 2017.
[10] F. Javed, M. K. Afzal, M. Sharif, and B. Kim, "Internet of Things (IoT) operating systems support, networking technologies, applications, and challenges: A comparative review," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 3, pp. 2062–2100, 3rd Quart., 2018.
[11] D. E. Kouicem, A. Bouabdallah, and H. Lakhlef, "Internet of Things security: A top-down survey," *Comput. Netw.*, vol. 141, pp. 199–221, Aug. 2018. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1389128618301208
[12] A. Chowdhury and S. A. Raut, "A survey study on Internet of Things resource management," *J. Netw. Comput. Appl.*, vol. 120, pp. 42–60, Oct. 2018. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1084804518302315
[13] D. Airehrour, J. Gutierrez, and S. K. Ray, "Secure routing for Internet of Things: A survey," *J. Netw. Comput. Appl.*, vol. 66, pp. 198–213, May 2016. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1084804516300133

[14] J. Guo, I.-R. Chen, and J. J. P. Tsai, "A survey of trust computation models for service management in Internet of Things systems," *Comput. Commun.*, vol. 97, pp. 1–14, Jan. 2017. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0140366416304959

[15] W. Ding, X. Jing, Z. Yan, and L. T. Yang, "A survey on data fusion in Internet of Things: Towards secure and privacy-preserving fusion," *Inf. Fusion*, vol. 51, pp. 129–144, Nov. 2019. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1566253518304731

[16] M. Mohammadi, A. Al-Fuqaha, S. Sorour, and M. Guizani, "Deep learning for IoT big data and streaming analytics: A survey," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 4, pp. 2923–2960, 4th Quart., 2018.

[17] M. Chen, U. Challita, W. Saad, C. Yin, and M. Debbah, "Artificial neural networks-based machine learning for wireless networks: A tutorial," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 4, pp. 3039–3071, 4th Quart., 2019.

[18] Y. Sun, M. Peng, Y. Zhou, Y. Huang, and S. Mao, "Application of machine learning in wireless networks: Key techniques and open issues," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 4, pp. 3072–3108, 4th Quart., 2019.

[19] S. K. Sharma and X. Wang, "Towards massive machine type communications in ultra-dense cellular IoT networks: Current issues and machine learning-assisted solutions," *IEEE Commun. Surveys Tuts.*, to be published.

[20] K. Gai and M. Qiu, "Optimal resource allocation using reinforcement learning for IoT content-centric services," *Appl. Soft Comput.*, vol. 70, pp. 12–21, Sep. 2018.

[21] R. Boutaba *et al.*, "A comprehensive survey on machine learning for networking: Evolution, applications and research opportunities," *J. Internet Services Appl.*, vol. 9, no. 1, p. 16, 2018.

[22] M. Wang, Y. Cui, X. Wang, S. Xiao, and J. Jiang, "Machine learning for networking: Workflow, advances and opportunities," *IEEE Netw.*, vol. 32, no. 2, pp. 92–99, Mar./Apr. 2018.

[23] O. Simeone, "A very brief introduction to machine learning with applications to communication systems," *IEEE Trans. Cogn. Commun. Netw.*, vol. 4, no. 4, pp. 648–664, Dec. 2018.

[24] K. Doppler, M. Rinne, C. Wijting, C. B. Ribeiro, and K. Hugl, "Device-to-device communication as an underlay to LTE-advanced networks," *Commun. Mag.*, vol. 47, no. 12, pp. 42–49, Dec. 2009, doi: 10.1109/MCOM.2009.5350367.

[25] W. Wu, W. Xiang, Y. Zhang, K. Zheng, and W. Wang, "Performance analysis of device-to-device communications underlaying cellular networks," *Telecommun. Syst.*, vol. 60, no. 1, pp. 29–41, Sep. 2015, doi: 10.1007/s11235-014-9919-y.

[26] C.-H. Yu, K. Doppler, C. B. Ribeiro, and O. Tirkkonen, "Resource sharing optimization for device-to-device communication underlaying cellular networks," *IEEE Trans. Wireless Commun.*, vol. 10, no. 8, pp. 2752–2763, Aug. 2011.

[27] Y. L. Lee, T. C. Chuah, J. Loo, and A. Vinel, "Recent advances in radio resource management for heterogeneous LTE/LTE-A networks," *IEEE Commun. Surveys Tuts.*, vol. 16, no. 4, pp. 2142–2180, 4th Quart., 2014.

[28] *Radio Resource Management in Heterogeneous Networks, Functional Models and Implementation Requirements*, STAN Nat., Sydney, NSW, Australia, 2015.

[29] S. Tarapiah, K. Aziz, and S. Atalla, "Common radio resource management algorithms in heterogeneous wireless networks with KPI analysis," *Int. J. Adv. Comput. Sci. Appl.*, vol. 6, no. 10, pp. 53–58, 2015, doi: 10.14569/IJACSA.2015.061007.

[30] M. Al-Imari, P. Xiao, and M. A. Imran, "Receiver and resource allocation optimization for uplink noma in 5G wireless networks," in *Proc. Int. Symp. Wireless Commun. Syst. (ISWCS)*, Aug. 2015, pp. 151–155.

[31] L. Song, Y. Li, Z. Ding, and H. V. Poor, "Resource management in non-orthogonal multiple access networks for 5G and beyond," *IEEE Netw.*, vol. 31, no. 4, pp. 8–14, Jul. 2017.

[32] U. Raza, P. Kulkarni, and M. Sooriyabandara, "Low power wide area networks: An overview," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 2, pp. 855–873, 2nd Quart., 2017.

[33] K. Mekki, E. Bajic, F. Chaxel, and F. Meyer, "A comparative study of LPWAN technologies for large-scale IoT deployment," *ICT Exp.*, vol. 5, no. 1, pp. 1–7, 2019. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S2405959517302953

[34] A. Mahmood, E. Sisinni, L. Guntupalli, R. Rondón, S. A. Hassan, and M. Gidlund, "Scalability analysis of a LoRa network under imperfect orthogonality," *IEEE Trans. Ind. Informat.*, vol. 15, no. 3, pp. 1425–1436, Mar. 2019.

[35] R. Zhang, Y.-C. Liang, and S. Cui, "Dynamic resource allocation in cognitive radio networks," *IEEE Signal Process. Mag.*, vol. 27, no. 3, pp. 102–114, May 2010.

[36] N. Nie and C. Comaniciu, "Adaptive channel allocation spectrum etiquette for cognitive radio networks," *Mobile Netw. Appl.*, vol. 11, no. 6, pp. 779–797, Dec. 2006. [Online]. Available: https://doi.org/10.1007/s11036-006-0049-y

[37] A. Alnahdi and S.-H. Liu, "Mobile Internet of Things (MIoT) and its applications for smart environments: A positional overview," in *Proc. IEEE Int. Congr. Internet Things (ICIOT)*, Jun. 2017, pp. 151–154.

[38] K. Nahrstedt, H. Li, P. Nguyen, S. Chang, and L. Vu, "Internet of mobile things: Mobility-driven challenges, designs and implementations," in *Proc. IEEE 1st Int. Conf. Internet Things Design Implement. (IoTDI)*, Apr. 2016, pp. 25–36.

[39] I. Tcarenko *et al.*, "Smart energy efficient gateway for Internet of mobile things," in *Proc. 14th IEEE Annu. Consum. Commun. Netw. Conf. (CCNC)*, Jan. 2017, pp. 1016–1017.

[40] M. Haenggi, J. G. Andrews, F. Baccelli, O. Dousse, and M. Franceschetti, "Stochastic geometry and random graphs for the analysis and design of wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 27, no. 7, pp. 1029–1046, Sep. 2009, doi: 10.1109/JSAC.2009.090902.

[41] M. Chen, W. Saad, and C. Yin, "Echo state networks for self-organizing resource allocation in LTE-U with uplink–downlink decoupling," *IEEE Trans. Wireless Commun.*, vol. 16, no. 1, pp. 3–16, Jan. 2017.

[42] Q. Zheng, K. Zheng, P. Chatzimisios, and F. Liu, "Joint optimization of link scheduling and resource allocation in cooperative vehicular networks," *EURASIP J. Wireless Commun. Netw.*, vol. 2015, no. 1, p. 170, Jun. 2015, doi: 10.1186/s13638-015-0392-4.

[43] L. Liang, S. Xie, G. Y. Li, Z. J. Ding, and X. Yu, "Graph-based radio resource management for vehicular networks," in *Proc. IEEE Int. Conf. Commun. (ICC)*, 2018, pp. 1–6.

[44] E. Castañeda, A. Silva, A. Gameiro, and M. Kountouris, "An overview on resource allocation techniques for multi-user MIMO systems," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 1, pp. 239–284, 1st Quart., 2017.

[45] K. Zheng, L. Zhao, J. Mei, B. Shao, W. Xiang, and L. Hanzo, "Survey of large-scale MIMO systems," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 3, pp. 1738–1760, 3rd Quart., 2015.

[46] T. O. Olwal, K. Djouani, and A. M. Kurien, "A survey of resource management toward 5G radio access networks," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 3, pp. 1656–1686, 3rd Quart., 2016.

[47] M. Qiao, Y. Ma, Y. Bian, and J. Liu, *Real-Time Multi-Application Network Traffic Identification Based on Machine Learning*, vol. 92. Cham, Switzerland: Springer, 2015, pp. 473–480.

[48] T. S. Tabatabaei, M. Adel, F. Karray, and M. Kamel, *Machine Learning-Based Classification of Encrypted Internet Traffic*. Heidelberg, Germany: Springer, 2017, pp. 578–592.

[49] R. Masiero, G. Quer, D. Munaretto, M. Rossi, J. Widmer, and M. Zorzi, "Data acquisition through joint compressive sensing and principal component analysis," in *Proc. IEEE Glob. Telecommun. Conf. (GLOBECOM)*, Nov. 2009, pp. 1–6.

[50] M. S. Jamal, V. S. Keerthy, H. Ochiai, H. Esaki, and K. Kataoka, "INSTRUCT: A clustering based identification of valid communications in IoT networks," in *Proc. 5th Int. Conf. Internet Things Syst. Manag. Security*, Oct. 2018, pp. 228–233.

[51] R. Han, F. Zhang, and Z. Wang, "AccurateML: Information-aggregation-based approximate processing for fast and accurate machine learning on MapReduce," in *Proc. IEEE Conf. Comput. Commun. (IEEE INFOCOM)*, May 2017, pp. 1–9.

[52] F. Hussain, A. Anpalagan, A. S. Khwaja, and M. Naeem, "Resource allocation and congestion control in clustered M2M communication using *Q*-learning," *Trans. Emerg. Telecommun. Technol.*, vol. 28, no. 4, 2017, Art. no. e3039.

[53] B. Frankovič and B. Budinska, "Advantages and disadvantages of heuristic and multi agents approaches to the solution of scheduling problem," *IFAC Proc. Vol.*, vol. 33, no. 13, May 2000, pp. 367–372.

[54] J. Moura and D. Hutchison, "Game theory for multi-access edge computing: Survey, use cases, and future trends," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 1, pp. 260–288, 1st Quart., 2019.

[55] S. Ali, A. Ferdowsi, W. Saad, and N. Rajatheva. (Oct. 2018). *Sleeping Multi-Armed Bandit Learning for Fast Uplink Grant Allocation in Machine Type Communications*. [Online]. Available: https://arxiv.org/abs/1810.12983

[56] K. Kanoun, C. Tekin, D. Atienza, and M. van der Schaar, "Big-data streaming applications scheduling based on staged multi-armed bandits," *IEEE Trans. Comput.*, vol. 65, no. 12, pp. 1–15, Nov. 2016.

[57] M. Mohammadi and A. Al-Fuqaha, "Enabling cognitive smart cities using big data and machine learning: Approaches and challenges," *IEEE Commun. Mag.*, vol. 56, no. 2, pp. 94–101, Feb. 2018.

[58] J. Oueis, E. C. Strinati, and S. Barbarossa, "The fog balancing: Load distribution for small cell cloud computing," in *Proc. IEEE VTC*, May 2015, pp. 1–6.

[59] M. Kim and I.-Y. Ko, "An efficient resource allocation approach based on a genetic algorithm for composite services in IoT environments," in *Proc. IEEE Int. Conf. Web Services*, Oct. 2018, pp. 136–142.

[60] P. Semasinghe, S. Maghsudi, and E. Hossain, "Game theoretic mechanisms for resource management in massive wireless IoT systems," *IEEE Commun. Mag.*, vol. 55, no. 2, pp. 121–127, Feb. 2017.

[61] H. Kharrufa, H. Al-Kashoash, and A. H. Kemp, "A game theoretic optimization of RPL for mobile Internet of Things applications," *IEEE Sensors J.*, vol. 18, no. 6, pp. 2520–2530, Mar. 2018.

[62] H. Zhang, Y. Zhang, Y. Gu, D. Niyato, and Z. Han, "A hierarchical game framework for resource management in fog computing," *IEEE Commun. Mag.*, vol. 55, no. 8, pp. 52–57, Aug. 2017.

[63] Z. Zhou, M. Dong, K. Ota, G. Wang, and L. T. Yang, "Energy-efficient resource allocation for D2D communications underlaying cloud-RAN-based LTE-A networks," *IEEE Internet Things J.*, vol. 3, no. 3, pp. 428–438, Jun. 2016.

[64] H. Zhang, Y. Xiao, S. Bu, D. Niyato, F. R. Yu, and Z. Han, "Computing resource allocation in three-tier IoT fog networks: A joint optimization approach combining Stackelberg game and matching," *IEEE Internet Things J.*, vol. 4, no. 5, pp. 1204–1215, Oct. 2017.

[65] S. F. Abedin, M. G. R. Alam, S. A. Kazmi, N. H. Tran, D. Niyato, and C. S. Hong, "Resource allocation for ultra-reliable and enhanced mobile broadband IoT applications in fog network," *IEEE Trans. Commun.*, vol. 67, no. 1, pp. 489–502, Jan. 2019.

[66] F. Hussain, R. Hussain, S. A. Hassan, and E. Hossain, "Machine learning in IoT security: Current solutions and future challenges," *ArXiv*, vol. abs/1904.05735, 2019.

[67] R. H. Javed, A. Siddique, R. Hafiz, O. Hasan, and M. Shafique, "ApproxCT: Approximate clustering techniques for energy efficient computer vision in cyber-physical systems," in *Proc. 12th Int. Conf. Open Source Syst. Technol. (ICOSST)*, Dec. 2018, pp. 64–70.

[68] S. Gupta, A. Thakral, and S. Sharma, "Novel technique for prediction analysis using normalization for an improvement in *k*-means clustering," in *Proc. IEEE Int. Conf. Inf. Technol. Next Gener. IT Summit Theme Internet Things Connect Worlds (InCITe)*, 2016, pp. 32–36.

[69] V. V. Kumari and P. R. K. Varma, "A semi-supervised intrusion detection system using active learning SVM and fuzzy C-means clustering," in *Proc. IEEE Int. Conf. IoT Soc. Mobile Anal. Cloud (I-SMAC)*, 2017, pp. 481–485.

[70] C.-M. Chung, C.-C. Chen, W.-P. Shih, T.-E. Lin, R.-J. Yeh, and I. Wang, "Automated machine learning for Internet of Things," in *Proc. IEEE Int. Conf. Consum. Electron. Taiwan (ICCE-TW)*, 2017, pp. 295–296.

[71] M. A. Khan, A. Khan, M. N. Khan, and S. Anwar, "A novel learning method to classify data streams in the Internet of Things," in *Proc. IEEE Nat. Softw. Eng. Conf.*, 2014, pp. 61–66.

[72] C. Zhang, P. Patras, and H. Haddadi, "Deep learning in mobile and wireless networking: A survey," *IEEE Commun. Surveys Tuts.*, vol. 1, no. 3, pp. 1–43, 3rd Quart., 2019.

[73] S. Park, M. Sohn, H. Jin, and H. Lee, "Situation reasoning framework for the Internet of Things environments using deep learning results," in *Proc. IEEE Int. Conf. Knowl. Eng. Appl. (ICKEA)*, 2016, pp. 133–138.

[74] J. Zhu, Y. Song, D. Jiang, and H. Song, "A new deep-*Q*-learning-based transmission scheduling mechanism for the cognitive Internet of Things," *IEEE Internet Things J.*, vol. 5, no. 4, pp. 2375–2385, Apr. 2018.

[75] N. Chauhan, N. Choudhary, and K. George, "A comparison of reinforcement learning based approaches to appliance scheduling," in *Proc. 2nd Int. Conf. Contemporary Comput. Informat. (IC3I)*, Dec. 2016, pp. 253–258.

[76] S. Maghsudi and S. Stanczak, "Channel selection for network-assisted D2D communication via no-regret bandit learning with calibrated forecasting," *IEEE Trans. Wireless Commun.*, vol. 14, no. 3, pp. 1309–1322, Mar. 2015.

[77] S. Chinchali *et al.*, "Cellular network traffic scheduling with deep reinforcement learning," in *Proc. AAAI*, 2018, pp. 766–774.

[78] H. Mao, M. Alizadeh, I. Menache, and S. Kandula, "Resource management with deep reinforcement learning," in *Proc. 15th ACM Workshop Hot Topics Netw. (HotNets)*, 2016, pp. 50–56. [Online]. Available: http://doi.acm.org/10.1145/3005745.3005750

[79] G. Rjoub and J. Bentahar, "Cloud task scheduling based on swarm intelligence and machine learning," in *Proc. IEEE 5th Int. Conf. Future Internet Things Cloud (FiCloud)*, Aug. 2017, pp. 272–279.

[80] S. Bohez, T. Verbelen, P. Simoens, and B. Dhoedt, "Discrete-event simulation for efficient and stable resource allocation in collaborative mobile cloudlets," *Simulat. Model. Pract. Theory*, vol. 50, pp. 109–129, Jul. 2015. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1569190X14000768

[81] J. S. Jang, Y. L. Kim, and J. H. Park, "A study on the optimization of the uplink period using machine learning in the future IoT network," in *Proc. IEEE Int. Conf. Pervasive Comput. Commun. Workshops (PerCom Workshops)*, Mar. 2016, pp. 1–3.

[82] T. Otoshi *et al.*, "Traffic prediction for dynamic traffic engineering," *Int. J. Comput. Telecommun. Netw.*, vol. 85, Jul. 2015, pp. 34–60.

[83] T. Zhou, G. Han, X. Xu, C. Han, Y. Huang, and J. Qin, "A learning-based multimodel integrated framework for dynamic traffic flow forecasting," *Neural Process. Lett.*, vol. 85, no. 1, pp. 407–430, Mar. 2018.

[84] H.-Y. Kim and J.-M. Kim, "A load balancing scheme based on deep-learning in IoT," *Cluster Comput.*, vol. 20, pp. 873–878, Oct. 2016.

[85] C. Caruso, D. Malerba, and D. Papagni, "Learning the daily model of network traffic," in *Proc. ISMIS*, 2005, pp. 131–141.

[86] F. Meng, P. Chen, L. Wu, and J. Cheng, "Power allocation in multi-user cellular networks: Deep reinforcement learning approaches," *ArXiv*, vol. abs/1901.07159, 2019.

[87] W. Kang and D. Kim, "Poster abstract: DeepRT: A predictable deep learning inference framework for IoT devices," in *Proc. IEEE/ACM 3rd Int. Conf. Internet Things Design Implement (IoTDI)*, Apr. 2018, pp. 279–280.

[88] N. Liu *et al.*, "A hierarchical framework of cloud resource allocation and power management using deep reinforcement learning," in *Proc. IEEE 37th Int. Conf. Distrib. Comput. Syst. (ICDCS)*, Jun. 2017, pp. 372–382.

[89] Z. Fan, X. Gu, S. Nie, and M. Chen, "D2D power control based on supervised and unsupervised learning," in *Proc. 3rd IEEE Int. Conf. Comput. Commun. (ICCC)*, Dec. 2017, pp. 558–563.

[90] P. Lynggaard, "Using machine learning for adaptive interference suppression in wireless sensor networks," *IEEE Sensors J.*, vol. 18, no. 21, pp. 8820–8826, Nov. 2018.

[91] L. Li, Y. Xu, Z. Zhang, J. Yin, W. Chen, and Z. Han, "A prediction-based charging policy and interference mitigation approach in the wireless powered Internet of Things," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 2, pp. 439–451, Feb. 2019.

[92] S. Grimaldi, A. Mahmood, and M. Gidlund, "Real-time interference identification via supervised learning: Embedding coexistence awareness in IoT devices," *IEEE Access*, vol. 7, pp. 835–850, 2019.

[93] V. M. Suresh, R. Sidhu, P. Karkare, A. Patil, Z. Lei, and A. Basu, "Powering the IoT through embedded machine learning and LoRa," in *Proc. IEEE 4th World Forum Internet Things (WF-IoT)*, Feb. 2018, pp. 349–354.

[94] S. K. Sharma and X. Wang, "Collaborative distributed *Q*-learning for RACH congestion minimization in cellular IoT networks," *IEEE Commun. Lett.*, vol. 23, no. 4, pp. 600–603, Apr. 2019.

[95] A. Iranfar, S. N. Shahsavani, M. Kamal, and A. Afzali-Kusha, "A heuristic machine learning-based algorithm for power and thermal management of heterogeneous MPSoCs," in *Proc. IEEE/ACM Int. Symp. Low Power Electron. Design (ISLPED)*, Jul. 2015, pp. 291–296.

[96] S. Toumi, M. Hamdi, and M. Zaied, "An adaptive *Q*-learning approach to power control for D2D communications," in *Proc. Int. Conf. Adv. Syst. Elect. Technol. (IC_ASET)*, Mar. 2018, pp. 206–209.

[97] M. Kulin, T. Kazaz, I. Moerman, and E. De Poorter, "End-to-end learning from spectrum data: A deep learning approach for wireless signal identification in spectrum monitoring applications," *IEEE Access*, vol. 6, pp. 18484–18501, 2018.

[98] J. Oksanen, J. Lundón, and V. Koivunen, "Reinforcement learning method for energy efficient cooperative multiband spectrum sensing," in *Proc. IEEE Int. Workshop Mach. Learn. Signal Process.*, Aug. 2010, pp. 59–64.

[99] W. Lee, M. Kim, and D.-H. Cho, "Deep cooperative sensing: Cooperative spectrum sensing based on convolutional neural networks," *IEEE Trans. Veh. Technol.*, vol. 68, no. 3, pp. 3005–3009, Mar. 2019.

[100] X. Ma, S. Ning, X. Liu, H. Kuang, and Y. Hong, "Cooperative spectrum sensing using extreme learning machine for cognitive radio networks with multiple primary users," in *Proc. IEEE 3rd Adv. Inf. Technol. Electron. Autom. Control Conf. (IAEAC)*, 2018, pp. 536–540.

[101] S. M. Zafaruddin, I. Bistritz, A. Leshem, and D. Niyato, "Distributed learning for channel allocation over a shared spectrum," *IEEE J. Sel. Areas Commun.*, vol. 37, pp. 2337–2349, 2019.

[102] D. A. Awan, R. L. G. Cavalcante, and S. Stanczak, "A robust machine learning method for cell-load approximation in wireless networks," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, 2018, pp. 2601–2605.

[103] C. Jiang, H. Zhang, Y. Ren, Z. Han, K.-C. Chen, and L. Hanzo, "Machine learning paradigms for next-generation wireless networks," *IEEE Wireless Commun.*, vol. 24, no. 2, pp. 98–105, Apr. 2017.

[104] C.-K. Wen, S. Jin, K.-K. Wong, J.-C. Chen, and P. Ting, "Channel estimation for massive MIMO using Gaussian-mixture Bayesian learning," *IEEE Trans. Wireless Commun.*, vol. 14, no. 3, pp. 1356–1368, Mar. 2015.

[105] Y. Xu, P. Cheng, Z. Chen, Y. Li, and B. Vucetic, "Mobile collaborative spectrum sensing for heterogeneous networks: A Bayesian machine learning approach," *IEEE Trans. Signal Process.*, vol. 66, no. 21, pp. 5634–5647, Nov. 2018.

[106] X. Zhou, M. Sun, G. Y. Li, and B.-H. F. Juang, "Intelligent wireless communications enabled by cognitive radio and machine learning," *China Commun.*, vol. 15, no. 12, pp. 16–48, 2018.

[107] Y. Li, H. Ji, X. Li, and V. C. M. Leung, "Dynamic channel selection with reinforcement learning for cognitive WLAN over BER," *J. Commun. Syst.*, vol. 25, no. 8, pp. 1077–1090, Aug. 2012.

[108] A. M. Koushik, F. Hu, and S. Kumar, "Intelligent spectrum management based on transfer actor–critic learning for rateless transmissions in cognitive radio networks," *IEEE Trans. Mobile Comput.*, vol. 17, no. 5, pp. 1204–1215, May 2018.

[109] Z. Li, W. Wu, X. Liu, and P. Qi, "Improved cooperative spectrum sensing model based on machine learning for cognitive radio networks," *IET Commun.*, vol. 12, no. 19, pp. 2485–2492, Nov. 2018.

[110] X.-L. Huang, Y. Gao, X.-W. Tang, and S.-B. Wang, "Spectrum mapping in large-scale cognitive radio networks with historical spectrum decision results learning," *IEEE Access*, vol. 6, pp. 21350–21358, 2018.

[111] T. Park, N. Abuzainab, and W. Saad, "Learning how to communicate in the Internet of Things: Finite resources and heterogeneity," *IEEE Access*, vol. 4, pp. 7063–7073, 2016.

[112] T. Wang, C.-K. Wen, H. Wang, F. Gao, T. Jiang, and S. Jin, "Deep learning for wireless physical layer: Opportunities and challenges," *IEEE China Commun.*, vol. 14, no. 11, pp. 92–111, Oct. 2017.

[113] Q. Zhang, L. T. Yang, Z. Chen, and P. Li, "A survey on deep learning for big data," *Inf. Fusion*, vol. 42, pp. 146–157, Jul. 2018. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1566253517305328

[114] Z. Pan, W. Yu, X. Yi, A. Khan, F. Yuan, and Y. Zheng, "Recent progress on generative adversarial networks (GANs): A survey," *IEEE Access*, vol. 7, pp. 36322–36333, 2019.

[115] Z. Wang, Q. She, and T. E. Ward, "Generative adversarial networks: A survey and taxonomy," *ArXiv*, vol. abs/1906.01529, 2019. [Online]. Available: http://arxiv.org/abs/1906.01529

[116] Y. Hua, R. Li, Z. Zhao, H. Zhang, and X. Chen, "GAN-based deep distributional reinforcement learning for resource management in network slicing," *ArXiv*, vol. abs/1905.03929, 2019. [Online]. Available: http://arxiv.org/abs/1905.03929

[117] M. Mahmud, M. S. Kaiser, A. Hussain, and S. Vassanelli, "Applications of deep learning and reinforcement learning to biological data," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 6, pp. 2063–2079, Jun. 2018.

[118] M. Mohammadi, A. Al-Fuqaha, M. Guizani, and J.-S. Oh, "Semisupervised deep reinforcement learning in support of IoT and smart city services," *IEEE Internet Things J.*, vol. 5, no. 2, pp. 624–635, Apr. 2018.

[119] N. D. Nguyen, T. Nguyen, and S. Nahavandi, "System design perspective for human-level agents using deep reinforcement learning: A survey," *IEEE Access*, vol. 5, pp. 27091–27102, 2017.

[120] K. Arulkumaran, M. P. Deisenroth, M. Brundage, and A. A. Bharath, "A brief survey of deep reinforcement learning," *IEEE Signal Process. Mag.*, vol. 34, no. 6, pp. 26–38, Nov. 2017.

[121] B. Lake, T. Ullman, J. Tenenbaum, and S. Gershman, "Building machines that learn and think like people," in *The Behavioral and Brain Sciences*. Cambridge, U.K.: Cambridge Univ. Press, 2016.

[122] A. Zappone, M. Di Renzo, and M. Debbah, "Wireless networks design in the era of deep learning: Model-based, AI-based, or both?" *IEEE Trans. Commun.*, vol. 67, pp. 7331–7376, 2019.

[123] N. C. Luong *et al.* (Oct. 2018). *Applications of Deep Reinforcement Learning in Communications and Networking: A Survey*. [Online]. Available: https://arxiv.org/abs/1810.07862

[124] W. Lee, M. Kim, and D.-H. Cho. (2017). *Deep Cooperative Sensing: Cooperative Spectrum Sensing Based on Convolutional Neural Networks*. [Online]. Available: https://arxiv.org/pdf/1705.08164.pdf

[125] L. Zhang, J. Tan, Y.-C. Liang, G. Feng, and D. Niyato, "Deep reinforcement learning-based modulation and coding scheme selection in cognitive heterogeneous networks," *IEEE Trans. Wireless Commun.*, vol. 18, pp. 3281–3294, 2018.

[126] H. Huang, J. Yang, H. Huang, Y. Song, and G. Gui, "Deep learning for super-resolution channel estimation and DOA estimation based massive MIMO system," *IEEE Trans. Veh. Technol.*, vol. 67, no. 9, pp. 8549–8560, Sep. 2018.

[127] M. Kim, N.-I. Kim, W. Lee, and D.-H. Cho, "Deep learning-aided SCMA," *IEEE Commun. Lett.*, vol. 22, no. 4, pp. 720–723, Apr. 2018.

[128] T. J. O'Shea, T. Erpek, and T. C. Clancy, "Deep learning based MIMO communications," *ArXiv*, vol. abs/1707.07980, 2017. [Online]. Available: http://arxiv.org/abs/1707.07980

[129] H. Ye, G. Y. Li, and B.-H. Juang, "Power of deep learning for channel estimation and signal detection in OFDM systems," *IEEE Wireless Commun. Lett.*, vol. 7, no. 1, pp. 114–117, Feb. 2018.

[130] K. I. Ahmed and E. Hossain. (Apr. 2019). *A Deep Q-Learning Method for Downlink Power Allocation in Multi-Cell Networks*. [Online]. Available: https://arxiv.org/pdf/1904.13032.pdf

[131] U. Challita, L. Dong, and W. Saad, "Proactive resource management in LTE-U systems: A deep learning perspective," *ArXiv*, vol. abs/1702.07031, 2017. [Online]. Available: http://arxiv.org/abs/1702.07031

[132] H. Ye and G. Y. Li, "Deep reinforcement learning for resource allocation in V2V communications," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2018, pp. 1–6.

[133] X. Du, H. Zhang, H. V. Nguyen, and Z. Han, "Stacked LSTM deep learning model for traffic prediction in vehicle-to-vehicle communication," in *Proc. IEEE 86th Veh. Technol. Conf. (VTC-Fall)*, Sep. 2017, pp. 1–5.

[134] Y. He, N. Zhao, and H. Yin, "Integrated networking, caching, and computing for connected vehicles: A deep reinforcement learning approach," *IEEE Trans. Veh. Technol.*, vol. 67, no. 1, pp. 44–55, Jan. 2018.

[135] L. T. Tan and R. Q. Hu, "Mobility-aware edge caching and computing in vehicle networks: A deep reinforcement learning," *IEEE Trans. Veh. Technol.*, vol. 67, no. 11, pp. 10190–10203, Nov. 2018.

[136] K. I. Ahmed, H. Tabassum, and E. Hossain, "Deep learning for radio resource allocation in multi-cell networks," *IEEE Netw.*, vol. 33, pp. 188–195, 2018.

[137] J.-B. Wang *et al.*, "A machine learning framework for resource allocation assisted by cloud computing," *IEEE Netw.*, vol. 32, no. 2, pp. 144–151, Mar./Jun. 2018.

[138] A. Testolin *et al.*, "A machine learning approach to QoE-based video admission control and resource allocation in wireless systems," in *Proc. 13th Annu. Mediter. Ad Hoc Netw. Workshop (MED-HOC-NET)*, Jun. 2014, pp. 31–38.

[139] I. AlQerm and B. Shihada, "Enhanced machine learning scheme for energy efficient resource allocation in 5G heterogeneous cloud radio access networks," in *Proc. IEEE 28th Annu. Int. Symp. Pers. Indoor Mobile Radio Commun. (PIMRC)*, Oct. 2017, pp. 1–7.

[140] O. Y. Al-Jarrah, Y. Al-Hammdi, P. D. Yoo, S. Muhaidat, and M. Al-Qutayri, "Semi-supervised multi-layered clustering model for intrusion detection," *Digit. Commun. Netw.*, vol. 4, no. 4, pp. 227–286, Nov. 2018.

[141] R. Li, Z. Zhao, X. Chen, J. Palicot, and H. Zhang, "TACT: A transfer actor–critic learning framework for energy saving in cellular radio access networks," *IEEE Trans. Wireless Commun.*, vol. 13, no. 4, pp. 2000–2011, Apr. 2014. [Online]. Available: https://doi.org/10.1109

[142] M. Qiao, Y. Ma, Y. Bian, and J. Liu, "Real-time multi-application network traffic identification based on machine learning," in *Proc. Adv. Neural Netw. (ISNN)*, 2015, pp. 473–480.

[143] Y. Zuo, Y. Wu, G. Min, and L. Cui, "Learning-based network path planning for traffic engineering," *Future Gener. Comput. Syst.*, vol. 92, pp. 59–67, Mar. 2019.

[144] P. Perera, Y.-C. Tian, C. Fidge, and W. Kelly, "A comparison of supervised machine learning algorithms for classification of communications network traffic," in *Proc. ICONIP*, 2017, pp. 445–454.

[145] W. Xing, N. Wang, C. Wang, F. Liu, and Y. Ji, "Resource allocation schemes for D2D communication used in VANETs," in *Proc. IEEE 80th Veh. Technol Conf. (VTC-Fall)*, Sep. 2014, pp. 1–6.

[146] Y. Xu, H. Zhou, X. Wang, and B. Zhao, "Resource allocation for scalable video streaming in highway VANET," in *Proc. Int. Conf. Wireless Commun. Signal Process. (WCSP)*, Oct. 2015, pp. 1–5.

[147] M. Li, L. Zhao, and H. Liang, "An SMDP-based prioritized channel allocation scheme in cognitive enabled vehicular ad hoc networks," *IEEE Trans. Veh. Technol.*, vol. 66, no. 9, pp. 7925–7933, Sep. 2017.

[148] X. Cao, L. Liu, Y. Cheng, L. X. Cai, and C. Sun, "On optimal device-to-device resource allocation for minimizing end-to-end delay in VANETs," *IEEE Trans. Veh. Technol.*, vol. 65, no. 10, pp. 7905–7916, Oct. 2016.

[149] Y. Qi, H. Wang, L. Zhang, and B. Wang, "Optimal access mode selection and resource allocation for cellular-VANET heterogeneous networks," *IET Commun.*, vol. 11, no. 13, pp. 2012–2019, Sep. 2017.

[150] L. P. Portugal-Poma, C. A. C. Marcondes, H. Senger, and L. Arantes, "Applying machine learning to reduce overhead in DTN vehicular networks," in *Proc. Braz. Symp. Comput. Netw. Distrib. Syst.*, May 2014, pp. 94–102.

[151] R. Hussain, Z. Rezaeifar, and H. Oh, "A paradigm shift from vehicular ad hoc networks to VANET-based clouds," *Wireless Pers. Commun.*, vol. 83, no. 2, pp. 1131–1158, Jul. 2015. [Online]. Available: https://doi.org/10.1007/s11277-015-2442-y

[152] R. Hussain, Z. Rezaeifar, Y.-H. Lee, and H. Oh, "Secure and privacy-aware traffic information as a service in VANET-based clouds," *Pervasive Mobile Comput.*, vol. 24, pp. 194–209, Dec. 2015. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1574119215001455

[153] R. Hussain *et al.*, "Secure and privacy-aware incentives-based witness service in social Internet of Vehicles clouds," *IEEE Internet Things J.*, vol. 5, no. 4, pp. 2441–2448, Aug. 2018.

[154] R. Hussain and H. Oh, "Cooperation-aware VANET clouds: Providing secure cloud services to vehicular ad hoc networks," *J. Indian Prosthodontic Soc.*, vol. 10, no. 1, pp. 103–118, 2014.

[155] L. Liang, H. Ye, and G. Y. Li, "Towards intelligent vehicular networks: A machine learning framework," *IEEE Internet Things J.*, vol. 6, no. 1, pp. 124–135, Feb. 2019.

[156] H. R. Arkian, R. E. Atani, A. Diyanat, and A. Pourkhalili, "A cluster-based vehicular cloud architecture with learning-based resource management," *J. Supercomput.*, vol. 71, no. 4, pp. 1401–1426, Apr. 2015. [Online]. Available: https://doi.org/10.1007/s11227-014-1370-z

[157] M. A. Salahuddin, A. Al-Fuqaha, and M. Guizani, "Reinforcement learning for resource provisioning in the vehicular cloud," *IEEE Wireless Commun.*, vol. 23, no. 4, pp. 128–135, Aug. 2016.

[158] M. S. Omar *et al.*, "Multiobjective optimization in 5G hybrid networks," *IEEE Internet Things J.*, vol. 5, no. 3, pp. 1588–1597, Jun. 2018.

[159] M. Simsek, M. Bennis, and N. Güvenç, "Context-aware mobility management in HetNets: A reinforcement learning approach," in *Proc. IEEE Wireless Commun. Netw. Conf (WCNC)*, Mar. 2015, pp. 1536–1541.

[160] K. Vasudeva, S. Dikmese, N. Güvenç, A. Mehbodniya, W. Saad, and F. Adachi, "Fuzzy logic game-theoretic approach for energy efficient operation in HetNets," in *Proc. IEEE Int. Conf. Commun. Workshops (ICC Workshops)*, May 2017, pp. 552–557.

[161] J. S. Perez, S. K. Jayaweera, and S. Lane, "Machine learning aided cognitive RAT selection for 5G heterogeneous networks," in *Proc. IEEE Int. Black Sea Conf. Commun. Netw. (BlackSeaCom)*, 2017, pp. 1–5.

[162] S. Fan, H. Tian, and C. Sengul, "Self-optimization of coverage and capacity based on a fuzzy neural network with cooperative reinforcement learning," *EURASIP J. Wireless Commun. Netw.*, vol. 2014, no. 1, p. 57, Apr. 2014. [Online]. Available: https://doi.org/10.1186/1687-1499-2014-57

[163] I. Alqerm and B. Shihada, "Sophisticated online learning scheme for green resource allocation in 5G heterogeneous cloud radio access networks," *IEEE Trans. Mobile Comput.*, vol. 17, no. 10, pp. 2423–2437, Oct. 2018.

[164] R. I. Ansari *et al.*, "5G D2D networks: Techniques, challenges, and future prospects," *IEEE Syst. J.*, vol. 12, no. 4, pp. 3970–3984, Dec. 2018.

[165] S. Liu, Y. Wu, L. Li, X. Liu, and W. Xu, "A two-stage energy-efficient approach for joint power control and channel allocation in D2D communication," *IEEE Access*, vol. 7, pp. 16940–16951, 2019.

[166] M. Ahmed, Y. Li, M. Waqas, M. Sheraz, D. Jin, and Z. Han, "A survey on socially aware device-to-device communications," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 3, pp. 2169–2197, 3rd Quart., 2018.

[167] P. Cheng *et al.*, "Localized small cell caching: A machine learning approach based on rating data," *IEEE Trans. Commun.*, vol. 67, no. 2, pp. 1663–1676, Feb. 2019.

[168] M. Haus, M. Waqas, A. Y. Ding, Y. Li, S. Tarkoma, and J. Ott, "Security and privacy in device-to-device (D2D) communication: A review," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 2, pp. 1054–1079, 2nd Quart., 2017.

[169] A. Asheralieva and Y. Miyanaga, "An autonomous learning-based algorithm for joint channel and power level selection by D2D pairs in heterogeneous cellular networks," *IEEE Trans. Commun.*, vol. 64, no. 9, pp. 3996–4012, Sep. 2016.

[170] Y. Luo, Z. Shi, X. Zhou, Q. Liu, and Q. Yi, "Dynamic resource allocations based on *Q*-learning for D2D communication in cellular networks," in *Proc. 11th Int. Comput. Conf. Wavelet Activ Media Technol. Inf. Process. (ICCWAMTIP)*, Dec. 2014, pp. 385–388.

[171] M. I. Khan, M. M. Alam, Y. L. Moullec, and E. Yaacoub, "Throughput-aware cooperative reinforcement learning for adaptive resource allocation in device-to-device communication," *Future Internet*, vol. 9, no. 4, p. 72, 2017. [Online]. Available: http://www.mdpi.com/1999-5903/9/4/72

[172] I. AlQerm and B. Shihada, "Energy-efficient power allocation in multitier 5G networks using enhanced online learning," *IEEE Trans. Veh. Technol.*, vol. 66, no. 12, pp. 11086–11097, Dec. 2017.

[173] J. Zhang, Y. Zhuo, and Y. Zhao, "Mobile location based on SVM in MIMO communication systems," in *Proc. Int. Conf. Inf. Netw. Autom. (ICINA)*, vol. 2, 2010, p. 360.

[174] K. Kim, J. Lee, and J. Choi, "Deep learning based pilot allocation scheme (DL-PAS) for 5G massive MIMO system," *IEEE Commun. Lett.*, vol. 22, no. 4, pp. 828–831, Apr. 2018.

[175] Z. Dong, J. Shi, W. Wang, and X. Gao, "Machine learning based link adaptation method for MIMO system," in *Proc. IEEE 29th Annu. Int. Symp. Pers. Indoor Mobile Radio Commun. (PIMRC)*, 2018, pp. 1226–1231.

[176] R. C. Daniels, C. M. Caramanis, and R. W. Heath, "Adaptation in convolutionally coded MIMO-OFDM wireless systems through supervised learning and SNR ordering," *IEEE Trans. Veh. Technol.*, vol. 59, no. 1, pp. 114–126, 2010.

[177] W. V. Mauricio, D. C. Araujo, F. H. C. Neto, F. R. M. Lima, and T. F. Maciel, "A low complexity solution for resource allocation and SDMA grouping in massive MIMO systems," in *Proc. IEEE 15th Int. Symp. Wireless Commun. Syst. (ISWCS)*, 2018, pp. 1–6.

[178] I. Ahmed and H. Khammari, "Joint machine learning based resource allocation and hybrid beamforming design for massive MIMO systems," in *Proc. IEEE Globecom Workshops (GC Wkshps)*, Dec. 2018, pp. 1–6.

[179] Y. Saito, Y. Kishiyama, A. Benjebbour, T. Nakamura, A. Li, and K. Higuchi, "Non-orthogonal multiple access (NOMA) for cellular future radio access," in *Proc. IEEE 77th Veh. Technol. Conf. (VTC Spring)*, 2013, pp. 1–5.

[180] M. N. Jamal, S. A. Hassan, D. N. K. Jayakody, and J. J. P. C. Rodrigues, "Efficient nonorthogonal multiple access: Cooperative use of distributed space-time block coding," *IEEE Veh. Technol. Mag.*, vol. 13, no. 4, pp. 70–77, Dec. 2018.

[181] M. Vaezi, R. Schober, Z. Ding, and H. V. Poor, "Non-orthogonal multiple access: Common myths and critical questions," *IEEE Wireless Commun.*, vol. 26, pp. 174–180, 2018.

[182] G. Gui, H. Huang, Y. Song, and H. Sari, "Deep learning for an effective nonorthogonal multiple access scheme," *IEEE Trans. Veh. Technol.*, vol. 67, no. 9, pp. 8440–8450, 2018.

[183] D. A. Awan, R. L. Cavalcante, M. Yukawa, and S. Stanczak, "Detection for 5G-NOMA: An online adaptive machine learning approach," in *Proc. IEEE Int. Conf. Commun. (ICC)*, 2018, pp. 1–6.

[184] L. Xiao, Y. Li, C. Dai, H. Dai, and H. V. Poor, "Reinforcement learning-based NOMA power allocation in the presence of smart jamming," *IEEE Trans. Veh. Technol.*, vol. 67, no. 4, pp. 3377–3389, Apr. 2018.

[185] J. Cui, Z. Ding, P. Fan, and N. Al-Dhahir, "Unsupervised machine learning-based user clustering in millimeter-wave-NOMA systems," *IEEE Trans. Wireless Commun.*, vol. 17, no. 11, pp. 7425–7440, Nov. 2018.

[186] F. Nikjoo, A. Mirzaei, and A. Mohajer, "A novel approach to efficient resource allocation in NOMA heterogeneous networks: Multi-criteria green resource management," *Appl. Artif. Intell.*, vol. 32, nos. 7–8, pp. 583–612, 2018.

[187] B. Wang, H. Zhang, K. Long, G. Liu, and X. Li, "Green resource allocation in intelligent software defined NOMA networks," in *Proc. Int. Conf. Mach. Learn. Intell. Commun.*, 2017, pp. 418–427.

[188] M. Liu, T. Song, and G. Gui, "Deep cognitive perspective: Resource allocation for NOMA based heterogeneous IoT with imperfect SIC," *IEEE Internet Things J.*, vol. 6, no. 2, pp. 2885–2894, Apr. 2019.

[189] S. Yao *et al.*, "Deep learning for the Internet of Things," *IEEE J. Comput.*, vol. 51, no. 5, pp. 32–41, May 2018.

[190] W. Samek, T. Wiegand, and K.-R. Müller, "Explainable artificial intelligence: Understanding, visualizing and interpreting deep learning models," *ArXiv*, vol. abs/1708.08296, 2017. [Online]. Available: http://arxiv.org/abs/1708.08296

[191] A. Adadi and M. Berrada, "Peeking inside the black-box: A survey on explainable artificial intelligence (XAI)," *IEEE Access*, vol. 6, pp. 52138–52160, 2018.

[192] R. Kozik, M. Choraś, M. Ficco, and F. Palmieri, "A scalable distributed machine learning approach for attack detection in edge computing environments," *J. Parallel Distrib. Comput.*, vol. 119, pp. 18–26, Sep. 2018. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0743731518302004

[193] N. Akhtar and A. Mian, "Threat of adversarial attacks on deep learning in computer vision: A survey," *IEEE Access*, vol. 6, pp. 14410–14430, 2018.

[194] K. Grosse, N. Papernot, P. Manoharan, M. Backes, and P. D. McDaniel, "Adversarial perturbations against deep neural networks for malware classification," *ArXiv*, vol. abs/1606.04435, 2016. [Online]. Available: http://arxiv.org/abs/1606.04435

[195] J. Su, D. V. Vargas, and K. Sakurai, "One pixel attack for fooling deep neural networks," *IEEE Trans. Evol. Comput.*, vol. 23, pp. 828–841, 2017.

[196] J. Konecný, H. B. McMahan, F. X. Yu, P. Richtárik, A. T. Suresh, and D. Bacon, "Federated learning: Strategies for improving communication efficiency," *ArXiv*, vol. abs/1610.05492, 2018. [Online]. Available: http://arxiv.org/abs/1610.05492

[197] A. Choudhary. (2019). *IA Hands on Introduction to Deep Q Learning Using OpenAI Gym in Phython*. [Online]. Available: https://www.analyticsvidhya.com/blog/2019/04/introduction-deep-q-learning-python/

Fatima Hussain (Senior Member, IEEE) received the Ph.D. and M.A.Sc. degrees in electrical and computer engineering from Ryerson University, Toronto, Canada. She joined the Network-Centric Applied Research Team as a Postdoctoral Fellow, where she worked on various NSERC-funded projects in the realm of the Internet of Things. She is working as a Security Analyst in "API Security and Governance" squad, Royal Bank of Canada, Toronto. She is leading the development and promotion of new APIs and API development learning curriculum along with API security and governance duties. She is the Technical Lead with Ryerson-RBC Research Collaborative Lab Initiative. She is also an Adjunct Professor with Ryerson University and her role includes the supervision of graduate research projects. Her background includes number of distinguished professorships with Ryerson University and the University of Guelph, where she has been awarded for her research, teaching and course development accomplishments within wireless telecommunication and Internet of Things. She is a prolific author with various conference and journal publications to her credit. Her research interests include API security, cyber security, and machine learning.

Syed Ali Hassan (Senior Member, IEEE) received the B.E. degree (Hons.) in electrical engineering from the National University of Sciences and Technology (NUST), Islamabad, Pakistan, in 2004, the M.S. degree in electrical engineering from the University of Stuttgart, Stuttgart, Germany, in 2007, and the M.S. degree in mathematics and the Ph.D. degree in electrical engineering from the Georgia Institute of Technology, Atlanta, GA, USA, in 2011. He is currently an Associate Professor with the School of Electrical Engineering and Computer Science, NUST, where he is also the Director of Information Processing and Transmission Research Group, which focuses on various aspects of theoretical communications. He was a Visiting Professor with Georgia Tech in Fall 2017. He has coauthored more than 150 publications in international conferences and journals. His research interest mainly focuses on signal processing for wireless communications. He served as a TPC Member for IEEE WCSP 2014, IEEE PIMRC 2013-16, IEEE VTC Spring 2013-2019, and MILCOM 2014-2018.

Rasheed Hussain (Senior Member, IEEE) is currently working as an Associate Professor and the Director of the Institute of Information Security and Cyber-Physical Systems, Innopolis University, Innopolis, Russia, where he is also the Director of the Networks and Blockchain Lab and serves as an ACM Distinguished Speaker. He is currently working on machine and deep learning for IoT security and API security. His research interests include information security and privacy and particularly security and privacy issues in vehicular ad hoc networks, vehicular clouds, and vehicular social networking, applied cryptography, Internet of Things, content-centric networking, cloud computing, API security, and blockchain. He serves as the Editorial Board Member for various journals, including IEEE ACCESS, IEEE Internet Initiative, *Internet Technology Letters* (Wiley), and serves as a Reviewer for most of the IEEE TRANSACTIONS, Springer, and Elsevier Journals. He also serves as the Technical Program Committee Member of various conferences, such as IEEE VTC, IEEE VNC, IEEE Globecom, IEEE ICCVE, IEEE ICC, and ICCCN. He is a member of ACM. He is a Certified Trainer for Instructional Skills Workshop and a recipient of Netherland's University Teaching Qualification (Basis Kwalificatie Onderwijs).

Ekram Hossain (Fellow, IEEE) is a Professor and an Associate Head (Graduate Studies) with the Department of Electrical and Computer Engineering, University of Manitoba, Canada. His research works have received over 27 000 citations to date (in Google Scholar, with H-index = 87). His current research interests include design, analysis, and optimization beyond 5G cellular wireless networks. He received the 2017 IEEE ComSoc TCGCC (Technical Committee on Green Communications and Computing) Distinguished Technical Achievement Recognition Award for outstanding technical leadership and achievement in green wireless communications and networking. He has won several research awards, including the 2017 IEEE Communications Society Best Survey Paper Award and the 2011 IEEE Communications Society Fred Ellersick Prize Paper Award. He was listed as a Clarivate Analytics Highly Cited Researcher in computer science in 2017, 2018, and 2019. He served as the Editor-in-Chief for the IEEE COMMUNICATIONS SURVEYS & TUTORIALS from 2012 to 2016. He currently serves as the Editor-in-Chief of IEEE Press and an Editor for the IEEE TRANSACTIONS ON MOBILE COMPUTING. He is a Distinguished Lecturer of the IEEE Communications Society and the IEEE Vehicular Technology Society. He is an Elected Member of the Board of Governors of the IEEE Communications Society from 2018 to 2020. He is a member (Class of 2016) of the College of the Royal Society of Canada. He is a fellow of the Canadian Academy of Engineering. He was elevated to an IEEE Fellow for contributions to spectrum management and resource allocation in cognitive and cellular radio networks.