Towards Massive Machine Type Communications in Ultra-Dense Cellular IoT Networks: Current Issues and Machine Learning-Assisted Solutions

Shree Krishna Sharma, *Senior Member, IEEE*, and Xianbin Wang, *Fellow, IEEE*

***Abstract*— The ever-increasing number of resource-constrained Machine-Type Communication (MTC) devices is leading to the critical challenge of fulfilling diverse communication requirements in dynamic and ultra-dense wireless environments. Among different application scenarios that the upcoming 5G and beyond cellular networks are expected to support, such as enhanced Mobile Broadband (eMBB), massive Machine Type Communications (mMTC) and Ultra-Reliable and Low Latency Communications (URLLC), the mMTC brings the unique technical challenge of supporting a huge number of MTC devices in cellular networks, which is the main focus of this paper. The related challenges include Quality of Service (QoS) provisioning, handling highly dynamic and sporadic MTC traffic, huge signalling overhead and Radio Access Network (RAN) congestion. In this regard, this paper aims to identify and analyze the involved technical issues, to review recent advances, to highlight potential solutions and to propose new research directions. First, starting with an overview of mMTC features and QoS provisioning issues, we present the key enablers for mMTC in cellular networks. Along with the highlights on the inefficiency of the legacy Random Access (RA) procedure in the mMTC scenario, we then present the key features and channel access mechanisms in the emerging cellular IoT standards, namely, LTE-M and Narrowband IoT (NB-IoT). Subsequently, we present a framework for the performance analysis of transmission scheduling with the QoS support along with the issues involved in short data packet transmission. Next, we provide a detailed overview of the existing and emerging solutions towards addressing RAN congestion problem, and then identify potential advantages, challenges and use cases for the applications of emerging Machine Learning (ML) techniques in ultra-dense cellular networks. Out of several ML techniques, we focus on the application of low-complexity Q-learning approach in the mMTC scenario along with the recent advances towards enhancing its learning performance and convergence. Finally, we discuss some open research challenges and promising future research directions.**

***Index Terms*— Cellular IoT, mMTC, 5G and beyond wireless, RAN congestion, Machine learning, Q-learning, LTE-M, NB-IoT.**

I. INTRODUCTION

The convergence of emerging wireless communication technologies, ubiquitous wireless infrastructure and vertical Internet of Things (IoT) applications such as industrial automation, connected cars and smart-grid is leading to an integrated enabling platform for future smart and connected societies. This platform envisions to synergistically integrate the ever-increasing number of smart devices (forecasted by IHS Markit to be around 125 billion by 2030), intelligent industry processes, people and societies together to enhance the overall quality of our daily life [1]. Towards supporting connected IoT devices, there are several recent developments in the area of licensed cellular technologies such as Long Term Evolution (LTE) for Machine-Type Communications (LTE-M) and Narrow-Band IoT (NB-IoT), and unlicensed technologies such as WiFi, ZigBee and LoRa [2]. Out of these, cellular technologies are considered to be promising due to their several advantages including Quality of Service (QoS) provisioning, wide coverage area and tight coordination, and therefore, cellular IoT is of the main focus in this paper.

This work was supported in part by NSERC Discovery and CREATE programs under project number RGPIN-2018-06254 and 432280-2013. (*Corresponding author: Xianbin Wang*).

S. K. Sharma was with the Department of Electrical and Computer Engineering, Western University, London, ON N6A 5B9, Canada. He is now with the SnT, University of Luxembourg, L-1855 Luxembourg City, Luxembourg (e-mail: shree.sharma@ieee.org). X. Wang is with the Department of Electrical and Computer Engineering, Western University, London, ON N6A 5B9, Canada (e-mail: xianbin.wang@uwo.ca)

A. Recent Developments in Cellular IoT

In recent years, cellular IoT has gained significant importance from academia, industries, regulators and standardization bodies to enable the incorporation of IoT devices in the existing cellular infrastructures. The ITU-R has categorized the emerging diversified telecommunication services in the upcoming 5G and beyond cellular networks into the following three classes [3]: (i) enhanced Mobile Broadband (eMBB),(ii) massive Machine Type Communications (mMTC), and (iii) Ultra-Reliable and Low Latency Communications (URLLC).

Out of the above-mentioned categories, the eMBB comprises of high data rate services of 5G systems while the mMTC deals with the scalable connectivity to a massive number of devices in the order of 10^6 devices per square kilometers with diverse QoS requirements [4]. On the other hand, URLLC aims to provide robust connectivity with very low latency. The main challenge in the mMTC case is to support a huge number of devices with the limited radio resources whereas the key challenge for the URLLC scenario is to provide extremely high reliability in the order of 99.999% within a very short duration in the order of 1 ms [5]. Among these usage scenarios, the mMTC has to deal with various non-conventional challenges including QoS provisioning, Random Access Network (RAN) congestion, highly dynamic and sporadic traffic, and large signalling overhead. To this end, this paper focuses on the involved issues and the potential enablers of the mMTC scenario with a particular emphasis on the RAN congestion problem and emerging Machine Learning (ML)-based solutions.

In terms of ongoing standardization efforts, the main cellular IoT standards introduced by the 3GPP are LTE-M and NB-IoT. Out of these, LTE-M is intended for mid-range IoT applications which can support voice and video services, while NB-IoT systems target to provide very large coverage and support for ultra-low cost devices [6]. Since MTC devices usually do not require high channel throughput, the existing LTE-M and NB-IoT standards allocate a small bandwidth for IoT devices, i.e., LTE-M assigns 1.4 MHz bandwidth while the NB-IoT allocates a significantly lower bandwidth of 180 kHz [7]. Despite these recent developments, there are still several challenges to be addressed while supporting MTC devices in cellular systems.

B. Challenges in Cellular IoT

Although centralized cellular systems provide several advantages in terms of providing large coverage, tight time synchronization and handover operations for mobile users, they are sluggish in terms of handling low-end devices and face several challenges in supporting a large number of MTC devices with diverse QoS requirements. While incorporating MTC devices in the existing LTE/LTE-A based cellular systems, cellular operators have to face a lot of challenges both at the operational and planning levels. More specifically, there arise various issues related to the MTC device deployment, mMTC traffic, energy efficiency of low-cost MTC devices and the network protocol aspects such as signalling overhead [8]. Also, sensing, processing and communications tasks in IoT-edge networks comprising of the massive number of sensor nodes and smart devices demand for highly energy-efficient IoT architectures and techniques [9]. Furthermore, network congestion may occur in different segments of LTE/LTE-A based cellular network including RAN, core network and signalling network [10]. Out of these, RAN congestion problem is crucial in ultra-dense cellular IoT networks due to the limited available radio resources at the access-side and the massive number of sporadic access attempts from heterogeneous MTC devices.

Existing contention-based protocols are effective to support the conventional Human-Type Communications (HTC), however, their performance significantly degrades in mMTC scenarios due to infrequent and massive number of access requests [11]. Also, due to limited available preambles in the existing LTE-based systems, several MTC devices may need to select the same preambles at the same time, resulting in a significantly high probability of collision in the access network. Furthermore, the number of transmission attempts from the massive number of heterogeneous IoT devices could be significantly large [12], and their activation periods and frame sizes could be very different [13]. This sporadic and dynamic nature of mMTC access attempts and data transmissions may result in the peak traffic in both the access and traffic channels well beyond the capacity of the IoT access network, thus leading to the inevitable congestion in an IoT access network [14]. Moreover, although data packets transmitted by IoT devices are relatively short, very high signalling overhead per data packet becomes another critical issue [8, 15]. To this end, it is significantly important to investigate suitable transmission scheduling and efficient signalling reduction techniques in ultra-dense scenarios by utilizing emerging tools such as ML.

C. Need of Machine Learning and Associated Challenges in IoT/mMTC Networks

Optimizing the operation of cellular networks in dynamic wireless environments has been challenging over the generations since the number of configurable parameters of a cellular network has been rapidly increasing from one cellular generation to the next one [16, 17]. The widely-used link adaptation techniques in the existing wireless systems, which adapt different physical layer parameters including transmission power and modulation and coding scheme based on the reliability/link of a communication link, may not be efficient in ultra-dense cellular IoT networks. This adaptation is based on the prediction of reliability of a wireless link in the form of some metrics such as Packet Error Rate (PER) and this prediction process becomes extremely complex due to the increasing trend of using multiple antennas, wideband signals and a number of advanced signal processing algorithms [18]. Furthermore, the prediction of PER with good accuracy becomes difficult in practice by using the conventional signal processing tools. Moreover, due to a significantly large number of environmental parameters such as channel state information, signal power, noise variance, non-Gaussian noise effect and transceiver hardware impairments, it becomes complicated to provide the near-optimal/optimal tuning of the transmission parameters to achieve the efficient link adaptation [19]. The severity of this problem greatly increases in ultra-dense networks due to the involvement of massive number of devices and system parameters.

Understanding the context of the surrounding wireless environment significantly facilitates in developing context-aware adaptive communication protocols and in taking optimized decisions. In wireless networks, context information may refer to various aspects including battery levels of devices, user activities, geospatial information, link quality, environmental parameters, network states and energy consumption [20]. Nevertheless, handling self-configuration, self-optimization and self-healing operations in the ultra-dense cellular networks becomes challenging since the networks need to observe dynamic environmental variations, learn uncertainties, plan response actions and configure the associated network parameters effectively. To this end, emerging ML-assisted techniques seem promising since they can play significant roles in learning the system variations/parameter uncertainties, classifying the involved cases/issues, predicting the future results/challenges and investigating potential solutions/actions [17]. Moreover, the conventional link adaptation techniques are more localized to a particular network and a geographical region, and do not usually consider their impacts on the other systems. However, future ultra-dense cellular networks will need to handle mutual impact among the involved entities to maximize the overall system performance. To this end, by utilizing the emerging collaborative edge-cloud processing platform [21], ML-assisted solutions can enable the utilization of global

TABLE I
CLASSIFICATION OF SURVEY/OVERVIEW WORKS IN THE AREAS OF IOT/MMTC, ML AND UDNS.

Main domain	Sub-topics	Survey References	Overview References
IoT/mMTC	Enabling technologies/protocols, challenges and applications	[1, 2, 6, 12, 49]	[8, 15, 47, 48, 50, 51]
	Random access schemes	[52]	[24, 53, 54]
	Traffic characterization and issues	[55]	
	Transmission scheduling	[57]	[56]
	IoT big data analytics	[58–60]	[21]
	QoS provisioning		[61]
Machine Learning (ML)	Intelligence in 5G networks	[22]	[17]
	Learning in IoT/sensor networks	[63]	[45, 62]
	Reinforcement learning	[64, 65]	
Ultra-Dense Networks (UDNs)		[68, 69]	[66, 67]

network knowledge at the edge-side, and also facilitate the coordination among different distributed systems. In this direction, the application of ML techniques to address various issues in dynamic wireless environments has recently received an important attention [17, 22] and in the context of MTC environments, some existing works have already studied the applications of different ML techniques in learning various system parameters [11, 23–26, 43].

However, the direct application of conventional ML techniques to complex and dynamic wireless IoT environments is not straight-forward due to several underlying constraints such as low computational capability of MTC devices, distributed nature and heterogeneous QoS requirements of IoT devices, and the distinct features of mMTC traffic as compared to the conventional HTC traffic [45]. Furthermore, due to the limited computed power and low memory size of IoT devices, implementing sophisticated learning techniques in IoT devices becomes challenging [46]. In this regard, this paper identifies the implementation issues of the ML techniques in ultra-dense IoT scenarios and provides an emphasis on computationally simpler Q-learning based solutions.

D. Review of Related Overview/Survey Articles

In this subsection, we provide a brief review of the existing survey and overview works in the main domains related to this paper, namely, IoT/mMTC, ML and Ultra-Dense Networks (UDNs). Also, we present the classification of the existing references related to these domains into different sub-topics which are listed in Table I.

Several existing papers have provided a survey and an overview of enabling technologies, protocols, challenges and applications of IoT/mMTC in different contexts [1, 2, 6, 8, 12, 15, 47–50]. Authors in [1] provided a comprehensive survey of existing IoT protocols including application protocols, service discovery protocols, infrastructure protocols, and discussed some enabling technologies including cloud computing, edge computing and big data analytics for IoT systems. Furthermore, the authors in [6] presented a detailed survey of MTC systems including its features, requirements and the required architectural enhancements in LTE/LTE-A based networks. In the context of short packet transmissions in mMTC/IoT environment, the contribution in [12] provided a review of the recent advances in information theoretic principles governing the transmissions of short data packets and discussed the applications of these principles to different scenarios including a two-way channel, a downlink broadcast channel and an uplink Random Access Channel (RACH). Another article [2] provided a comprehensive survey of three main low power and long range M2M solutions, namely, Low Power Wide Area Network (LPWAN), IEEE 802.11ah-based network and cellular M2M including LTE-M and NB-IoT. Also, the survey article [49] provided a comprehensive tutorial on the development of MTC design over different releases of LTE and recent user equipments belonging to the MTC and the NB-IoT categories, called CAT-M and CAT-N, respectively.

In addition, the article [48] presented a review on various features defined by the 3GPP to support Machine-to-Machine (M2M) communications in LTE-based cellular systems and discussed recent advances in different layers including the physical layer improvements under the enhanced MTC (eMTC), and MAC and higher layer enhancements brought by the extended Discontinuous Reception (eDRX). Furthermore, the article [15] highlighted the requirements and design challenges for mMTC systems and discussed various physical and Medium Access Control (MAC) layer solutions for energy-efficient and massive access. Also, another overview paper [47] discussed the physical limitations of MTC devices while operating in cellular networks, and then analyzed the impact of these device limitations on the link performance and the link budget design. Besides, the article [8] presented the new requirements and challenges in large-scale MTC applications, and discussed some enabling techniques including efficient overhead signalling protocols, data aggregation and in-device intelligent processing. Moreover, another overview article [50] provided a review of various features of NB-IoT introduced in LTE Release 14 including the increased positioning accuracy, multi-casting, enhanced non-anchor carrier operation and lower device power class, and the applicability of these features for NB-IoT systems. Additionally, the authors in [51] provided an overview on physical layer aspects of wireless IoT in different application scenarios, and subsequently discussed potential physical layer enabling technologies for wireless IoT systems.

The design of effective Random Access (RA) schemes in the mMTC environment is an important challenge due to massive access requests and sporadic device transmissions from a huge number of resource-constrained MTC devices. In this regard, Furthermore, the article [52] presented a comprehensive survey of various RA solutions attempting to enhance the RACH operation of LTE/LTE-A based cellular networks, and carried out the performance evaluation of LTE RACH from the energy efficiency perspective.

Furthermore, authors in [24] provided an overview of different RA overload control mechanisms to avoid the RAN congestion caused by the random channel access from the MTC devices. Moreover, the authors in [53] provided a review of emerging LPWAN technologies both in the unlicensed band (LoRa and SIGFOX) and in the licensed band (LTE-M and NB-IoT) while considering three common fundamental objectives of these access mechanisms, namely, high system capacity, wide coverage and long battery life. In addition, the article [54] provided an overview of the existing RA solutions towards supporting MTC devices in LTE/LTE-A based networks and these solutions are compared in terms of five key metrics, namely, access success rate, access delay, QoS guarantee, energy efficiency and the impact on the HTC.

The nature of MTC traffic is significantly different from the HTC traffic as detailed later in Section II-D, however, existing cellular networks are mainly optimized to support the HTC traffic. Therefore, it is important to understand and characterize MTC traffic to facilitate the incorporation of MTC devices in cellular networks. To this end, the authors in [55] provided a discussion on the traffic issues of MTC and the associated congestion problems on the access channels, traffic channels and core network, and presented a comprehensive review of the existing solutions towards addressing these problems along with their advantages and disadvantages.

One of the promising solutions to handle massive access requests in the mMTC scenarios is to employ suitable transmission scheduling techniques at the distributed MTC devices. In this context, the article [57] provided a detailed survey on the uplink scheduling techniques for M2M devices over LTE/LTE-A based cellular networks by considering various aspects of M2M communications such as scalability, energy efficiency, QoS support and multi-hop connectivity. Furthermore, the authors in [56] identified limitations for signalling and scheduling of M2M devices over the existing LTE-based cellular infrastructures and discussed some of the existing proposals.

Another important aspect in ultra-dense IoT networks is how to handle the massive amount of data generated from the resource-constrained sensors and MTC devices. In this regard, authors in [58, 60] discussed the connection between IoT and big data analytics, and provided a survey on the existing research attempts in the domain of big IoT data analytics. Also, authors in [60] provided an overview of the existing network methodologies suitable for real-time IoT data analytics along with the fundamentals of real-time IoT analytics, software platforms and use cases, and highlighted real-time IoT analytics issues related to network scalability, network fault tolerance, spectral efficiency and network delay. Another article [59] presented a panoramic survey on the big data for cyber physical systems from various perspectives including data collection, storage, processing, analytics, energy-efficiency and cybersecurity. Moreover, the article [21] presented basic features, challenges and enablers for big data analytics in wireless IoT networks, and discussed the importance of collaborative cloud-edge processing for live data analytics along with the associated challenges and potential enablers.

Besides, providing QoS support in ultra-dense IoT networks is challenging due to massive connectivity, heterogeneity and the resource constraints of the MTC devices as detailed later in Section II. While analyzing from the energy efficiency perspective, maximizing QoS usually becomes energy costly and higher energy efficiency can be achieved by considering satisfactory QoS levels. Motivated by this, authors in [61] provided a discussion on the need for QoS satisfaction and the methods to achieve QoS satisfaction efficiently. Also, game theory-based fully distributed algorithms were presented to enhance the energy efficiency of IoT systems while maintaining a desired QoS threshold.

In the direction of incorporating intelligence in 5G and beyond networks, there have been some recent attempts in applying Artificial Intelligence (AI)-based techniques to address various issues in wireless communications. The AI techniques can provide significant benefits in achieving efficient management, organization and optimization of various system resources in emerging ultra-dense 5G Heterogeneous Networks (HetNets). In this regard, the article [22] discussed the state-of-the-art AI-based techniques for intelligent HetNet systems by considering the objective of achieving self-configuration, self-optimization and self-healing. Furthermore, authors in [17] recently introduced the fundamental concepts of AI and its relation with 5G candidate technologies. Also, the challenges and opportunities for the application of AI in managing network resources in intelligent 5G networks were discussed.

In the IoT/mMTC environment, learning techniques need to consider the unique features such as heterogeneity, resource constraints and QoS requirements. In this direction, the recent article [63] provided an overview of existing context-aware computing studies along with the learning and big data related works in the direction of intelligent IoT systems. Furthermore, the article [45] discussed the applicability of different types of learning techniques in the IoT scenarios by taking their learning performance, computational complexity and required input information into account. Furthermore, authors in [62] discussed various aspects of deep Reinforcement Learning (RL) and its application in building cognitive smart cities while considering the use cases in the areas of water consumption, energy and agriculture.

In the context of RL techniques, a comprehensive survey of multi-agent RL is provided in [64] by considering the aspects of stability of the learning dynamics of the agents and adaptation to the varying behavior of other learning agents. Also, the recent article [65] provided a brief survey of the existing deep RL algorithms along with the highlights on current research areas and the associated challenges.

Additionally, there exist a few survey and overview papers in the area of UDNs [66–69]. The survey article [68] highlighted the key issues in incorporating M2M communications in the emerging UDNs and also identified different ways to support M2M communications the UDNs from the perspective of different protocol layers including physical, MAC, network and application. Moreover, the authors in [69] provided a comprehensive review on the recent advances and enabling technologies for UDNs along with a discussion on the widely-used performance metrics and modeling techniques. Besides, the authors in [66] provided an overview of the operation of UDNs in the millimeter-wave band and presented wireless self-backhauling across multiple hops to improve the deployment flexibility. In addition, the authors in [67] presented a potential architecture for an ultra-dense heterogeneous network and investigated the random access problem by considering dense deployment of MTC devices.

E. Contributions

Although several existing survey/overview articles reviewed in Section I-D have considered different aspects of mMTC systems, ML techniques and UDNs, a comprehensive analysis of the research issues involved in supporting the massive number of MTC devices in ultra-dense cellular IoT networks and a detailed review of the recent advances including ML-assisted solutions attempting to address these challenges are missing in the literature. As highlighted earlier in Section I-B, there arise several challenges while incorporating MTC devices in the existing LTE/LTE-A based cellular systems. The main issues include QoS provisioning to heterogeneous MTC devices, addressing random and dynamic MTC traffic, transmission scheduling with QoS support and RAN congestion. To this end, we highlight the main contributions of this survey paper below.

1) The major challenges faced by the existing cellular IoT networks in supporting the massive number of MTC devices are identified and the potential enabling technologies are highlighted along with the key features, traffic characterization and the application scenarios of the mMTC.
2) The inefficiency of the legacy LTE RA procedure in supporting MTC devices is pointed out and its adaptation for mMTC systems is presented along with the main features and channel access mechanisms of emerging cellular IoT standards (LTE-M and NB-IoT).
3) A mathematical framework for the performance analysis of transmission scheduling with the QoS support in an mMTC system is presented, and several limitations and the design aspects of short data packet transmission are identified.
4) Existing solutions towards addressing the RAN congestion problem in cellular IoT networks are reviewed along with the highlights on three emerging techniques.
5) The potential benefits, challenges and promising use case scenarios for the applications of emerging ML techniques in ultra-dense cellular networks are identified and the existing ML techniques are reviewed by broadly categorizing them into supervised, unsupervised and RL techniques.
6) A framework for the application of low-complexity Q-learning in addressing the RACH congestion problem is presented along with different exploration strategies, and some performance enhancement techniques are suggested in multi-agent and dynamic wireless environments.
7) Various research issues are identified and some interesting future directions are presented to stimulate future research activities in the related domains.

F. Paper Organization

The remainder of this paper is organized as follows: Section II identifies the main features, application areas and the potential enablers for mMTC in cellular networks along with a discussion on various issues associated with QoS provisioning in ultra-dense IoT networks. Section III highlights the inefficiency of the conventional LTE RA procedure in mMTC systems along with the basics of RA procedure in legacy LTE systems, and then present key features and channel access mechanisms in two emerging cellular IoT standards, namely, LTE-M and NB-IoT. Also, the characterization of MTC traffic is presented along with the 3GPP-based MTC traffic models and related works. Subsequently, Section IV provides a mathematical framework for the performance analysis of transmission scheduling with QoS support in the mMTC systems, and also present various design aspects and limitations of short data packet transmission in the mMTC environment. Section V presents a review of the existing solutions for RAN congestion problem in cellular IoT networks along with some emerging solutions while Section VI presents the advantages and challenges of ML techniques in wireless IoT systems and also provides an overview of the existing ML techniques. Also, it provides a detailed explanation on the Q-learning mechanism from the perspective of addressing RAN congestion minimization along with some Q-learning performance enhancement techniques. Finally, Section VII provides some research challenges and future directions, and Section VIII concludes this paper. To improve the flow of this paper, we provide the structure of the paper in Fig. 1 and the definitions of acronyms in Table II.

II. QoS Provisioning in Ultra-Dense IoT Networks

In this section, we first discuss various aspects related to QoS provisioning in ultra-dense IoT networks in the general context, and then provide specific details related to cellular IoT scenarios. The International Mobile Telecommunication (IMT) vision for 2020 and beyond envisions to provide the connection density target of about 10^6 devices per square kilometers [3]. However, the available radio resources and communication infrastructures are limited and there is a significant amount of cost involved while acquiring new radio resources such as spectrum and building new infrastructures. Because of this issue, existing communication networks need to be made as efficient as possible to support the massive number of devices, thus leading to the concept of ultra-dense IoT networks. Due to diverse types of emerging IoT services and heterogeneous capabilities of MTC devices, QoS provisioning in ultra-dense IoT networks is a crucial challenge.

The overall network performance and QoS of emerging Internet protocol-based networks significantly depend on the effective management of instantaneous traffic flowing in the network. In wireless IoT networks, the instantaneous aggregated traffic at the IoT gateway can be bursty and may greatly exceed the average aggregated traffic since it aggregates periodic transmissions from a large number of sensor nodes with different periods and frame sizes [13]. Due to this, there may occur congestion during possible bursty intervals

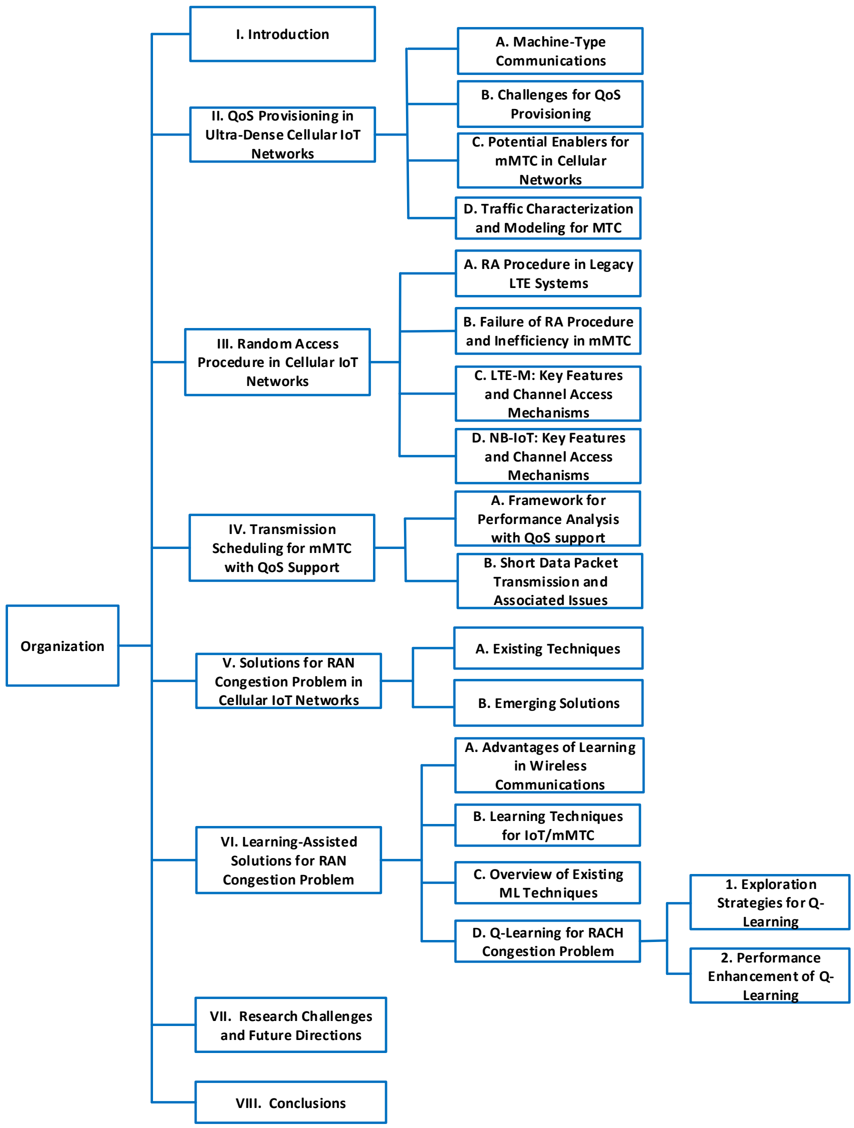

Fig. 1. Structure of the Paper

and much higher backhaul link (from aggregators to the cloud) bandwidth is needed than that required for the non-bursty traffic. In this direction, one of the important research challenges is how to make the aggregated traffic as close to the average traffic as possible.

In contrast to the conventional HTC traffic, there are several unique features of MTC traffic [70] which need to be considered while devising transmission scheduling and traffic management strategies for wireless IoT networks. The amount of small packets in IoT-type networks could become significantly large due to the resource-constrained sensor devices and the transmission of short packets from mMTC devices [12]. As compared to the dominant downlink traffic in the conventional cellular systems, uplink to downlink ratio for the MTC traffic is much higher. Furthermore, MTC devices usually have limited power budget and the MTC traffic consists of packets with the short payload length. Also, MTC traffic may arrive in the batch-mode due to high density of devices and correlated transmission [71]. Moreover, the standard Poisson process may not be suitable for modeling the MTC traffic since the MTC transmissions usually exhibit spatial and temporal synchronism. In addition, in contrast to the conventional voice traffic which has a constant sampling rate for a given codec, MTC traffic usually comprises of different packet sizes and inter-arrival patterns [72]. Different MTC applications have distinct characteristics and service requirements such as priority and delay constraints, thus leading to the need of separate traffic modelling and scheduling schemes to incorporate mMTC devices in the current LTE-based cellular networks.

In addition, MTC devices have completely different QoS requirements than that of the conventional HTC devices. MTC nodes are usually constrained in terms of battery power and the employed protocols need to be as energy-efficient as possible. Although the previous research in the area of wireless sensor network protocols mainly focused on monitoring applications based on low-rate delay-tolerant data collection, the current IoT-based research has moved to several new applications such as eHealthCare, industrial automation, military and smart home. These applications have different QoS requirements in terms of delay, throughput, priority, reliability and different

TABLE II
DEFINITIONS OF ACRONYMS

Acronyms	Definitions	Acronyms	Definitions
ACB	Access Class Barring	NPSS	Narrowband Primary Synchronization Signal
ACK	Acknowledgement	NSSS	Narrowband Secondary Synchronization Signal
AI	Artificial Intelligence	NPBCH	Narrowband Physical Broadcast Channel
BS	Base Station	NRS	Narrowband Reference Signal
CE	Coverage Enhancement	NPDCCH	Narrowband Physical Downlink Control Channel
CRA	Coded Random Access	NPDSCH	Narrowband Physical Downlink Shared Channel
CS	Compressive Sensing	NOMA	Non-Orthogonal Multiple Access
CIoT	Cellular IoT	OFDM	Orthogonal Frequency-Division Multiplexing
DCI	Downlink Control Information	OFDMA	Orthogonal Frequency-Division Multiple Access
DNC	Deterministic Network Calculus	PSK	Phase-Shift Keying
DQCA	Distributed Queuing Collision Avoidance	PDU	Periodic Update
eDRX	Extended Discontinuous Reception	PDCCH	Physical Downlink Control Channel
EPDCCH	Enhanced PDCCH	PRACH	Physical RACH
eMBB	Enhanced Mobile Broadband	PUSCH	Physical Uplink Shared Channel
ED	Event Driven	RB	Resource Block
FDD	Frequency Division Duplex	RRC	Radio Resource Control
HARQ	Hybrid Automatic Repeat Request	RAN	Random Access Network
HetNet	Heterogeneous Network	RA	Random Access
HTC	Human-Type Communications	RACH	Random Access Channel
H2H	Human to Human	RAR	Random Access Response
IoT	Internet of Things	RAW	Random Access Window
LSA	Licensed Shared Access	RL	Reinforcement Learning
LTE	Long Term Evolution	SAS	Spectrum Access System
LTE-A	LTE-Advanced	SCMA	Sparse Code Multiple Access
MAC	Medium Access Control	SC-FDMA	Single Carrier Frequency Division Multiple Access
MCL	Maximum Coupling Loss	SDN	Software Defined Networking
MDP	Markov Decision Process	SL	Sequential Learning
MTC	Machine-Type Communications	TA	Timing Alignment
mMTC	Massive MTC	TDD	Time Division Duplex
M2M	Machine-to-Machine	TTI	Transmission Time Interval
ML	Machine Learning	UE	User Equipment
MPRACH	MTC PRACH	UDN	Ultra-Dense Network
MPDCCH	MTC PDCCH	URLLC	Ultra-Reliable and Low-latency Communications
MUD	Multi-User Detection	UFMC	Universal Filtered Multi-Carrier
NB-IoT	Narrowband IoT	QoS	Quality of Service

traffic patterns such as event-driven, periodic and streaming [73]. To this end, it is crucial to consider these distinct QoS features while designing transmission and access techniques for the MTC devices.

In the above context, several existing works deal with the maximization of QoS while attempting to minimize the energy consumption. Several techniques such as sleep mode optimization, power control mechanisms, adaptation of the data rates and learning-assisted algorithms have been employed in various settings [61]. However, the objective of QoS maximization may lead to unnecessary energy consumption and achieving satisfactory levels of QoS may be sufficient to balance other performance metrics of systems such as energy efficiency. In contrast to the existing works related to the maximization of QoS, the objective of achieving satisfactory QoS levels can provide several benefits as highlighted in Table III [61].

Moreover, in time-critical MTC applications, several requirements in terms of low end-to-end delay, deterministic delay, bounds on systematic delay variations and linear delay-payload (packet size) dependence need to be considered [74]. The packet arrival period in the HTC systems such as multimedia ranges from 10 ms to 40 ms whereas this may range from about 10 ms to several minutes in MTC systems [75]. In addition, some applications such as data reporting in the smart grid have deterministic (hard) timing constraints and serious consequences may occur in the case of violence of these constraints. In this regard, multiplexing massive accesses with these diverse QoS characteristics effectively is a crucial challenge in ultra-dense IoT networks.

In the following subsections, we describe several aspects of MTC systems, highlight existing challenges for QoS provisioning in ultra-dense cellular IoT networks, present potential enablers for the incorporation of MTC devices in cellular IoT systems, and then present the characterization and modeling of the MTC traffic.

A. Machine-Type Communications

MTC has got a wide variety of application areas ranging from industrial automation and control to environmental monitoring towards building an information ambient society. The main applications of the MTC are listed below [6].

1) **Industrial automation and control**: This class includes several scenarios such as production on demand, quality control, automatic interactions among machines, optimization of packaging, logistics and supply chain, and inventory tracking.
2) **Intelligent transportation**: Under this category, MTC finds applications in different scenarios such as logistic services, M2M assisted driving, fleet management, e-ticketing and passenger services, smart parking and smart car counting.
3) **Smart-grid**: This category include various application scenarios including automatic meter reading, power demand management, smart electricity distribution and patrolling, online monitoring of transmission lines and transmission tower protection [76].
4) **Smart environment**: Several application scenarios including smart homes/offices/shops, smart lighting, smart

TABLE III
ADVANTAGES OF ACHIEVING DESIRED QOS LEVELS INSTEAD OF MAXIMIZING QOS

Advantages	**Related Causes**
1. Reduction of energy consumption	1. Extraneous energy will be wasted while maximizing QoS.
2. Better compliance with the fixed data rate services	2. No need to maximize data rates for fixed data rate services such as video surveillance and online gaming
3. A good perceived performance at the users' end	3. End-users are usually insensitive to small changes. in QoS levels, allowing the room for energy saving
4. Better support for emerging application-oriented networks	4. Desired QoS level is required only within a specific coverage area.
5. The set of feasible regions of optimization solutions are enlarged.	5. Relaxation of global optimum for the QoS maximization makes the mathematical problems less restrictive.
6. Adaptive resource allocation problem leads to the cost-effective solutions.	6. No need to waste additional radio resources by considering the conventional optimization assumptions such as full-buffer traffic.

industrial plants, smart water supply, environmental monitoring and green environment can be enabled with MTC.

5) **Security and public safety**: Several applications such as remote surveillance, personal tracking and public infrastructure protection can be considered under this category.
6) **e-Health**: In this application area, various scenarios exist such as tracking or monitoring a patient or a segment of an organ in a patient, identification and authentication of patients, diagnosing patient conditions and providing real-time information on patients health related data to the remote monitoring center.

Since MTC applications are significantly different from their HTC counterparts, they have distinct QoS requirements with different service features such as time-controlled, time-tolerant, small data transmission, low or no mobility, group-based connection, priority-based transmissions and low power consumption [77]. In this regard, the 3GPP has identified the following 14 features for M2M communications[1]: (i) low mobility, (ii) time-controlled, (iii) time-tolerant, (iv) packet switched only, (v) mobile originated only, (vi) small data transmission, (vii) infrequent mobile terminated, (viii) M2M monitoring, (ix) priority alarm message, (x) secure connection, (xi) location specific trigger, (xii) network-provided destination for uplink, (xiii) infrequent transmission, and (xiv) group-based policing and addressing.

Furthermore, the 3GPP has specified various general requirements for the MTC systems [6, 10] in order to effectively operate MTC devices and also to establish successful linkage between an MTC subscriber and the network operator. Some of the main technical requirements include: (i) providing a control mechanism to the network operators for the addition/removal/restriction of individual MTC device features, (ii) exploring a peak reduction mechanism for data and signaling traffic when a number of MTC devices concurrently attempt for their data transmissions, (iii) yielding a mechanism to restrict downlink data traffic and also limiting access towards a specific access point name in case of network overload, and (iv) investigating techniques to maintain efficient connectivity for a large number of MTC devices and to lower the corresponding energy consumption.

In comparison to the conventional HTC, the emerging MTC has the features of infrequent transmissions and low data rates. Also, the size of signalling data packets can be much larger than the size of user data packets in M2M applications [56]. Furthermore, although M2M devices need to transmit small amounts of data, communication infrastructure may get congested if a huge number of M2M devices attempt to access the network near-simultaneously [78]. Hence, the performance of the existing cellular standards has to be evaluated for this emerging type of traffic.

In addition, the 3GPP has identified the following performance objectives to support mMTC in the emerging air interface 5G New Radio (NR) [5, 79].

1) Very high connection density of about 10^6 devices per km^2 in an urban environment
2) Ultra-low complexity and low-cost IoT devices/networks
3) Battery life in extreme coverage beyond 10 years with the battery life evaluated at 164 dB MCL, and a battery capacity of 5 Wh.
4) Maximum Coupling Loss (MCL) of about 164dB for a data rate of 160 bps at the application layer
5) Latency of about 10 seconds or less on the uplink to deliver a 20-byte application layer packet (measured at 164dB MCL)

B. Challenges for QoS Provisioning in Ultra-Dense IoT Networks

There arise several challenges in incorporating MTC devices in LTE/LTE-A based cellular networks. First, the massive number of devices try to access the scarce network resources in a short period of time and there may arise the need of either utilizing the available resources efficiently or allocating additional bandwidth to incorporate these devices [80]. Secondly, there are significant differences in the transceiver properties and the applications of MTC devices from the existing LTE-based user terminals [47]. In most of the applications, MTC devices consume low power and have intermittent low rate transmissions. Furthermore, due to the need of cost-effective deployment of massive devices, MTC devices have degraded transceiver performance and reduced coverage as compared to the LTE user terminals. Besides, their effects in the communication performance of the existing LTE-A users need to be monitored and mitigated carefully. In this regard, one of the important research questions is how to provide concurrent access to a large number of MTC devices without degrading the QoS of the existing cellular users.

Since a network interface is fully utilized during the peak time, the devices may not be able to send or receive data and it

[1]For the description of these features, interested readers may refer to [10].

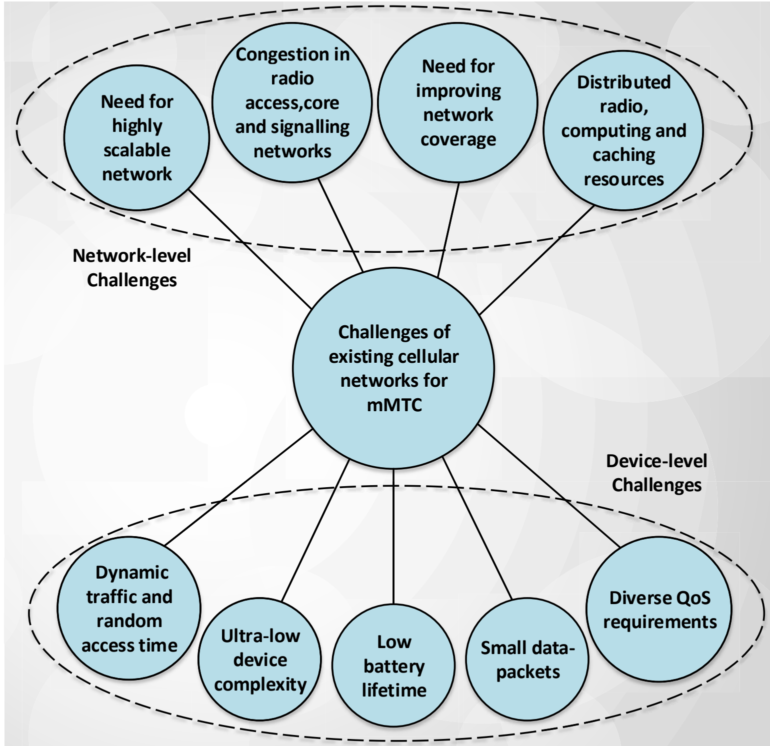

Fig. 2. Challenges of existing cellular networks to support emerging massive machine-type communications.

is crucial to optimize the peak traffic in the emerging content-centric wireless networks [81]. Furthermore, as highlighted earlier in Section II-A, the emerging MTC applications are quite different from the traditional HTC applications due to unique features such as group-based communications, time-controlled, small data transmissions, and low or no mobility [77, 82]. These distinct features of MTC applications result in diverse QoS requirements and it is important to take these QoS requirements into account while devising multiple access techniques for future cellular IoT networks. The main performance indicators of an mMTC system are the number of concurrent connections to be supported, energy efficiency and network coverage [83].

Existing cellular networks face the following major problems in supporting MTC devices [10, 84, 85]. In Figure 2, we present the pictorial representation of these issues in the form of device-level and network-level challenges.

1) **Highly dynamic traffic and random access time**: The data traffic arisen from the MTC devices is highly dynamic in nature as compared to more predictable HTC traffic. Furthermore, there arises a need to handle the mixed traffic models with the event-driven and periodic traffics. In addition, the existing contention-based radio access schemes will need to coordinate random transmissions from the massive number of devices [14].
2) **Ultra-low device complexity**: Due to the requirement of cheap MTC devices for mass deployment, the devices are constrained in terms of computational and memory resources, thus providing the limited performance.
3) **Low battery lifetime**: Because of the cost and space constraints, MTC devices are limited in their battery capacity. Furthermore, due to distributed nature of IoT devices and the involved cost-issues in replacing the batteries, the battery lifetime of MTC devices is expected to be more than 10 years with the battery capacity of 5 Wh, thus leading to the need of investigating power saving methods for ultra-dense cellular IoT networks.
4) **Small data packet transmissions**: In addition to the huge signalling burden associated with a large number of small packet transmissions from MTC devices, there arise other challenges such as the requirement of higher resource granularity and efficient channel coding for short block lengths in contrast to the channel coding schemes designed for long packets in the conventional cellular systems [15].
5) **Diverse QoS requirements**: MTC devices have diverse QoS requirements in terms of data rate and latency requirements and existing cellular technologies need to adapted to handle these features.
6) **Network congestion**: As highlighted earlier in Section I-B, the incorporation of massive MTC devices in the existing LTE/LTE-based cellular network may result in congestion in different segments of the network including RAN, the core network and the signalling network.
7) **Highly scalable network**: Because of the need to support a significantly large number of connected devices ranging from a factor of $10\times$ to $100\times$ as compared to the cellular devices, it is crucial to maintain the system performance with the increase in the connection density.
8) **Need for improving network coverage**: There arises significant shrinkage in the link budget due to the reduced capability of MTC devices. In order to increase

the coverage to the areas where MTC devices are deployed (such as deep inside a building), LTE release 13 targeted the coverage extension of at least 15 dB for the MTC devices. This coverage improvement enables the support of the devices in the locations where the conventional cellular networks face difficulty.

9) **Distributed radio, computing and caching resources**: With the recent trend of migrating communications networks from the connection-oriented to the content-oriented nature, it is important to investigate synergies among communications, computing and caching resources which are distributed across different devices in ultra-dense IoT networks [21]. However, the conventional cellular networks based on the centralized management are sluggish in terms of network resource management and they need to evolve to deal with the management of distributed resources.

Towards modeling and analysis of the QoS of wireless networks, one of the important mathematical tools is Deterministic Network Calculus (DNC), which is useful to calculate delay parameters such as delay bound, backlog bound and other service quality parameters by utilizing the traffic/packet arrival and service curves [86]. This DNC tool enables the determined boundary analysis for the system performance and offers a strict service guarantee by considering the worst-case scenarios. The main QoS metrics that can be evaluated include delay bound and backlog bound. The metric delay bound represents the maximum between arrival and service curves while the backlog bound denotes the maximum vertical deviation between these two curves.

In addition to providing high capacity to the fairly limited number of traditional user equipments to support high data rate services such as video streaming, the air interface of 5G cellular network has to provide connectivity to the massive number of concurrent transmissions coming from the MTC devices [87]. Also, the exchange of signalling information needs to be minimized both in the uplink and downlink due to a large number of MTC devices to be supported with the limited available radio resources. Furthermore, the RA procedure at the device-side should be simplified as much as possible by shifting the burden to the network side/eNodeB due to the resource constrained and low-cost nature of MTC devices.

Moreover, the conventional centralized approaches for congestion management in cellular networks are not scalable as desired by the MTC systems and also the distributed scheduling approaches can not easily acquire the knowledge about the network load and requirements of other applications [88]. Furthermore, the conventional congestion management process is mostly a reactive process instead of the proactive one needed for MTC devices. By deferring and shaping transmissions at the source itself in a network and being aware of the underlying application properties, better congestion management can be obtained for MTC devices [88].

Authors in [8] highlighted the difference between HTC over cellular and MTC over cellular in terms of various parameters such as uplink, downlink, subscriber load, device types and requirements in terms of delay, energy, signalling and cellular architecture. Furthermore, the signalling overhead in the uplink and downlink MTC links are analyzed considering the SMS-type raw data size (< 248 bytes) and email-type raw data size for smart metering and vehicular sensing applications via experimental measurements. It has been shown that the signalling overhead for the downlink control messages is considerably higher than for the uplink case, and higher signalling overhead occurs in vehicular applications than in the smart metering applications.

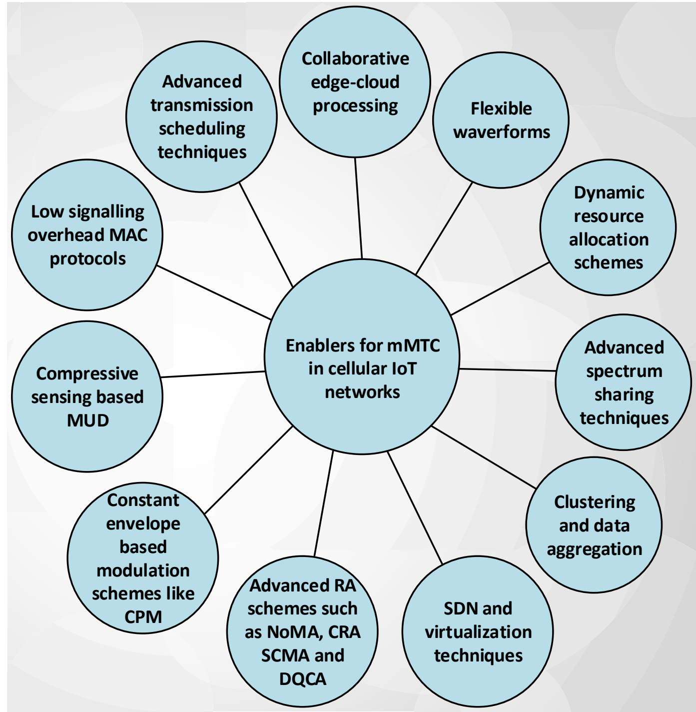

Fig. 3. Potential enabling techniques for mMTC in cellular IoT networks.

C. Potential Enablers for mMTC in Cellular Networks

Due to some distinct transitions while going from the conventional HTC platform to the emerging mMTC such as from the larger packet sizes to the smaller packet sizes, from the downlink-focused communication scenario to the uplink dominant, and from high data rate to low data rate transmissions from MTC devices, mMTC systems have new design requirements than those of the conventional HTC systems. To this end, it is important to look into the mMTC network design problem from a different perspective than the tradition approach followed for the HTC systems.

In Figure 3, we present the main enabling techniques being considered to facilitate the incorporation of mMTC in the upcoming cellular IoT networks. In the following, we briefly describe these enabling techniques.

1) **Flexible waveform design**: The design of flexible waveforms can enable the in-band mMTC channels within the LTE carrier [15]. In this regard, the traditional waveforms designed for HTC communications need to be adapted to support mMTC while considering various aspects such as end-to-end latency, robustness against time and synchronization errors, out-of-band radiations, spectral efficiency and transceiver complexity [89].

2) **Dynamic resource allocation techniques**: The 3GPP suggested to allocate RACH resources dynamically to address the RAN congestion problem [90]. For example, by deciding the number of preambles adaptively without knowing the number of devices and access probability, the RACH throughput can be maximized [91]. Furthermore, since this approach can dynamically change the size of the RACH resource pool and other resources, the total data collection time from the resource-constrained MTC devices can be minimized in delay-sensitive/emergency applications [92]. The two main issues for employing this process are the requirement of estimating the number of contending devices in an RA slot and determining the preamble pool size.
3) **Advanced spectrum sharing methods**: Although both the licensed and unlicensed bands can be exploited for mMTC applications, lack of QoS guarantees in the unlicensed band becomes highly problematic [4]. In this regard, emerging advanced spectrum sharing techniques such as Licensed Shared Access (LSA) and Spectrum Access System (SAS) [93] could be potential solutions for mMTC applications since they can provide better interference characterization.
4) **Clustering and data aggregation schemes**: By grouping MTC devices into smaller clusters based on some suitable criteria such as geographical locations or QoS requirements and then aggregating the individual device data at the MTC gateway/aggregator, the RAN congestion can be significantly minimized [83]. Furthermore, the investigation of energy-efficient clustering schemes facilitates the deployment of low-power MTC devices [83].
5) **Software Defined Networking (SDN) and virtualization techniques**: Based on the functionalities of MTC devices and their QoS requirements, a physical cellular network can be virtualized into different networks such as industrial, vehicular, smart grids and emergency networks, with all these networks sharing the same set of radio, computing and networking resources [94]. The dynamic sharing of resources and the reconfiguration of network elements among thus virtualized networks can be carried out by utlizing an SDN paradigm which decouples the control plane from the data plane and incorporates the capability of programming in the IoT network.
6) **Advanced RA schemes**: Several emerging RA schemes such as Non-Orthogonal Multiple Access (NOMA), Sparse Code Multiple Access (SCMA), Coded Random Access (CRA) [15] and distributed queueing based access protocol [95] can be considered as the promising enablers for the mMTC in cellular networks.
7) **Constant envelope coded-modulation schemes**: Due to space/cost constraints, MTC devices need to use low-cost amplifiers which are prone to non-linearities and hardware imperfections. In this scenario, constant envelope signals can enable the non-linear power-efficient and cost-effective operation at the MTC devices. Therefore, constant envelope coded modulation schemes such as Continuous Phase Modulation (CPM) can be considered as enablers for the mMTC [15].
8) **Compressed Sensing (CS)-based Multi-User Detection (MUD)**: The amount of collisions in the IoT access network can be further minimized by employing advanced interference cancellation receivers. As an example, the CS-MUD can enhance the resource efficiency and serve higher number of users by using the combination of non-orthogonal RA and joint detection of user data and activity [15]. In this regard, the combination of advanced MAC protocols with the CS-based MUD can be utilized by exploiting the sparse joint activity in the mMTC environment [96].
9) **Low signalling overhead MAC protocols**: One of the main technical challenges in an mMTC system is to reduce the amount of signalling overhead generated by the MTC devices and the design of low-signalling overhead protocols will facilitate the deployment of MTC devices in cellular networks [8].
10) **Advanced transmission scheduling techniques**: The transmission scheduling techniques designed for cellular IoT systems should be able to accommodate the MTC devices with heterogeneous QoS requirements in addition to the legacy cellular users. In this regard, advanced scheduling techniques such as latency-aware scheduling [97], fast uplink grant [98] and learning-assisted scheduling [11] seem promising to schedule the sporadic transmissions from a huge number of MTC devices over limited RACH resources.
11) **Collaborative cloud-edge processing**: Cloud computing platform has very high computational and storage capacity, and has a global view of the network but is not suitable for delay sensitive applications. On the other hand, edge-computing is suitable for applications demanding low delay and high QoS but has lower computational resources and storage capacity. In this regard, collaborative processing between these two platforms will be a promising approach to address various issues including latency minimization [99], dynamic spectrum sharing [100], peak traffic management and data offloading in ultra-dense IoT networks [21].
12) **Energy-efficient techniques for green IoT**: Due to high energy consumption caused by the massive number of IoT nodes and resource-constrained nature of IoT devices, it is crucial to optimize sensing, processing and communications operations to enhance the overall energy efficiency of cellular IoT networks. In this regard, a hierarchial framework comprising of sensing layer, gateway layer and control layer could be a promising energy-efficient architecture since it can balance the traffic load as well as elongate the system lifetime by utilizing energy-efficient mechanisms such as device sleep mode, sleep scheduling, and wake-up protocol [9]. Also, several base station switch-off strategies including random, distance-aware, load-aware and auction-based can be employed to balance the energy related performance trade-offs such as a trade-off between energy efficiency and data throughput [102]. Furthermore, vari-

ous green tag, sensing and internet technologies as well as energy-efficient scheduling, offloading and energy harvesting mechanisms can be utilized towards enabling green IoT networks [103]. In addition, the concept of energy internet [104] towards realizing the optimal usage of highly scalable and distributed energy sources seems promising to address the issues of energy shortage and greenhouse gas emissions in emerging IoT network.

The ML-assisted techniques, detailed later in Section VI, can address various issues related to self-configuration, self-optimization and self-healing in emerging wireless networks and seem promising in facilitating the implementation of the most of the technology enablers listed in Fig. 3 towards enhancing the performance of mMTC systems. However, the ML techniques should be as simple as possible to be applied in the MTC devices and the investigation of low-complexity adaptive ML techniques is one of the emerging future research directions as highlighted later in Section VII.

D. Traffic Characterization and Modeling for mMTC Systems

The characterization and modeling of mMTC traffic is crucial to support MTC devices in the existing cellular networks due to various reasons specified in the following. The incorporation of MTC devices in cellular networks may cause harmful interference to the existing cellular users and may significantly degrade the system performance of LTE/LTE-A based cellular systems. To this end, it is important to analyze the impact of MTC traffic on the existing cellular users by utilizing suitable interference modeling in realistic wireless environments [101]. Furthermore, suitable interference mitigation, resource allocation and resource sharing schemes need to be investigated to ensure the sufficient protection of the cellular users against harmful interference caused by the massive number of MTC devices and by utilizing the given MTC traffic models, these schemes can be designed in an efficient manner. Moreover, since MTC traffic is uplink dominant and the rigid QoS support framework of LTE designed for voice and data services may not be capable of addressing specific QoS requirements of MTC traffic in terms of latency, jitter and packet loss, suitable transmission scheduling techniques need to be investigated to support a large number of MTC devices while fulfilling their specific QoS requirements. Besides, to investigate suitable traffic management schemes such as peak traffic reduction in wireless IoT networks, it is essential to understand and characterize the traffic models applicable for a particular IoT application [14]. In addition, the traffic characteristics depend on the application scenarios and the MTC devices usually have heterogeneous traffic patterns in terms of their amplitudes, starting times and activation periods [73].

Existing traffic models in telecommunication systems can be categorized into: (i) source traffic models, mainly applicable for video, data and voice transmissions, and (ii) aggregated traffic model applicable for Internet, high-speed links and backbone networks [105]. Since an IoT network consists of a large number of sensors or MTC devices which are usually controlled by a gateway/server, IoT traffic at the gateway usually fits into the aggregated traffic model. In this regard,

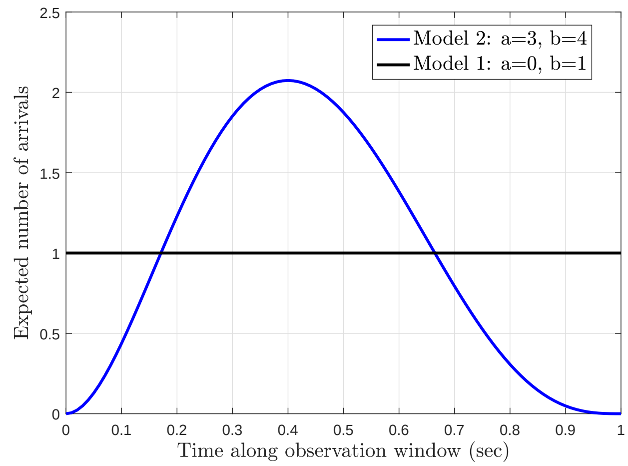

Fig. 4. Access intensity for two 3GPP based MTC traffic models.

the 3GPP has defined the following two types of aggregated MTC traffic models [90].

1) **Model 1**: Uniform distribution over a duration T in which MTC devices access the network uniformly over a period of time, i.e., in a non-synchronized manner. This model does not take account of the correlation between the transmissions of the devices.
2) **Model 2**: Beta distribution over T in which a large amount of MTC devices access the network in a highly synchronized manner. This model generates correlated traffic in a specific time interval.

For the first model, the following function can be considered.

$$f_i(t) = \begin{cases} P_i, & 0 \le t \le \tau_i, \\ 0, & \tau_i \le t \le T_i, \end{cases} \tag{1}$$

where P_i denotes the amplitude of the traffic profile for the ith device and τ_i/T_i represents the corresponding duty cycle. Similarly, for the second model, the probability density function of the Beta distribution, is given by [90]

$$p(t, \alpha, \beta) = \frac{1}{B(\alpha, \beta)} t^{(\alpha-1)}(1-t)^{(\beta-1)}, \tag{2}$$

where t denotes a single realization over the time axis, $\alpha > 0, \beta > 0$ are the scale parameters, and $B(\alpha, \beta)$ denotes the Beta function, which is a normalization constant to ensure that the total probability equals to 1.

Considering M number of MTC devices and their activation periods between $t = 0$ and $t = T$, the RA intensity for the mth device is given by the probability distributions either in (1) or (2) depending on the employed model. The number of arrivals in the ith access slot is given by [90]

$$I_{\text{access}}(i) = M \int_{t_i}^{t_{i+1}} p(t)dt, \tag{3}$$

where t_i denotes the time of the ith access opportunity and $p(t)$ given by (1) or (2). The distribution of access attempts should be limited over a certain observation period T in such way that $\int_0^T p(t)dt = 1$. Figure 4 illustrates the RA intensity of two 3GPP-based MTC traffic models with $T = 1$.

Although aggregated traffic modeling is suitable for scenarios involving a large number of devices and is less complex to realize, it is less precise than the source traffic modeling since it is not able to capture the real traffic features at the source level [70, 105]. On the other hand, source traffic modeling treats the traffic for every devices separately, and hence is more precise. However, modeling source traffic becomes complex for a large number of source devices. Therefore, it is crucial to investigate suitable traffic models which can combine the benefits of both the source and aggregate traffic models. In this regard, coupled Markov modulated Poisson processes [105] seems a promising approach, which has higher accuracy than the aggregated modeling and has lower complexity than the conventional source traffic modeling.

Considering a variety of applications, the MTC traffic can be categorized into the following three traffic patterns: (i) Periodic Update (PU), (ii) Event-Driven (ED) and (iii) payload exchange [106]. The PU traffic has a regular pattern, constant data size and is non-real time type (example: smart meter reading) while the ED traffic has a variable pattern, varying data size and is a real time traffic (example: health emergency alarming). On the other hand, payload exchange traffic follows either of the above traffic types and it may be of constant size or variable size, real time or non-real time depending on the application scenario. In addition, there exist three main types of traffic shaping policies [107]: (i) traffic shaping for bulk applications where each flow is assigned a fixed bandwidth, (ii) traffic shaping for the aggregate traffic, and (iii) time-based traffic shaping which is applied only at the peak-time to reduce congestion and cost.

The uplink traffic generated from the sensors in most of MTC applications is heterogeneous and can be classified into [108]: (i) non real-time with no task completion deadline, (ii) soft real-time with the decreased utility if the deadline not met and (iii) firm real-time having zero utility if the deadline is not met. As an example, industrial M2M traffic has very low latency requirements in the order of a few milliseconds [109]. In general, the PU traffic is periodic with tight service deadline while the ED traffic is random with all three traffic categories, i.e., non real-time, firm or soft real-time. From the scheduler designer perspective, it is crucial to maximize a system utility metric in order to maximally satisfy the delay requirements of all the classes [108].

Moreover, possible network applications in wireless IoT networks can be classified into the following [77].

1) **Elastic applications**: This category corresponds to more traditional HTC applications such as electronic email, file transfer as well as the downloading of remote data from the MTC servers. These applications are mostly delay tolerant in nature and the user utility usually has diminishing marginal improvements with the incremental increase in the achievable data rate.
2) **Hard real-time applications**: These applications have a desired delay constraint with hard real-time requirements. Beyond the desired time frame, there is no additional utility gain while increasing the data rate and the user utility becomes the step function of the achievable data rate.
3) **Delay adaptive applications**: Some delay sensitive applications can occasionally tolerate a small delay with a certain delay-bound violation and the packet dropping probability. The user utility in these applications (such as remote monitoring of e-Health services) deteriorates rapidly when the achievable data rate becomes less than the required intrinsic data rate.
4) **Rate-adaptive applications**: These applications try to adjust their transmission rates based on the available radio resources with the moderate delays. A highly efficient scheduler is needed to enhance the performance of these applications in time-varying channel conditions.

Traffic shaping, also called packet shaping, delays certain types of data packets in order to optimize the overall performance of a network. To achieve the optimized network performance, Internet traffic thesedays is intentionally shaped into ON/OFF pattern [110]. Also, ON/OFF pattern is generated due to some inherent characteristics of applications such as HTTP web browsing and MapReduce operation at the data center/server. The main benefits in performing ON/OFF traffic shaping include: (i) reduction of computing overhead at the server-side, (ii) energy saving at the wireless terminals and (iii) minimizing the bandwidth waste while delivering streaming services. However, this On/OFF traffic shaping faces several challenges such as the impact on packet drop probability, harmful effect on other real-time applications and weakening the congestion control function of the transmission control protocol [110]. To address these issues, it is crucial to design suitable models to characterize the relation among the associated parameters of ON/OFF traffic such as the ratio of ON/OFF duration, burst size, and burst transmission rate, and also the models for packet loss probability and temporary congestion caused by the bursty transmissions.

Another approach to manage the peak-traffic is to employ demand-side management, which adopts suitable measures at the customer-side/sensor-side to optimize the overall network performance [107]. On one hand, the demand profile can be flattened to limit the amplitude fluctuations while simultaneously accommodating the same amount of traffic volume. For example, a distribution reshaping concept can be employed to reshape the traffic arrival distribution having burstiness to a more flattened distribution of the RA attempts towards reducing the RA collisions as well as enhancing the utilization of RA resources [111]. On the other hand, the utilization profile of network resources can be flattened by rescheduling or delaying the services [81]. Some approaches for traffic scheduling include limiting the queue size and setting a bandwidth limit so that aggregate traffic size in the active queue does not exceed the limit. However, scheduling techniques cause delay for processing the network demand and it is crucial to investigate suitable techniques to minimize the delay introduced by traffic scheduling mechanisms. Furthermore, a suitable bandwidth limit should be applied to balance the trade-off between latency and energy saving [107].

III. Random Access Procedure in Cellular IoT Networks

An MTC device must go through the access procedure to establish a connection to the Base Station (BS)/eNodeB/access point mainly in the following situations [52]: (i) while establishing an initial access to the network, (ii) while receiving/transmitting new data and MTC device is not synchronized to the network, (iii) during the transmission of new data when no scheduling resources are allocated on the uplink control channel, (iv) to perform a seamless handover, (v) in order to re-connect to the network in the case of radio link failure.

The RA methods in the LTE-based cellular systems can be categorized into contention-based (for delay-tolerant access requests) and contention-free (for delay-sensitive requests) schemes [112]. Out of these, the contention-based scheme is of the main interest here due to the limitation in the number of available Resource Blocks (RBs) as compared to the massive number of access requests to be supported. In the contention-based RA approach, a huge number of MTC devices have to select the same preambles because of the limitation in the available preambles in the existing LTE-based cellular systems, and this results in significantly high number of collisions in the access network and subsequently leads to the RAN overload or radio access congestion problem in ultra-dense IoT networks. In this direction, one important question to be answered is how to concurrently support the massive number of MTC devices in ultra-dense cellular IoT networks without affecting the performance of the existing cellular devices by using the current communication technologies/standards.

In the following, we briefly describe the RA procedure in the legacy LTE systems and its inefficiency in handling the massive number of devices in the mMTC environment, then present some adaptations made to support MTC devices in cellular networks, and subsequently highlight the main features and access mechanisms in emerging cellular IoT standards, namely, LTE-M and NB-IoT.

A. RA Procedure in Legacy LTE Systems

After an eNodeB broadcasts the system information to the devices, the contention-based RA procedure follows a four-stage message handshake procedure as depicted in Fig. 5 [52, 113], which mainly involves the following four stages: (i) RA preamble transmission from the device to the eNodeB (Message 1), (ii) RA Response (RAR) from the eNodeB to the device (Message 2), (iii) connection request message from the device to the eNodeB (Message 3), and (iv) connection resolution message from the eNodeB to the device (Message 4).

In the first stage, each device randomly selects an RA preamble from the set of available preambles broadcasted by the eNodeB during the initial network synchronization phase and sends the RA request (Message 1) by transmitting thus selected preamble in an RACH. At this stage, the User Equipments (UEs) just transmit the selected preambles and not the device IDs. In the second stage, the eNodeB acknowledges the received distinct preambles with an RA response (Message 2) which includes the preamble index being acknowledged, instructions for the timing alignment and the command for the RB allocation. Subsequently, in the third stage, the UE recognizes the RA response addressed to it by noting the preamble index it has used for the RA request in the first step and utilizes the dedicated RB on the Physical Uplink Shared Channel (PUSCH). The devices which made current transmissions of the RA request with the same preamble in the first stage will be instructed to use the same RB in the step 3 and such transmissions will go through collisions. On the other hand, for the packets (which contain the corresponding device IDs) which are successfully decoded in step 3, the eNodeB sends a contention resolution message (Message 4) to the corresponding devices.

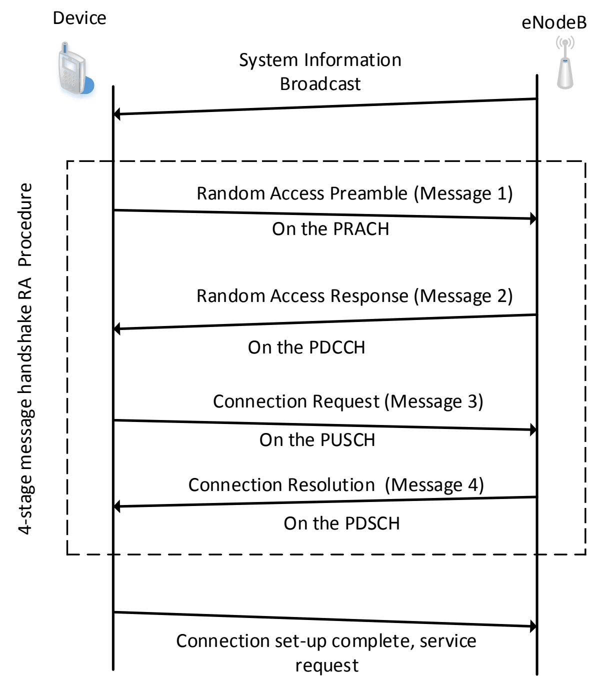

Fig. 5. Illustration of four-stage message handshake-based RA procedure in LTE-based cellular systems (PRACH: Physical Random Access Channel, PDCCH: Physical Downlink Control Channel, PUSCH: Physical Uplink Shared Channel, PDSCH: Physical Downlink Shared Channel).

In this RA procedure, after sending the preamble in the RA request (Message 1), the device sets an RAR window and waits for the eNodeB's response with an uplink grant (Message 2) in the RAR message. If the UE successfully receives its Message 2 within the defined RAR window, the UE sends the Radio Resource Control (RRC) connection request (Message 3) to the eNodeB. At this stage, the device starts the Message 4 timer and waits to receive its own RRC connection setup message (Message 4) from the eNodeB [114].

In the above-mentioned RA procedure, the physical-layer mapping of RACHs is called Physical RACHs (PRACHs) which are time-frequency blocks specified by the eNodeB. The periodicity of RA slots is broadcasted by the eNodeB in terms of the PRACH configuration index, which varies between every 1 ms (i.e., a maximum of 1 RA slot per 1 subframe) to 20 ms (i.e., the minimum of 1 RA slot every 2 frames) [52]. The transmission scheduling in terms of time and frequency depends on the configuration of the PRACH. For example, for the PRACH configuration index of 6, there will be RACHs in every 5 ms within a bandwidth of 180 kHz, with a duration ranging from 1 ms to 3 ms.

B. Failure of RA Procedure and Its Inefficiency in mMTC systems

The aforementioned RA procedure in the LTE/LTE-A based networks may fail due to the following reasons [114].

1) **Failure of preamble transmission**: This occurs either due to the collision of RACH preambles (caused due to concurrent transmission of the same set of preambles by more than one devices) or insufficient preamble transmission power.
2) **Failure of Message 2 reception**: This occurs mainly due to the lack of downlink radio resources (i.e., PDCCH) to send the RA response (Message 2) to all the received preambles within the devices' RAR windows.
3) **Failure of Message 3 transmission**: This occurs due to the failure in transmitting Message 3 to the eNodeB by employing the Hybrid Automatic Repeat Request (HARQ) process at the device.
4) **Failure of Message 4 reception**: This occurs due to the failure in receiving Message 4 by the devices within the Message 4 expiration time while using HARQ transmission from the eNodeB either due to insufficient PDCCH resources or an imperfect channel condition.

Out of 64 preambles used in LTE-A networks, 54 preambles are used for the contention-based access, while the remaining 10 preambles are reserved for the contention-free access which is needed for high-priority services such as handover. In every 5 ms, there arises an access opportunity and 200 access opportunities per second. This corresponds to the absolute maximum capacity of $10,800$ preambles per seconds in the absence of access collisions [52]. However, due to ALOHA type access protocol and random backoffs, performance becomes much lower than this maximum limit in practical cellular systems. Furthermore, the situation becomes worse while supporting the MTC devices. Besides the problem of supporting higher number of nodes in ultra-dense IoT networks, other performance metrics such as access delay and energy consumption are important to be considered.

To study the performance of the aforementioned 4-stage message handshake RA procedure in LTE systems, authors in [115] studied the stability limit of this legacy RA procedure, which indicates the probability of failure of the RA procedure and the maximum achievable throughput. It has been shown that the performance of the RA procedure deteriorates rapidly while sharing the Physical Downlink Control Channel (PDCCH) resources between Messages 2 and 4 with different priorities and the overall RA performance can be enhanced by increasing the size of the PDCCH resource.

Furthermore, in mMTC systems, the system performance may severely degrade in the presence of concurrent massive access requests due to high probability of collision caused by the signaling and traffic load spikes since the contention-based operation of the RACH in LTE-A networks is based on ALOHA-type access [116]. One of the possibilities to reduce the load of physical RACH is to increase the number of access opportunities scheduled in a frame, however, this approach will reduce the amount of resources needed for data transmission. In this regard, it is crucial to balance the tradeoff between the amount of resources available for data transmission and the amount of access opportunities to be scheduled per frame while designing an uplink scheduler for MTC applications by taking into account of limited available bandwidth. Besides, the main performance metrics to be improved include access success probability, preamble collision rate, access delay and device energy consumption [52].

The LTE RA procedure employed in legacy LTE systems is not efficient to support MTC devices mainly due to the following main reasons [87].

1) Because of the limited number of available preambles for the contention-based RA procedure, the massive number of concurrent transmissions of the same preambles would cause the overload of the RA procedure both in the uplink and the downlink and this will result in high collision probability, access failure rate and the access delay.
2) To support a huge number of access requests, additional downlink resources need to be allocated since each RAR message for one MTC device consists of 56 bits.
3) Even after an MTC device becomes successful in the RA procedure, the signalling overhead degrades the overall system efficiency since the size of the upload data payload from the MTC device is significantly smaller than the traditional cellular terminals.

C. Adaptation of RA Procedure for MTC devices

It should be noted that the RACH in the RA procedure is related to two different channels, namely, PRACH and PDCCH as illustrated in [112]. A single PRACH consists of six physical RBs and has a bandwidth of 1.08 MHz. Over the whole system bandwidth, a maximum of 6 RACHs can be deployed for time-division multiplexing with one RACH for frequency-division multiplexing. However, while sending RAR in the downlink channel, only a single PDCCH is responsible for handling multiple PRACHs. Although this is not a problem in the conventional HTC devices, this becomes a serious problem in MTC devices due to their hardware limitations in terms of their capacity to listen to the wideband PDCCH. In general, a low-cost MTC device consists of a single RF interface operating with 1.4 MHz bandwidth. To address this issue, the 3GPP has proposed Enhanced-PDCCH (EPDCCH) with the narrow bandwidth of 1.4 MHz for low-cost MTC devices and each PRACH has a dedicated NB EPDCCH. In this modified RACH structure adapted for low-cost MTC devices, the RA requests from the devices are distributed across multiple NB channels, thus reducing the congestion caused due to wideband nature of PDCCH in the conventional RA structure. Despite this enhanced RACH structure designed for low-cost MTC devices, the capacity of this RACH structure is not sufficient to handle the massive number of RA requests coming from the ever-increasing number of devices.

Towards addressing the problem of RACH overload in the cellular IoT systems, several methods have been proposed in the literature [112]. From the perspective that whether the device or the eNodeB employs the solution, the existing schemes can be broadly categorized into push-based and pull-based. In the first approach, the RA requests are controlled

TABLE IV
COMPARISON OF TWO MAIN CELLULAR IOT (LTE-M AND NB-IOT) TECHNOLOGIES

Feature/parameter	LTE-M	NB-IoT
Channel bandwidth	1.4 MHz	180 KHz
Transmission mode	HD-FDD/FDD/TDD	HD-FDD
Peak data rate	375 kbps (HD-FDD), 1 Mbps (TDD)	50 kbps (HD-FDD)
Latency	50-100 ms	1.5-10 seconds
Noise figure	9 dB (uplink), 5 dB (downlink)	5 dB (uplink), 3 dB (downlink)
Maximum coupling loss (MCL)	155.7 dB	164 dB
Modes of operation	In-band	Inband, guard-band and standalone
Power consumption	Best at medium data rates	Best at very low data rates
Mobility support	Full mobility	No connected mobility
Voice over LTE support	Yes	No

from the device-side while in the pull-based approach, the contention in the RA procedure is controlled from the eNodeB. Besides, there are some strict separation schemes and soft separation schemes to concurrently support both the HTC and MTC traffic in LTE-A networks [52, 116]. The strict separation schemes mainly comprise of the following: (i) resource separation: orthogonal allocation of resources between HTC and MTC traffic and dynamic shifting of resources among two classes, (ii) slotted access methods which define access cycles including the RA slots dedicated to the MTC device access, and (iii) pull-based scheme in which the MTC devices are allowed to access the PRACH only upon being paged by the corresponding eNodeB. On the other hand, soft-separation schemes include the following: (i) backoff tuning which assigns longer back-off intervals to the MTC devices which do not succeed during the preamble transmission of the RA procedure and (ii) Access Class Barring (ACB) scheme. A brief description of various existing and emerging solutions for the RAN congestion problem is provided in Section V.

Most of the existing MTC related works focus on the BS load balancing, radio resource management and the grouping of MTC devices, and only a few studies have been conducted in optimizing the access control of massive requests from the MTC devices [86]. The incoming requests from the MTC devices can be categorized into delay-sensitive and delay-tolerant based on the delay tolerance level of the underlying applications and the aggregator/BS can be equipped with two queues with one having higher priority over the other in order to deal with the two traffic classes. The criteria used for defining delay tolerant and delay sensitive may differ from one scenario to another [86].

The RA delay is one of the important aspects to be considered while designing RA techniques for mMTC in the existing cellular networks. In this regard, authors in [119] derived lower bounds for the LTE-A RA delay by considering uniformly distributed and Beta-distributed traffic arrivals and analyzed the effect of frequency of RA opportunities and the number of preambles. It has been shown that the RA delay can be reduced by several orders of magnitude by effectively tuning these system parameters.

As briefly highlighted in Section I, cellular IoT standards mainly comprise of two categories, namely, LTE-M and NB-IoT, which are described in the following subsections. Also, in Table IV, we highlight the key differences between these two technologies optimized to provide cellular connectivity to IoT devices [120, 121]. For the detailed differences among LTE-M, NB-IoT and legacy LTE in terms of supported features and functionalities for different uplink and downlink physical channels, interested readers may refer to [49].

D. LTE-M: Key Features and Channel Access Mechanisms

While looking at the history of MTC, the first generation of a full featured MTC device emerged in 3GPP Release R12. In this release R12, the 3GPP has defined the category 0, i.e., CAT-0 for the low-cost MTC operation [48]. In the subsequent releases, the efforts to incorporate mMTC devices continued and LTE release 13 (R13) in 2016 introduced two special categories, namely, CAT-M (also called LTE-M) for MTC and CAT-N for the NB-IoT to support various features of MTC/IoT applications. LTE Rel-14 enhancements were completed in June 2017, and the improvements under the Rel-15 are ongoing and are expected to be released by June 2018.

With respect to Cat-1 category which was the lowest UE category in LTE Release 11 from the perspective of transmission capability (peak rate of 10 Mbps in the downlink and 5 Mbps in the uplink), Cat-0 devices have a reduced complexity of about 50 % and have a reduced transmission rate of 1 Mbps for both the downlink and the uplink [49]. Also, Cat-0 category enables the use of only one receiver antenna with a maximum receiver bandwidth of 20 MHz and supports FDD half-duplex operation with relaxed switching time, eliminating the need of dual receiver chains and duplex filters for low cost MTC devices, respectively. In the subsequent LTE releases after the introduction of LTE-M in release 13, several new features have been added. The key features of LTE-M in different releases are included in the following [79].

1) **Release-13**: The main features included in this release include Coverage Enhancement (CE) mode A/B, bandwidth limited operations (1.4 MHz), half-duplex support, in-band operation mode, RRC connection, data transmission via a control plane, mobility support and eDRX.
2) **Release-14**: The key features incorporated in this version include multi-cast support with single-cell point-to-multipoint, positioning enhancements such as enhanced-cell ID requirements and observed time difference of arrival support, larger channel PDSCH/PUSCH bandwidth (up to 5 & 20 MHz), Voice over LTE enhancements, and support for HARQ-ACK bundling and inter-frequency measurements.
3) **Release-15**: The main features of this release are reduced latency and power consumption, lower UE power

class, improved spectral efficiency, improved load control of idle UEs, eDRX enhancements and support for higher UE velocity.

The four-stage message handshake procedure followed in the current LTE standard results in very high overhead for most of the IoT devices since the packets transmitted by the resource-constrained IoT devices are quite short as compared to the conventional cellular packets [52]. In this regard, several approaches are being investigated to design efficient channel access mechanisms to support MTC in the existing cellular systems. One approach investigated in the literature is to follow ALOHA-like immediate access without any reservation [122]. Although this scheme completely eliminates the channel reservation phase and provides very low latency, the system throughput is limited by the slotted ALOHA capacity of $1/e$. Besides, another approach is to utilize a preamble-initiated contention-based mechanism in which the nodes transmit a randomly selected preamble to reserve a time/frequency resource [7]. In contrast to the conventional RACH procedure followed in LTE, this method eliminates Message 3 and Message 4 of the four-stage message handshake procedure and the data is transmitted on the RB specified in the RAR message, thus significantly lowering the delay. However, if two or more nodes choose the same set of preambles for the RA request, the collisions occur which are detected by the lack of Acknowledgement (ACK) message.

From the performance analysis carried out in [7], it is shown that the preamble-initiated access achieves 86 % more capacity in comparison to both the conventional LTE access mechanism and ALOHA-like immediate transmission scheme for small data packet transmissions in IoT application scenarios. However, in terms of delay, the ALOHA-like scheme reduces the delay by about 62 % for low traffic loads in comparison to the preamble initiated access and by about 77 % as compared to the conventional LTE access mechanism.

To capture the signature of the LTE signal, an MTC device will need to receive the synchronization signals which occupy 6 RBs of the eNodeB's bandwidth [49]. Although decoding PDCCH becomes impossible due to the bandwidth limitation (only 1.4 MHz) of MTC devices, the enhanced version EPDCCH, which uses only one RB, is a good candidate, but is not sufficient for the required coverage enhancement [123]. However, increasing its bandwidth to 6 RBS in the 1.4 MHz bandwidth on the MTC device along with the repetition will provide the good coverage of about -14 dB and the EPDCCH also supports beamforming to enhance the coverage [47]. Therefore, 6 RBs, i.e., one narrowband is usually used as the basic unit for the MTC bandwidth [49].

The main changes incorporated in the physical layer operation of LTE-M as compared to the legacy LTE are briefly described below [49].

1) **Frequency hopping**: Due to narrow-bandwidth and a single receiver chain at the MTC device, the benefits due to spatial diversity and frequency diversity are not available. To compensate for the performance loss caused due to frequency diversity, the concept of frequency hopping, which allows MTC transmissions to hop from one NB channel to another, is employed. The challenges associated with frequency hopping in MTC devices include the need of retuning the RF chain, and the prior knowledge of the hopping pattern at the eNodeB and the device.
2) **Repetitions**: To achieve sufficient link budget in the downlink for the coverage enhancement, repeated copies of the same signal are transmitted over time to boost the link performance via time diversity. The main issue involved with this repeated transmission strategy is the requirement of increased decoding time, i.e., latency, demanding for longer wake-up time for the MTC device.
3) **MTC Physical RACH (MPRACH) and Physical Downlink Control Channel (MPDCCH)**: To compensate for the additional path-loss caused due to the extended coverage for the MTC device, the PRACH of legacy LTE needs to be modified. For this, frequency diversity and repetitions need to applied to the MPRACH to achieve the required diversity. Furthermore, additional features such as defining downlink control formats and enhancing control channel assignment procedure need to be added to support frequency hopping and repetitions in MPDCCH.
4) **MTC search spaces**: In order to reduce the number of decoding trials by the devices in the LTE systems, each MTC device can be assigned only a defined search space area of the whole control region to be monitored. In contrast to the legacy Enhanced PDCCH, there are mainly two classes of search spaces in MPDCCH, namely, device-specific search space and common search space.
5) **MTC Downlink Control Information (DCI) formats**: To reduce blind decoding iterations, i.e., device complexity as well as to facilitate the use of frequency hopping, repetition and enhanced coverage, three different DCI formats have been defined for uplink grant, downlink scheduling and paging in MTC devices.

E. Narrowband IoT: Key Features and Channel Access Mechanisms

To address various challenges of supporting MTC devices in cellular IoT networks specified in Section II-B, the 3GPP has proposed the concept of NB-IoT in its Release 13 [85]. The main objectives behind the NB-IoT concept include providing better indoor coverage and support to a massive number of low-throughput devices, with low power consumption and relaxed delay requirements [84]. To accomplish these objectives, the NB-IoT follows the procedures of optimizing control plane and user plane of Cellular-IoT (CIoT) evolved packet system towards reducing the signalling overhead for small data packet transmissions [124].

For both the uplink and downlink operations, NB-IoT can operate with an effective narrowband operation of 180 kHz bandwidth corresponding to one RB in the LTE network. In the downlink of an NB-IoT system, Orthogonal Frequency-Division Multiple Access (OFDMA) is employed with the subcarrier spacing of 15 kHz over 12 sub-carriers while in the uplink, both single tone and multi-tones are supported (single-tone with the subcarrier spacing of either 3.75 kHz or 15 kHz)

[125]. The NB-IoT usually can be operated in the following three operation modes [85, 126].

1) **In-band operation**: This mode of operation utilizes the RBs within an LTE carrier by reserving one RB for the NB-IoT system.
2) **Guard band operation**: This mode uses the unused resources within the guard band of the LTE carriers while ensuring that this does not affect the normal capacity of the LTE carrier.
3) **Stand-alone operation**: This mode utilizes the re-farmed GSM low band already existing in several countries (700 MHz, 800 MHz, and 900 MHz) [126].

The NB-IoT technology provides greater flexibility for the deployment of IoT devices in different applications such as smart city, smart home, smart metering and smart agriculture by reusing the existing network architectures. The main requirements for the NB-IoT system include the following [126].

1) **Low power consumption**: NB-IoT systems utilize the power saving mode and eDRX to maximize the battery life.
2) **Low channel bandwidth**: Due to low channel bandwidth of 200 kHz (180 kHz plus guard bands), GSM channel re-farming is applicable for NB-IoT systems since a single NB-IoT channel can utilize one GSM/GPRS channel.
3) **Low cost for the end-device**: Due to low channel bandwidth of 200 kHz, the front-end and digitizer of NB-IoT receivers are much simpler than that of the existing LTE-based systems operating on the bandwidth of 1.4 MHz, thus leading to low-complexity (cheaper) devices.
4) **Low deployment cost**: Besides the device cost, due to the capability of reusing existing GSM bands, the deployment cost for the network operators is significantly reduced.
5) **Extended coverage**: NB-IoT can provide about ten times better coverage area compared to the legacy GPRS systems as it can be achieve the additional 20 dB link budget gain.
6) **Support for massive number of connections**: Due to improved coverage and low channel bandwidth, it can support significantly higher number of MTC devices.

The main signals and channels involved in the downlink of an NB-IoT system are Narrowband Primary Synchronization Signal (NPSS), Narrowband Secondary Synchronization Signal (NSSS), Narrowband Physical Broadcast Channel (NPBCH), Narrowband Reference Signal (NRS), Narrowband Physical Downlink Control Channel (NPDCCH) and Narrowband Physical Downlink Shared Channel (NPDSCH) [127]. Out of these, NPSS and NSSS are used by an NB-IoT device to carry out cell search procedure including cell identity detection, and frequency and time synchronization. The NPBCH includes the master information block while the NRS is used to provide phase reference required for the demodulation of downlink signals. Similarly, NPDCCH includes the scheduling information for both the uplink and downlink data channels while the NPDSCH carries various information such as system information, paging message, RAR message and also data from the higher layers.

Besides, the uplink transmission scheduling of devices in the NB-IoT mainly comprises of Narrowband PRACH (NPRACH) and Narrowband PUSCH (NPUSCH) [125]. Out of these, NPRACH corresponds to the time-frequency resource used to transmit RA preambles and the NPUSCH is used for carrying the uplink data. The differences of the above-mentioned uplink and downlink channels from the legacy LTE systems are highlighted in [127]. The key technique employed by an NB-IoT system to obtain enhanced coverage with low complexity is repetition, which utilizes the repeated transmission of both data transmission and the involved control signalling transmission [125].

The RA procedure in the NB-IoT system is responsible for establishing a radio link during the initial access, for scheduling the transmission requests and to achieve uplink synchronization among the NB-IoT devices [127]. Three different types of NPRACH resource can be configured by assigning separate repetition values for a basic RA preamble to serve the devices belonging to different coverage classes with different ranges of path loss. The device estimates its coverage level by measuring the downlink received signal power, and then the device transmits an RA preamble in the NPRACH resources configured for the estimated coverage level. The configuration of NPRACH resources is made flexible in a time-frequency resource grid to enable the deployment of NB-IoT systems in different scenarios.

IV. Transmission Scheduling for MTC Systems with QoS Support

Most of the models used to analyze the capacity of wireless systems are based on physical layer models and they can not capture the link-layer QoS requirements such as bounds on the delay [128]. Therefore, physical-layer only models are not suitable for QoS support mechanisms such as resource reservation and admission control. Furthermore, in contrast to the wired links, it is challenging to guarantee QoS requirements in wireless systems due to low reliability, multi-path fading, co-channel interference and time-varying capacities. In order to incorporate complex QoS requirements into account, it is important to understand the queuing behavior of the connections and to capture the QoS requirements while characterizing the performance of packet-switching based wireless networks. For analyzing the queuing behavior, the characterization of source traffic as well as the services is an important aspect to be considered [128].

Due to distinct QoS requirements of MTC devices, it is crucial to provide QoS support for MTC devices in future wireless networks. For example, non-real time MTC applications such as data transmissions aim to enhance the reliability of transmission and do not have strict delay constraint. Whereas, real-time MTC applications such as video surveillance/demand, the important QoS metrics are strict latency and data rate requirements rather than high spectral efficiency [129]. To meet the QoS requirements of different network applications highlighted in Section II-D, it is crucial to design efficient radio resource allocation algorithms for the MTC devices in the uplink while considering the constraints on the available radio spectrum.

One of the potential candidate platforms to support MTC devices is the LTE-A standard and the 3GPP has been working on several enhancements of LTE-A standard towards this direction. The 3GPP uses Single Carrier Frequency Division Multiple Access (SC-FDMA) [130] as the multiple access scheme in the uplink of LTE cellular networks due to its main advantage of low Peak-to-Average Power Ratio (PAPR) as compared to that of the OFDMA. Due to this feature, the reduced requirements on the processing power and battery are suitable for the resource-constrained MTC devices [129]. However, the allocation of RBs in the SC-FDMA becomes complex as compared to that in the OFDMA scheme due to the sequential transmission of the RBs in the SC-FDMA in contrast to the transmission of orthogonal RBs in the OFDMA-based systems.

The minimum resource unit used for scheduling downlink and uplink transmissions in the LTE-A based cellular systems is referred to as an RB. Each RB comprises of 12 sub-carriers with each sub-carrier having the bandwidth of 180 kHz in the frequency domain and one sub-frame of 1 ms duration in the time domain [24]. The RB can be considered as a time-frequency resource in which an UE performs RA and each RA slot comprises of the bandwidth equivalent to the 6 RBs, i.e., 1.08 MHz and its duration in the time domain is 1 ms. In LTE-based cellular networks, the eNodeB broadcasts the periodicity of the RA slots by means of a variable referred to as the Physical RACH (PRACH) configuration Index [52], and subsequently, the MTC devices and the legacy cellular users can perform RA by using the PRACH channel. Even though the data size from the MTC devices is significantly small, the massive number of devices attempt to concurrently communicate over the same radio channel, thus leading to the network overload problem [24]. In contrast to the conventional HTC services such as multimedia for which the packet arrival periods range from 10 ms to 40 ms, the packet arrival periods in MTC applications may range from 10 ms to several minutes [75].

A. Framework for Performance Analysis with QoS Support

In this subsection, we present a mathematical framework to carry out the performance analysis of mMTC systems with QoS support in terms of the effective capacity, effective SNR and the estimated number of MTC devices. For this analysis, we consider the uplink in a single-cell of 3GPP LTE-A networks serving multiple MTC devices with the SC-FDMA scheme. This scenario can also be studied in conjunction with the legacy cellular/HTC users as in [131], however, herein, we deal only the case of MTC devices since we are interested in providing QoS support for MTC devices while maximizing some network performance metric subject to the constraints on the available radio spectrum. In practice, these devices can be grouped based on the employed transmission protocols and QoS requirements, and can be deployed on the cluster basis by using different wireless technologies such as WiFi, Bluetooth and Zigbee [129]. Furthermore, MTC devices can communicate to the eNodeB via an MTC gateway and the total available RBs can be divided between the access link (MTC devices to the MTC gateway) and the backhaul link (from the MTC gateway to the eNodeB) in the time domain as considered in [131].

Let us assume that there are M number of total MTC devices in the coverage area of the eNodeB, indexed by the set $\mathcal{M} = \{1, \ldots, m, \ldots, M\}$ and there are L number of available RBs, indexed by the set $\mathcal{L} = \{1, \ldots, l, \ldots, L\}$. We assume Poisson distribution for the traffic arrival rate of the MTC devices and block fading wireless channel between MTC devices and the eNodeB/gateway as in [129]. Also, we assume that channel coherence time is greater than the Transmission Time Interval (TTI) and the channel gain remains constant during a TTI.

To incorporate QoS requirements of MTC devices into the problem formulation, one way is to define a QoS exponent for each MTC device and to introduce this exponent in the definition of system capacity. Let θ_m denote the QoS exponent of the mth MTC device indicating a steady-state delay violation probability of the mth M2M device. Considering a queue of infinite buffer size required due to a constant arrival rate λ, the delay violation probability is given by [129]

$$\delta = \Pr(d_m > d_{\max}) \approx \phi_m(\lambda) e^{-\theta_m} d_{\max}, \tag{4}$$

where $\Pr(.)$ denotes the probability operation, d_m represents the delay experienced by a source packet of the mth MTC device, $d_{\max}$ is a delay bound, and $\phi_m(\lambda) = \Pr(d_m > 0)$ indicates the probability of non-empty buffer. In this formulation, the pair of $(\phi_m(\lambda), \theta_m(\lambda))$ can be used to characterize the link from the mth device to the gateway/eNodeB.

In contrast to the conventional physical layer-based capacity, we define the effective capacity to take the link-layer QoS requirements into account. The effective capacity [128] is defined as the maximum constant arrival rate that a given service process can support to guarantee a QoS requirement specified by θ and can be defined for the mth MTC device as

$$R_e^m(\theta_m) = -\frac{1}{\theta_m} \ln E[e^{-\theta_m R_m}], \tag{5}$$

where R_e^m denotes the effective capacity for the mth MTC device, θ_m represents the statistical QoS exponent of the mth MTC device, $E(.)$ denotes the expectation, and R_m is the data rate of the mth MTC device. In order to guarantee a QoS requirement of θ_m for the mth MTC device, the following condition should be satisfied [129]: $R_e^m(\theta_m) \geq \lambda_m$, where λ_m is the traffic arrival rate for the mth MTC device. By solving the above relation, one can obtain θ_m. Subsequently, by using the Shannon's capacity formula, the maximum achievable transmission rate for the mth MTC device can be expressed as

$$R_m = B\log_2(1 + \gamma_m) = B\log_2\left(1 + \frac{P_m |h_m|^2}{\sigma_n^2}\right), \tag{6}$$

where B is the bandwidth of each RB, P_m is transmission power of the mth MTC device, $|h_m|^2$ is the channel gain, σ_n^2

is the Additive White Gaussian Noise (AWGN) power, and $\gamma_m = \frac{P_m|h_m|^2}{\sigma_n^2}$ is the SNR for the mth MTC device.

Despite the significant benefits of SC-FDMA in terms of power and battery requirements, there arise some restrictions for uplink resource allocation (RB and power allocations) while employing SC-FDMA in the uplink [130]. The main aspects to be considered include: (i) a single RB can only be allocated to at most one user, (ii) multiple RBs allocated to a single user should be adjacent, and (iii) the transmit power on all the RBs allocated to a user should be equal. Let us assume that the set of RBs $\mathcal{L}_m$ is allocated to the mth MTC device in the current TTI, then the achievable rate (upper bound) from (6) in terms of effective SNR can be written as

$$R_m = B.L_m \log_2(1 + \gamma_{\text{eff},m}), \tag{7}$$

where $L_m = |\mathcal{L}_m|$ denotes the cardinality of the set $\mathcal{L}_m$ and $\gamma_{\text{eff},m}$ denotes the effective SNR for the mth MTC device. Since each data symbol is spread over the whole bandwidth in SC-FDMA transmission, the effective SNR can be computed as an average of SNRs over the allocated set of RBs to a particular MTC device as follows: $\gamma_{\text{eff},m} = \frac{1}{L_m}\sum_{l\in\mathcal{L}_m}\gamma_{m,l}$, where $\gamma_{m,l}$ is the SNR of the mth device for the lth RB. In the following, we summarize the main steps for calculating the maximum achievable rate for an MTC device in an SC-FDMA based cellular system: (i) determine the bandwidth of each RB and the set of RBs allocated to a particular MTC device, (ii) calculate the effective SNR as the average SNR over the set of RBs allocated to a particular MTC device, (iii) Utilize (7) to compute the maximum achievable rate for the considered MTC device.

Another aspect to be considered is how to effectively design the medium access scheme to support the massive number of devices. One approach is to determine the optimal size of the Random Access Window (RAW) based on the estimated number of MTC devices in the following way [132]. If there are idle slots available at the RAW, the eNodeB/access point can estimate the number of devices for the uplink access by using suitable estimation techniques such as maximum likelihood estimation method. Let $\tilde{I}$ be the measured number of idle slots in the uplink RAW, L_{UL} be the number of slots of the uplink RAW and N_{UL} be the number of devices for the uplink access. When N_{UL} devices contend in L_{UL}, the probability of selecting a slot by a device for the uplink access becomes $\frac{1}{L_{\text{UL}}}$ and the corresponding complementary probability is $(1 - \frac{1}{L_{\text{UL}}})$. Thus, the idle probability that no devices for the uplink access transmit the power save poll message, by which the device requests for the downlink data or the ACK frame from the eNodeB, is $p_{\text{idle}} = (1 - 1/L_{\text{UL}})^{N_{\text{UL}}}$ and the probability p_{idle} is estimated as: $\hat{p}_{\text{idle}} = \frac{\tilde{I}}{L_{\text{UL}}}$. Subsequently, the estimated number of devices for the uplink access by utilizing the aforementioned idle probability can be calculated as: $\hat{N}_{\text{UL}} = \frac{\log(\hat{p}_{\text{idle}})}{\log(1-\frac{1}{L_{\text{UL}}})}$.

On the other hand, the existing packet schedulers are mainly designed for a specific wireless system such as LTE and do not fully capture the heterogeneous characteristics of ultra-dense IoT networks. In this regard, authors in [108] proposed delay-efficient joint packet scheduling and subcarrier assignment by considering the classification of the uplink MTC traffic aggregated at the MTC aggregator into multiple classes based on traffic features such as packet size, arrival rate and delay requirements. By employing an MTC specific traffic model, the incoming data from the sensors at the aggregator is categorized either as ED or PU types and the delay requirements of these PU and ED traffic types are mapped onto sigmoidal and step utility functions, respectively. In addition, in order to ensure that the packets transmitted by an MTC device are within the delay budget, authors in [133] introduced a new MAC element, called Packet Age, with which the device informs the scheduler about the waiting time of the oldest packet in the device buffer along with the buffer size specified in the buffer status report.

B. Short Data Packet Transmission and Associated Issues

In this subsection, we briefly discuss various issues related to short data packet transmission in IoT systems along with its information theoretic perspective. One of the emerging areas in the MTC systems is ultra-reliable and low-latency communications, also known as mission-critical MTC. Some of the applications of mission-critical MTC are industrial control, intelligent transportation systems and smart grids for power distribution automation [134]. As an example, industrial control applications may need to transmit about 100 bits within 100 μseconds with 10^{-9} PER [135].

Existing wireless systems are designed to support the conventional HTC traffic having long packet sizes and each packet consists of information payload and the control information (metadata) which usually contains various information about logical addresses, packet initiation and termination, synchronization and security. As compared to the transmission of long packets, the transmission of short packets in the wireless IoT systems differs mainly in the following two ways [12]. First, existing transmission techniques are based on the assumption that the metadata is negligible as compared to the size of the information payload. However, this assumption does not apply to the transmission of short packets since the metadata size becomes no longer negligible, resulting in the need of highly efficient encoding schemes. Secondly, for the case of long packets, there exist channel codes which enable the reconstruction of information payload with high probability. The thermal noise and channel distortions average out for the case of long packets due to the law of large numbers, however, this averaging does not occur for the case of short packets and the classical law of large numbers is not applicable for mMTC applications, resulting in the need of new information theoretic principles. In this regard, authors in [12] discussed various information theoretic approaches to characterize the transmission of short packets in wireless communication systems and applied these principles on the transmission of short packets in various channels such as a two-way channel, a downlink broadcast channel and the uplink RACH. In addition, authors in [136] investigated the tradeoffs among reliability, throughput, and latency for the transmission of information over multiple-antenna Rayleigh block-fading channels.

In the above context, several recent works have investigated different physical layer approaches to support small packet

TABLE V
RECENT RESEARCH WORKS TOWARDS SUPPORTING SMALL DATA PACKET TRANSMISSIONS IN WIRELESS NETWORKS

Main Theme	**Applicable systems**	**References**
Optimization of pilot overhead	IoT Sensor networks	[137]
Design of air interface and waveforms	Multicarrier 5G systems	[138]
Coding and modulation schemes	IoT Sensor networks	[139]
Non-orthogonal multiple access	MTC scenarios	[140]
Minimization of the core network signalling	5G cellular network	[141]
Joint encoding of grouped messages	Wireless broadcast channel	[142]
Autonomous transmission mode	Delay tolerant IoT/MTC scenarios	[143]
Receiver algorithms to enhance the reception quality	5G wireless networks	[144]
Energy and information outage performance analysis	Wireless powered network	[145]
Exploitation of frequency diversity to enhance reliability	Tactile Internet	[146]

transmissions in mMTC/IoT environment by considering their specific characteristics, which are briefly reviewed in the following paragraphs. Also, in Table V, we list the recent research works towards supporting small data packet transmission in wireless networks with their main themes and applicable systems.

In short data packet transmissions, one effective way of enhancing the packet transmission efficiency is to optimize the pilot overhead [137]. However, most of the existing pilot overhead optimization works considering the objective of ergodic channel capacity maximization are based on the assumption of sufficiently large packet length resulting in small packet error probability, which is not suitable for short-packet transmission. In this regard, authors in [137] formulated the optimization of approximate achievable rate as a function of block length, pilot length and error probability, and illustrated the importance of considering packet size and error probability while optimizing pilot overhead via numerical results.

Another potential enabling approach to support short packet transmissions in IoT/mMTC environments is to design suitable transmit waveforms. In the IoT environment, there are some applications with very small packet sizes such as the data transmitted from sensors like temperature sensors while some other applications such as car to car and car to infrastructure communications demand very fast response time. In order to support these diverse set of applications, the 5G and beyond air interface should be able to support transmissions with very small air interface latency enabled by very short transmission frames [138]. In this regard, it is important to investigate suitable waveforms for supporting diverse applications in an IoT environment. Among potential multi-carrier waveform contenders such as filtered Cyclic Prefix-OFDM, Filter bank multi-carrier and Universal Filtered Multi-Carrier (UFMC), authors in [138] concluded UFMC as the best choice for IoT systems with short burst transmissions due to its several benefits in terms of supporting fast Time Division Duplex (TDD) switching, low latency modes, low energy consumption and small packet transmission.

In addition, investigating suitable coding and modulation schemes is crucial to achieve high energy efficiency for short-packet transmissions having low-duty cycles. Due to lower duty cycle, time synchronization and phase coherency for short-packet transmissions become non-trivial. Furthermore, because of short packet length, a large coding can not be achieved as in the conventional voice or data networks. Moreover, the overhead required to maintain time synchronization and phase coherency becomes significantly large while using the conventional coherent modulation schemes [139]. In this regard, the time synchronization overhead can be reduced by employing either non-coherent modulation/demodulation schemes such as Phase-Shift Keying (PSK) with differential encoding or orthogonal modulations. To this end, authors in [139] analyzed the tradeoff between energy efficiency and bandwidth in non-coherent short packet transmission systems.

Towards addressing the problem of scalability and efficient connectivity to the massive number of MTC devices with short packets, the NOMA scheme is considered as one candidate multiple access solution [140]. Due to its benefit of improving fairness and spectral efficiency for low-latency transmission with respect to the orthogonal multiple access technique, it is considered promising for IoT applications. In this regard, authors in [140] analyzed a trade-off among the transmission rate, transmission delay (in terms of block-length) and decoding error probability by considering a two user downlink NoMA system with finite block-length constraints.

Furthermore, towards minimizing the signalling overhead for small data packet transmissions, authors in [141] proposed a framework based on 5G RAN controlled user-centric mobility, in which an anchor node is allocated and updated for each end-device and it maintains the connection of the device to the core network within its coverage area. In this approach, an user centric area is dynamically allocated so that an user/device can move freely and communicate with the network without any state transitions signaling required in the existing connection management schemes with RRC protocols [141].

Moreover, while implementing multiple-antenna based interference suppression in IoT systems with small data packet structures, the insufficient training period may result in severe degradation in the estimation of the desired channel and interference covariance matrix. The main challenge here is to obtain the reliable channel estimation without significantly affecting the data transmission duration, i.e., to balance the trade-off between the pilot training period and the data transmission period. In this regard, authors in [144] investigated an efficient receiver structure which can exploit information received during the data transmission period to enhance the reception quality for the short packet transmissions. Moreover, in the context of energy harvesting networks, authors in [145] provided a comprehensive analysis of the backscatter wireless powered communication with sporadic short data packets by using a stochastic geometry framework.

V. Solutions for RAN Congestion Problem in Cellular IoT networks

In this section, we first review several existing techniques towards addressing RAN congestion problem in cellular IoT networks, and then discuss some emerging solutions.

A. Existing Techniques

Towards addressing the RAN congestion problem in LTE-based cellular networks, 3GPP has specified the following six different solutions of LTE RA congestion [90]:: (i) ACB, (ii) MTC-Specific backoff, (iii) dynamic resource allocation, (iv) Slotted random access, (v) separate RA resources and (vi) pull-based RA. In the following, we briefly describe the principles of these techniques along with other related solutions in the literature [54, 90, 147]. Also, in Table VI, we provide the list of these schemes along with their main principles and the corresponding references.

1) **Back-off based scheme**: In this scheme, the devices retransmit after a backoff time if they encounter a collision. This scheme can enhance the network performance under a low congestion level, however, becomes problematic in high-level congestions [148] This is the conventional approach followed in contention-based wireless networks and 3GPP has suggested several improvements to solve the RAN overload problem. To support MTC devices in the existing cellular networks, 3GPP has suggested the use of MTC-specific backoff scheme in which MTC devices are subject to a larger backoff interval than the HTC devices [90].
2) **Access Class Barring (ACB) scheme**: This scheme classifies the contending devices into multiple access classes with different access probabilities and each class is assigned to an ACB parameter and an access barring timer [90]. The working principle of the ACB scheme can be summarized in the following way. First, the BS broadcasts the ACB parameter, i.e., $0 \leq p \leq 1$ to the MTC devices and each MTC device trying to connect to the BS generates a random number $0 \leq r \leq 1$ uniformly. Then, the MTC device is allowed to start the RA procedure if $r < p$ and otherwise, the access to that particular device is barred and the device has to wait for a random backoff time determined based on the barring duration of that class. Therefore, by controlling the ACB parameter p, the BS can control the stabilization of RA to optimize some network performance metrics such as throughput [116].
   However, in the presence of severe congestion caused by the presence of massive number of IoT devices, the value of p may be set to be extremely low, thus leading to the intolerable delay. Also, the ACB scheme is not suitable for event-driven applications in which the contention may arise within a short time duration [52]. Furthermore, the operating parameters such as transmission probability should be adjusted based on the network status and estimating the number of devices/network status becomes challenging due to highly bursty traffic in event-driven MTC communications [148].
   To address the above drawbacks of the ACB scheme, there have been some attempts in the literature. Some of the important ones include the following.
   a) Extended Access Barrier (EAB) scheme [149]: In this scheme, devices belonging to a certain access class are barred from the channel access to provide some form of service differentiation [149]. The operation of this scheme depends mainly on the following two factors: (i) the sets of barred access classes and (ii) the time of turning EAB on or off. The larger the set of barred access classes, the higher will be the access success probability which comes at the cost of increased mean access delay. Furthermore, the timing of turning EAB on or off relies on the input network load which is proportional to the number of devices concurrently accessing the network.
   b) Cooperative ACB scheme [117]: In this scheme, ACB parameters are determined across the network jointly by many BSs interconnected via the X_2 interface [117] rather than individually calculated at each BS. This scheme aims to balance the traffic load among the BSs in a heterogeneous multi-tier cellular network with the objective of reducing the congestion level and also improving the access delay.
   c) Dynamic ACB scheme [118]: In this approach, the ACB parameters are updated dynamically based on the information about the number of collisions in the previous time slots.
   d) Prioritized RA with dynamic ACB [150, 151]: This scheme pre-allocates the RACH resources for different classes of MTC devices with class-dependent backoff procedures and reduces the number of concurrent requests for the RACH by employing the dynamic ACB method.
3) **Dynamic Resource Allocation**: In this scheme, the BS predicts the congestion level of the access network overload caused due to MTC devices and allocates additional RACH resources dynamically in the time domain or frequency domain or both for the MTC devices [54, 90]. However, the allocation of more radio resources for RACH will reduce the radio resources available for the traffic channels and this trade-off needs to considered while implementing this solution.
4) **Slotted Random Access**: In this method, a dedicated RA opportunity is provided to each MTC device and is allowed to perform RA only in the access slot allocated to it [90]. However, in ultra-dense IoT scenarios, this method will result in very high access delay since the duration for each RA cycle will be significantly large.
5) **Separation of RA Resources**: In this approach, different RACHs are allocated to MTC devices and HTC devices to avoid the impact of RA congestion on HTC devices. The separation of RA resources can be done either by splitting the available preambles into MTC and HTC subsets or by allocating different RA slots for MTC and

TABLE VI
SUMMARY OF EXISTING SOLUTIONS FOR RAN CONGESTION IN CELLULAR IOT NETWORKS

RA Scheme	Main principle	References
Back-off based scheme	In the occurrence of collisions, devices retransmit after an MTC-specific backoff period.	[90, 148]
Access Class Barring (ACB)	Multiple access classes of devices are assigned different access probabilities.	[116]
Extended Access Barrier (EAB)	A certain access class of devices is barred from the channel access.	[149]
Cooperative ACB scheme	ACB parameter is designed by many BSs in a collaborative way.	[117]
Dynamic ACB	ACB parameter is updated dynamically based on previous collisions.	[118]
Prioritized RA with dynamic ACB	Utilizes class-dependent back-offs and dynamic ACB.	[150, 151]
Dynamic resource allocation	Congestion level is predicted at the BS and additional RACH resources are allocated dynamically.	[54]
Slotted random access	Each device is assigned a dedicated RA slot and is allowed to perform RA only in that slot.	[90]
Separation of RA Resources	Available preambles or RA slots are divided between MTC and HTC devices.	[54, 90]
Pull-based/Paging-based	Devices perform RA attempts only after receiving paging messages from the BS.	[90]
Group-based RA	RACH resources are allocated on the basis of groups formed based on some defined criterion.	[152, 153]
Code-expanded RA	RA codewords are generated and each device utilizes a set of preambles in each RA slot.	[154]

HTC devices [54, 90].

6) **Pull-based/Paging-based scheme**: All the schemes described above fall under the category of push-based approach in which RA attempts are done randomly by the devices. However, in the pull-based method, the devices perform RA attempts only after receiving paging messages from the BS. To reduce the number of paging load in this approach, a number of MTC devices can be paged together by following a group paging method [90].
7) **Group-based RA Scheme** [152, 153]: The MTC devices can be grouped based on some criterion such as having similar QoS/delay requirements and being deployed in a specific geographical region, and RACH resources can be allocated on the group-basis to reduce the access network congestion. In a group-based RA scheme proposed in [152], the devices within one paging group are partitioned into different access groups based on some criterion and only one device within each access group, called group delegate/header, is made responsible for communicating with the BS. The group delegate can be decided by the BS based on some suitable metrics such as transmission power and channel conditions.
Another grouping approach is to divide the cell coverage area into a different spatial groups and to enable the use of same preambles at the same RA slot by the MTC devices located in different groups if the minimum distance of these MTC devices is larger than the multi-path delay spread [112, 155]. While sending the RAR message, the BS sends distinct RARs to all the detected devices having different Timing Alignment (TA) values even if they use the same preamble during the RA preamble transmission phase.
8) **Code-Expanded RA Scheme** [154]: In this approach, the contention space is expanded to the code domain by creating the RA codewords. While initiating an RA attempt, each device sends a set of preambles over the given RA slots instead of transmitting only a single preamble at any random RA slot, thus creating a set of preambles in each RA slot.
9) **Tree-based RA scheme** [156, 157]: This category of RA schemes utilizes the tree-based algorithms such as q-ary tree splitting technique [158], which rely on the utilization of feedback obtained after each contention attempt [157]. This RA scheme is mostly used to address the contention problem caused due to synchronized arrivals of the traffic from a large number of MTC devices. Furthermore, the combination of collision avoidance techniques such as access barring can be used in combination with tree-based collision resolution in order to form a hybrid RA scheme [156].

In Table VII, we present the qualitative comparison of the main existing RA schemes in terms of access delay, energy efficiency, access success rate and QoS guarantee, which are important to characterize the performance of an RA scheme in LTE-based cellular networks [54]. It can be depicted from Table VII that none of the techniques can perform equally better in terms of all the desired performance metrics and there arise trade-offs among these performance metrics. For example, group-based and code-expanded RA schemes perform better in terms of access delay but may not provide QoS guarantee. Also, the slotted RA scheme performs better in terms of energy efficiency and access success rate but is worse in terms of access delay and QoS guarantee. On the other hand, the prioritized RA scheme can provide higher QoS guarantee and medium performance in terms of energy efficiency and access success rate, but do not have a fixed access delay.

B. Emerging Solutions

In the following, we provide some of the emerging research directions to address RAN congestion problem in wireless IoT networks.

1) Learning-based Techniques: Recently, learning-based techniques have received important attention in addressing the RAN congestion problem in cellular IoT networks. In this direction, an RL scheme has been applied in [24] for the selection of an appropriate BS for the MTC devices with the objective of avoiding access network congestion and minimizing the packet delay. In addition, a Q-learning based access scheme has been studied in [11] to support MTC traffic in the existing cellular networks. In this Q-learning based approach, MTC devices learn to avoid collisions among each other without involving a central entity and after the learning convergence, each MTC device gets a unique RACH slot. Furthermore, authors in [25] applied a Q-learning based unsupervised algorithm in order to select an appropriate BS for MTC devices on the basis of QoS parameters in dynamic

TABLE VII
QUALITATIVE COMPARISON OF THE MAIN EXISTING SOLUTIONS FOR RACH CONGESTION.

RA Scheme	Access delay	Energy efficiency	Access success rate	QoS guarantee
Back-off based scheme	high	low	low	no
ACB, EAB and cooperative ACB schemes	varied	medium	high	high
Prioritized RA	varied	medium	medium	high
Dynamic resource allocation	medium	medium	medium	no
Slotted RA	high	very high	very high	very low
RA resource separation	high	low	low	no
Pull-based/Paging-based RA	medium	medium	medium	no
Group-based RA	low	medium	high	no
Code-expanded RA	low	very low	high	no

network traffic conditions. Moreover, a hierarchical stochastic learning algorithm has been applied in [159] to enable each device to make the access decision with the assistance of common control information broadcasted from the BS. In addition, in [23], a Q-learning algorithm has been applied to dynamically adjust the value of a barring factor to be allocated to the MTC device in the ACB scheme.

2) Distributed Queueing: The existing approaches to enhance the RACH performance are mainly based on the ALOHA-type mechanisms which suffer from some level of inefficiency, instability and uncertainty in the outcome of the access opportunities to be assigned to the devices [52]. In this regard, one promising approach is Distributed Queuing Collision Avoidance (DQCA) [160] which is a distributed and always-stable high performance protocol. This MAC protocol behaves as an RA mechanism for low traffic load and switches automatically and smoothly to a reservation scheme when the traffic volume increases [160, 161]. More specifically, the DCQA protocol utilizes two distributed queues which operate in parallel [161]. The first queue, called collision resolution queue, deals with the resolution of access-request signal collisions, while the other queue, called data transmission queue, helps to manage the data transmission. The main features of this protocol are the following [160].

1) It can eliminate back-off periods and avoid collisions in data packet transmissions.
2) Its performance is independent of the number of transmitting nodes.
3) It is stable independently of the traffic conditions.
4) As compared to other centralized or distributed MAC, it utilizes very few bits for signaling operation purposes.

Furthermore, authors in [95] proposed a distributed queuing-based access protocol for LTE with the objective of improving the RA performance for MTC systems without altering the existing frame structure of LTE systems. The original version of distributed queueing protocol envisions orthogonal mini-slots as access opportunities and its implementation requires a change in the LTE frame structure since the preambles in LTE are not orthogonal in time domain. To address this issue, the authors in [95] considered the distribution of allocated preambles for MTC devices among N_g virtual groups, with each virtual group having N_p number of preambles and each preamble being equivalent to one mini-slot considered in the original distributed queueing protocol.

3) SDN and Virtualization for RAN Management: To support differentiated MTC services with diverse QoS requirements, a physical wireless network can be abstracted and sliced into multiple virtual networks by employing suitable network function virtualization techniques [94]. On the other hand, Software Defined Networking (SDN) enables the separation of a data plane and a control plane, and provides the capability of programming a network via a centralized controller. Due to the global view of the underlying network, an SDN controller enables the efficient management of radio resources in dynamic network traffic and channel conditions. In cellular IoT networks, a hypervisor can divide the physical network into different IoT networks based on device classes and functionalities, and the SDN controller can dynamically allocate the available radio resources among these virtual networks to meet the QoS requirements of different IoT networks. Among these virtual networks, each MTC device can select one of the virtual networks to access to the physical network while meeting its connection requirements [162]. In addition to radio resources, it is also possible to enhance the utilization of other network resources such as computing, caching and networking resources [94].

VI. LEARNING-ASSISTED SOLUTIONS FOR RAN CONGESTION PROBLEM IN CELLULAR IOT NETWORKS

A. Advantages of Learning Techniques in Wireless Communications

The main questions this section attempts to answer are why learning techniques are important in wireless communication systems, which parameters to learn and for what purposes. First, we present the main advantages of learning techniques in wireless communications systems in general, and then discuss why learning techniques are needed on the top of the conventional link adaptation techniques. Subsequently, we discuss various parameters which can be learnt by using learning techniques in different application scenarios.

As highlighted earlier in Section I-C, the number of configurable system parameters has increased significantly from one cellular generation to the next one. For example, the number of configurable parameters has increased to about 1500 in a 4G node from about 500 in a 2G node and from about 1000 in a 3G node, and it is predicted to be around 2000 in a 5G node [16, 17]. In this regard, the process of optimizing these reconfigurable parameters in 5G and beyond systems becomes extremely complex and performing self-configuration, self-optimization and self-healing operations will be challenging. Also, emerging ultra-dense networks will need to observe environmental variations, learn uncertainties, plan response actions and configure the network parameters effectively to

TABLE VIII
USE CASES FOR THE APPLICATIONS OF LEARNING TECHNIQUES IN ULTRA-DENSE CELLULAR SYSTEMS

Use Case	Sub-cases/related description	Reference
RACH congestion minimization	Learning to find the unique access slot for each MTC device Learning to adapt an ACB parameter Learning to associate MTC devices with the best eNodeBs	[11] [23] [24, 25]
Autonomous adaptive resource allocation	Learning to find the existence of the critical delay-sensitive messages	[26]
Dynamic spectrum sharing	Learning to predict the occupancy status of radio channels	[163]
Learning-assisted edge-side processing	Extracting useful information from the raw sensor data at the edge devices to reduce the communication burden	[43] [43]
Learning-assisted traffic offloading	Learning to decide when and where to offload the traffic demands from the user-devices to enhance energy efficiency	[44]
Selection of a suitable RAT	Learning to select a suitable RAT among different RAT technologies under network conditions and user preferences constraints	[164]
Network traffic control	Learning to control network traffic to enhance computational efficiency and scalability	[165]
Adaptation of transmission parameters	Learning link quality/reliability to adapt parameters such as MCS and transmission slot	[18, 19, 171–173]
Data analytics	To extract user activity/mobility patterns, temporal, spatial and social correlations	[62]
Provisioning of personalized services	To learn network contexts for a global view of communications, computing and caching resources	[174]

handle these operations. To this end, emerging ML techniques could bring potential benefits in efficient handling of these operations. The main role of learning techniques include learning the system variations/parameter uncertainties, classifying the involved cases/issues, predicting the future results/challenges and investigating potential solutions/actions [17].

ML techniques can be utilized to address various issues in wireless networks including link adaptation, resource allocation and user scheduling (as highlighted in Table VIII). Typical problems, which are suitable to use ML, are usually too complex to be modelled but have the hidden patterns which can be explored with the assistance of ML. In this context, the widely-used link adaptation could be an appropriate example when it becomes a complex problem due to the impact of dynamic communication environment, resource condition and link quality. Therefore, in the following, first, we provide the justification about the need of ML over the existing link adaptation techniques, and then highlight its importance in addressing other different issues.

Wireless systems utilize link adaptation techniques to adapt the physical layer parameters such as modulation and coding scheme based on the reliability/quality of the communication link. In practice, different applications such as wireless video broadcasting and VoIP demand for different reliability constraints. Before performing this link adaptation, the reliability of a wireless link in the form of some metrics such as PER is predicted for each set of physical layer parameters, needs to be predicted [18]. During the link adaptation process, there arises a tradeoff between data rate and reliability since the PER calculated for a set of physical layer parameters is in general inversely related to the data rate. In order to predict the reliability, existing systems form explicit input/output models of a wireless channel and then analyze the performance of physical layer for each set of the parameters.

However, due to the increasing trend of using multiple antennas, wideband signals and a number of advanced signal processing algorithms, the above-mentioned reliability prediction process becomes extremely complex [18] and the prediction of PER with good accuracy becomes difficult in practice [175]. Furthermore, due to a significantly large number of environmental parameters such as channel state information, signal power, noise variance, non-Gaussian noise effect, transceiver hardware impairments such as power amplifier non-linearity and quantization error, it becomes challenging to provide the near-optimal/optimal tuning of the transmission parameters to achieve the efficient link adaptation [19]. The severity of this problem greatly increases in ultra-dense networks due to the involvement of various agents and system parameters such as Signal to Interference plus Noise Ratio (SINR) mismatch in ultra-dense small cell networks [176], and therefore, the link adaptation in emerging ultra-dense networks becomes extremely challenging. Also, the existing link adaptation systems are localized to individual links and small coverage areas, and do not take into account of the consequences on other systems from the system-level perspective.

In order to make the link/system adaptation more flexible and efficient, existing works have applied ML techniques in different settings [18, 19, 171–173]. The contribution in [18] investigated an online learning framework for the link adaptation by using a modified k nearest neighbor (kNN) algorithm to learn the mappings between the channel conditions and PER values for all possible Modulation and Coding Schemes (MCS) supported by the system. In this online learning framework, when a new packet is delivered, the predicted PER associated with each MCS is calculated by using the kNN algorithm and the best MCS is selected. After a packet is transmitted, the packet is stored as a prior data of the selected MCS for future prediction purpose. Although the accuracy of PER prediction becomes more accurate with the increase in the number of packet transmissions, this kNN-based approach requires to store all the previous samples and has higher time complexity, not suitable for a real time operation [18]. In this regard, the authors in [19] proposed an online kernelized support vector regression method which can work with the minimal memory size and has low computational complexity while providing a comparable performance to those of the existing algorithms.

Learning techniques are expected to provide significant benefits by adaptively learning numerous parameters in various application scenarios as listed below. Also, we provide a brief summary of these use cases along with the corresponding references in Table VIII.

1) Learning to exploit an unique RA slot for each MTC device within the considered transmission frame in a

way that concurrent transmissions in the same RACH opportunity can be avoided [11].

2) Learning to adapt an access control parameter, i.e., access barring factor for the RACH congestion [23].
3) Learning to associate MTC devices with suitable BSs/eNodeBs with the objective minimizing overall access network congestion [24, 25].
4) Learning the existence of delay sensitive/critical messages by IoT devices in heterogeneous ultra-dense IoT networks so that enough resources can be dynamically allocated and critical information can be successfully transmitted to the BS/eNodeB/aggregator as soon as they are generated [26]. Based on the learned information about critical messages, IoT devices can collectively adjust their uplink transmission parameters such as orthogonal codes, transmission slot period, periodicity of transmission and the received power for performing autonomous resource allocation and the coordination for the usage of available codes.
5) Learning the radio spectrum by dynamic spectrum sharing among the systems/nodes in a collaborative manner to predict the occupancy status of radio channels [163].
6) Learning the relationship of the contextual information (related to the surrounding radio environment) collected from IoT sensors to extract knowledge and to predict the future context at the edge devices [43]. Instead of transferring all the raw contextual data to the network/cloud-centre, only the inferred knowledge can be transferred, thus reducing the communication burden. This approach deals with pushing learning intelligence from the network to the distributed edge devices having heterogeneous computing abilities.
7) Learning-assisted traffic offloading in a heterogeneous cellular network for offloading time-varying traffic to the nearby small cells by employing Q-learning with the compact state representation algorithm (QC-learning) with the objective of minimizing the total discounted energy consumption while maintaining the QoS of the cellular users [44].
8) Learning for the selection of Radio Access Technology (RAT) while performing vertical handovers among heterogeneous networks having different RAT technologies under the constraints of network conditions and user preferences, which can be designed on the device-side, network side or in a hybrid manner [164].
9) Learning for network traffic control (such as routing) in ultra-dense heterogeneous networks to alleviate the issues of computational efficiency and scalability of the existing approaches [165].
10) Learning link quality/reliability to adapt the transmission parameters (such as MCS, transmission slot, received power etc.) of a wireless link. [18, 19, 171–173]
11) Learning to extract user activity/mobility patterns, temporal, spatial and social correlations from raw unstructured/semi-structured data coming from massive number of sensors.

B. Overview of Existing Machine Learning Techniques

ML techniques learn necessary information either from the available data-sets or by interactions with the surrounding environments and make suitable decisions on future actions to be followed by the agents either based on some learned models or in a model-free manner. The classification of existing machine/deep learning techniques in terms of different bases such as learning principles, objectives and the employed algorithm is provided in Fig. 6 [165, 166]. On the basis of the employed learning mechanisms, the existing ML and deep learning techniques can be broadly categorized into the following [45, 165, 178]: (i) supervised learning, (ii) unsupervised learning, and (iii) reinforcement learning. Figures 7 and 8 provide the illustrations of the principles of these three categories of ML techniques [179]. Furthermore, in Table IX, we have pointed out the main advantages and disadvantages of supervised, unsupervised, reinforcement learning and deep learning techniques along with the highlights on their applicability in IoT/mMTC scenarios.

The first category of ML techniques, i.e., supervised learning requires the need of training data to be labelled and the output of the algorithm needs to be already fed to the machine. Being aware of the output, the learning agent builds a model to move from the input to the output guided by the input training set. Based on the employed learning algorithms, the supervised learning techniques can be classified into [63, 165]: Artificial Neural Networks (ANNs), Deep NNs, Bayesian Networks (BNs), Support Vector Machine (SVM), Deep Learning (DL), Case-based Reasoning (CBR), Decision Trees (DTs), K-Nearest Neighbor (KNN), Instance-based Reasoning (IBR) and Naive Bayesian Classifier. The main difficulty of applying supervised ML techniques in IoT scenarios is that they require the processing of extensive data-set to learn from the dynamic environment but IoT devices are limited in terms of computing and caching/memory resources [45].

On the other hand, the second category of ML techniques, i.e., unsupervised learning does not require the need of labels of data-sets and is more complex than supervised learning techniques in terms of the computational cost [165]. Although this learning class has not been widely used as compared to the supervised learning in the current context, it can be considered as a promising future ML paradigm since the main objective of ML is to make the learning agent capable of learning without any supervision or human intervention. Since the learning agent does not have any training data-set and the knowledge of the output, the learning process is quite complex as compared to the supervised counterpart. This learning approach divides the unlabelled heterogeneous data into smaller homogeneous sub-sets which can be easily understood and managed [63]. Therefore, unsupervised learning techniques are mainly based on following three objectives [165]: (i) clustering, (ii) dimensionality reduction, and (iii) density estimation. For clustering-based unsupervised learning, different ML algorithms such as K-means, spectral clustering, Principal Component Analysis (PCA) and Dirchlet processes can be utilized. Similarly, unsupervised learning for dimensionality reduction can be employed by using ML

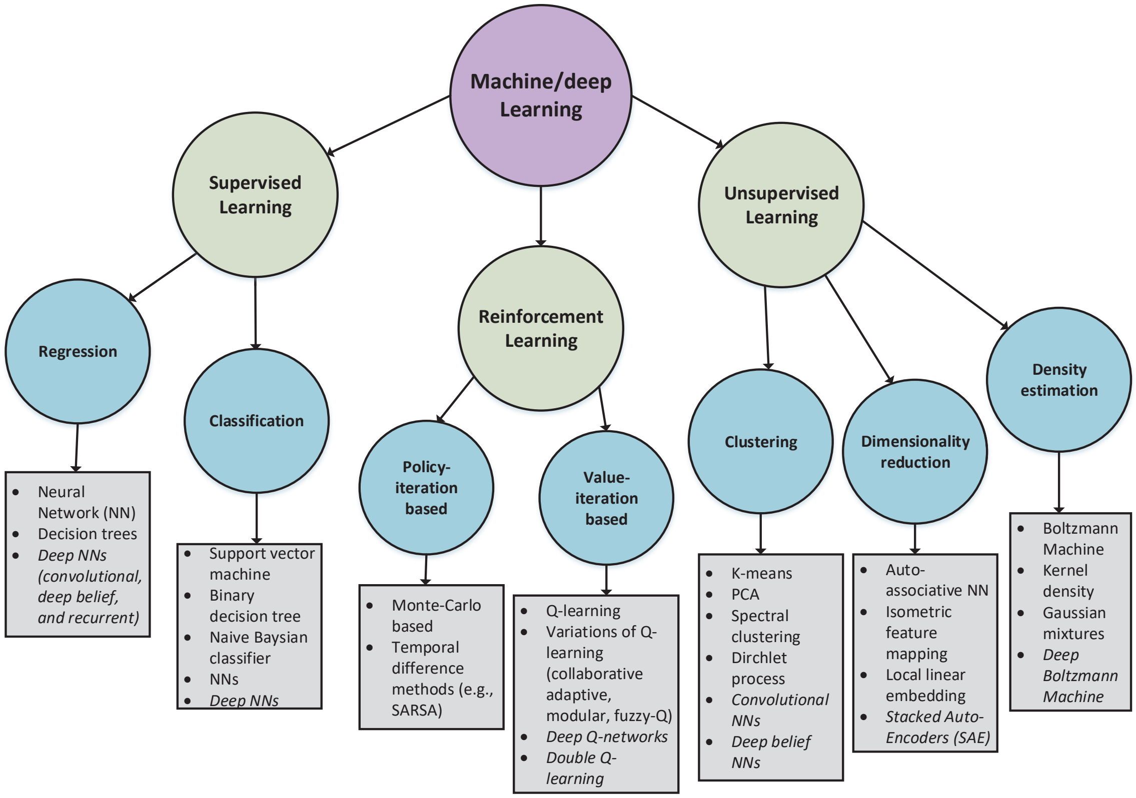

Fig. 6. Classification of existing machine/deep learning techniques (Deep learning techniques are indicated in the italic format).

algorithms like auto-associative NN, local linear embedding and Isometric Feature Mapping (ISOMAP). Furthermore, the density estimation-based unsupervised learning can be realized by using ML algorithms such as Kernel density, Gaussian mixtures, Boltzmann machine and Deep Boltzmann machine.

In order to take the advantages of supervised and unsupervised learning, there is another intermediate type of ML techniques, called semi-supervised learning technique [180] which uses both the labelled and unlabelled data for learning. The main objective of semi-supervised learning scheme to improve the learning accuracy by utilizing a small amount of the labelled data along with the large amount of the unlabelled data. The above-mentioned types of ML techniques have been applied mainly in the centralized framework [178]. For example, cloud-based processing can enable the operation of big data analytics to handle the massive amount of data gathered from the heterogeneous IoT sensors/devices [21].

The third category of learning techniques, i.e., RL enables a number of agents to interact with the environment and involves an environment of states, actions to be taken by the agents, state transition functions, an immediate reward and an initial observation function [45]. In this approach, an agent learns from the previous experience in the absence of the training data set. This learning mechanism deals with finding a proper balance between exploration for the random actions and exploitation of current knowledge, i.e., exploration-exploitation trade-off. The role of exploration phase in RL learning is to attempt some random actions towards searching better rewarding actions, while the exploitation phase attempts to utilize the previously learned utility to maximize the reward for the agent [181].

Besides the above-discussed three categories of ML techniques, some researchers have also considered another category of learning techniques, called as Sequential Learning (SL) [45], which helps the autonomous agents to learn the true underlying state of the environment having binary states. In this approach, the agents learn the system state in the sequence by following a given order while observing the environment and the actions or observations of previous agents, and then eventually converge to a true underlying state with repeated hypothesis testing [45]. The main advantage of employing SL techniques in the IoT systems is its flexibility in terms of memory requirements since it enables the convergence of finite memory SL in the resource-constrained IoT devices [177]. However, the main drawback of the SL approach is that it relies on direct communication links among MTC devices since the information required for SL comes from other agents, thus leading to the requirement of additional network resources.

Among several RL techniques, Q-learning requires low computational resources for its implementation and does not require the knowledge of the model of the environment, thus being a suitable learning technique for the resource-constrained IoT devices [181]. Furthermore, it is possible to implement this technique in a distributed way. Therefore, in the following subsection, we utilize the Q-learning technique to address the problem of RACH congestion in cellular IoT networks.

In addition to the aforementioned classical ML techniques, several Deep Learning (DL) techniques are being investigated

TABLE IX
ADVANTAGES AND DISADVANTAGES OF SUPERVISED LEARNING, UNSUPERVISED LEARNING, REINFORCEMENT LEARNING AND DEEP LEARNING ALONG WITH THEIR APPLICABILITY IN IOT/MMTC SYSTEMS

Learning paradigms	Advantages	Disadvantages	Applicability to IoT/mMTC environments
Supervised Learning	1. Data analysist has a full centralized control. 2. Output is known in advance to learn the model. 3. Suitable for learning problems where each input data point is labelled or belongs to a category.	1. Need of labelled data 2. Need of large training data-sets 3. Demands for high computational capacity.	1. The main supervised learning techniques investigated for IoT applications in the literature include: Artificial Neural Networks (ANNs), Bayesian networks, case-based reasoning, instance-based learning, ensembles of classifiers, decision trees, K-nearest neighbor, and support vector machines [62]. 2. Supervised learning is difficult to implement in distributed environments and in resource-constrained IoT/mMTC devices.
Unsupervised learning	1. No need of labelled data 2. Attempts to find hidden structure in the unlabeled data 3. Minimizes the human error that may arise in supervised learning. 4. Suitable for large and complex models where a data-point does not have a label or does not belong to a category.	1. Only the input data-set is known and no prior knowledge of training data-set and expected output. 2. The learning objective is of more subjective nature than that of the supervised learning. 3. Data analysist does not have much control.	1. The main unsupervised learning techniques studied in the context of IoT applications include clustering, association rule learning and ANNs [62]. 2. Unsupervised learning techniques are preferred over supervised learning in IoT applications which need faster results and require to extract hidden patterns from the massive IoT data.
Reinforcement Learning	1. No need of labelled data-and desired output 2. Low computational complexity compared to supervised and unsupervised learning 3. Easier distributed implementation and applicable for asynchronous network operation 4. Explores the trade-off between exploration and exploitation 5. Suitable for the learning problems having continuous interaction between the learning agent and environment via action-reward feedback loop.	1. No prior knowledge of underlying learning environment and feedback limited to only reward signal 2. Convergence to the steady state can be time consuming 3. The learning agent's observations depend on its actions and may contain strong temporal correlations. 4. May have to suffer from credit assignment problem in distributing credits to success among many involved decisions.	1. Simplicity of operation and distributed implementation make RL more applicable for IoT/mMTC applications 2. The continuous interaction between the learning agent and the environment in terms of action-reward feedback loop makes RL suitable for dynamic wireless IoT/mMTC environments. 3. The main RL technique used for IoT/mMTC applications is Q-learning.
Deep Learning	1. Reduces the part of feature extraction/engineering which is the most time consuming part of classical ML. 2. It is highly configurable and flexible than classical ML. 3. DL techniques can achieve much higher learning accuracies than the classical ML. 4. The performance scales much better with a large amount of data than with the classical ML.	1. Involves several hyper parameters and training process is slower. 2. DL techniques are sensitive to the data structure and size. 3. Lack of theoretical tools for determining the topology, training method and hyper-parameters. 4. DL techniques demand for high-end GPU platforms for training with big data in a reasonable time. 5. Interpretation of results with DL models becomes often difficult.	1. The main DL techniques considered in the literature for IoT/mMTC applications include deep CNNs, Recurrent Neural Networks (RNNs), deep belief networks, Long Short-Term Memory (LSTM) networks and deep Boltzmann machine [62]. 2. Due to its multilayer structure, DL is advantageous for extracting accurate information from raw IoT data in complex IoT systems. 3. DL can be utilized for privacy preservation of IoT systems since the intermediate data usually have different semantics than that of the source data [35]. 4. Due to huge need for energy, battery and memory resources, the deployment of DL in resource-constrained devices becomes difficult, leading to the need of efficient and compressed DL models [34].

in wireless networks including IoT networks [35]. Although feed-forward NN models have been employed in the past, the limitation in the computational capabilities of the available hardware was the major bottleneck in implementing more deeper architectures. The recent advances in Graphics Processing Units (GPUs) and hardware accelerators have led to the development of different DL algorithms and architectures. The main advantage of DL architectures over the classical ML is that DL models comprise of several hidden layers with the innermost layer being capable of recognizing more complex features, and can learn hidden features from the raw data-set.

As illustrated in Fig. 6, the existing DL techniques fall into three categories, i.e., supervised, unsupervised and RL. Some of the important DL techniques under the category of supervised learning include Convolutional Neural Network (CNN), Deep Belief Network (DBN) and Recurrent Neural Network (RNN). Similarly, deep Q learning and deep Q-networks fall under the category of deep RL techniques. Furthermore, the unsupervised DL incorporates various techniques including CNN, DBN, Stacked Auto-Encoders (SAE) and Deep Boltzmann Machine (DBM). Furthermore, depending on the objective of how a learning architecture will be utilized, DL architectures can be broadly categorized into [35, 166]: generative, discriminative and hybrid. The main objective of a generative DL architecture is to extract the high-order correlation features of the input data for further analysis while a discriminative DL architecture aims for pattern classification or recognition. On the other hand, a hybrid DL architecture combines the advantages of both generative and discriminative architectures.

As compared to the classical NNs, CNNs provide much better performance in terms of learning task specific features due to its deeper network structure. The CNNs can be trained by using supervised or unsupervised approaches but the supervised approach requires a large number of input-output pairs and a very large training data-set [166]. Also, as the CNNs become more deeper, there arises the need of a large-scale data-set and massive computing power for training the learning models. Thus, a proper trade-off between the computational complexity and the depth of the network needs to be investigated.

The RNNs are mostly used in the sequential or time-series problems with various length in which future prediction is dependent on several previous samples. The depth of an RNN can be adjusted to be as large as the length of the input data sequence, and the extended version of backpropagation algorithm, called Backpropagation through time [167] can be utilized to train the network. The main issue with the RNNs is that the long-term information has to sequentially pass through various layers before it reaches to the final layer and there may arise the issue of vanishing gradients across the layers [168]. This issue can be addressed partially with the Long-Short Term Memory (LSTM) version of RNNs and the modified stochastic gradient descent methods [169].

Another DL technique Stacked Auto-Encoders (SAE) belongs to the category of unsupervised learning and can be utilized to perform dimensionality reduction or data compression [166]. Its architecture comprises of an input layer followed by a small number of hidden layers capable of encoding the input data-set and an output layer which attempts to reconstruct the input layer. Furthermore, DBM is another most commonly used DL technique which comprises of stacked Restricted Boltzmann Machines (RBMs) and enables the efficient training of several layers of hidden units. Furthermore, some examples of the deep version of RL, i.e., deep RL include the deep Q networks and double Q-learning. Deep Q network, already patented by Google, combines Q-learning with a deep NN and can address the overestimation issue of Q-learning in certain conditions. In addition, double Q-learning algorithm can be generalized to any arbitrary function approximation to enhance the performance of deep Q-networks [170].

C. Learning Techniques for IoT/MTC Systems

The main challenges of applying learning techniques in an IoT environment include the following [45].

1) MTC devices have low computational capability, however, the widely-used ML techniques such as RL and decision trees can be computationally complex to implement.
2) Because of the distributed nature of IoT devices and high energy required to maintain constant communication with the BS/centralized aggregator, distributed learning needs to be investigated for an IoT environment.
3) Due to limited radio resources and energy constraints, only the limited amount of information is available at the IoT devices, and therefore, it is necessary to adapt the learning mechanisms based on the limited amount of information.
4) In some critical applications such as eHealthCare and industrial control, IoT devices need to learn quickly in order to satisfy the ultra-reliable and low-latency requirements. For this purpose, learning time should be as small as possible to quickly adjust the performance parameters.
5) To enable the harmonious coexistence of MTC and HTC systems, learning techniques should consider both the existing traffic as well as the new traffic from the MTC devices.

Existing works have applied learning techniques in the context of MTC/IoT in the following ways: (i) adaptation of an access control parameter, i.e., access barring factor to minimize the RACH overload [23], (ii) learning a dedicated slot within the MTC transmission frame by using an intelligent slot assignment strategy to avoid the collisions of access requests [11], (iii) BS/eNodeB selection by using Q-learning/RL techniques to minimize the access network overload [24, 25], and (iv) sequential learning with finite memory in order to learn transmission parameters under stringent memory and computational constraints [26, 177]. In Table IX, we highlight the main aspects related to the applicability of supervised, unsupervised , reinforcement and deep learning in IoT/mMTC scenarios.

Due to resource-constrained nature of IoT devices in terms of computational capability and power, it becomes challenging to employ supervised and unsupervised ML algorithms which are usually computationally difficult to implement and require the centralized mode of operation. Therefore, for wireless IoT applications, it is crucial to investigate ML algorithms which are computationally simpler and also have the distributed nature of operation [45]. To this end, RL can be a promising solution due to its simplicity of operation as well as its distributed implementation feature. The RL techniques are considered to be computationally simpler since the function used for deciding the next action is simple, for example, the operation of Q-learning is purely algebraic. Also, an RL technique exploits the interaction between the learning agents and the underlying environment and a learning agent can utilize future rewards to decide about the current action to maximize its long-term rewards [182]. Furthermore, RL techniques can work in model-free wireless environments where the system dynamics are usually unknown [183]. More importantly, the action-reward feedback loop of RL enables resource-constrained IoT devices to interact with the environment continuously without requiring the need of a supervisor, a training data-set and large input data-sets as required in the supervised learning. Although an RL technique requires the knowledge of state transition function and it has slow convergence, its unique feature of action-reward feedback to the agents makes this suitable for the applications in IoT/mMTC systems.

In an RL technique, a learning agent interacts with the underlying environment and alters the state of the environment by taking some actions. Then, based on the executed action, the environment provides some reward to the agent which then attempts to maximize its rewards over time by choosing those actions which result in higher rewards, thus leading to the unique aspect of action-reward feedback loop [182]. The learning agent's knowledge about the environment is reinforced during this learning process, i.e., the agent learns to compromise with the rewards and risks from their past experience.

Existing IoT related works have already exploited the action-reward feedback to employ an RL technique in different settings. Authors in [184] employed RL for the real-time task assignment among fog/edge servers in a fog IoT network consisting of 200 IoT devices and 10 fog servers of heterogeneous capabilities for task processing with the objective of minimizing the total computation latency over a long period. This RL approach exploited the pattern of the state-reward pair (with the inverse of the latency as the reward) to reinforce the task assignment action towards minimizing the long-term latency and it was shown that the proposed RL based method overcomes the conventional approach by about 16 % in terms of long-term latency. Furthermore, another article [185] employed an RL-based scheduling algorithm for access control problem in edge-IoT systems, and also to minimize prediction error in a battery prediction problem without any knowledge of the model about the energy source and the arrival process. In the access control problem, the BS employs the RL-based Long Short-Term Memory (LSTM) Deep Q-Network (DQN) with all the possible User Equipment (UE) selection choices as the action space, and the received sum-rate at the BS as the reward signal with the objective of maximizing the long-term expected uplink sum-rate. In the joint problem of action control and battery prediction, the mixture of sum-rate and the prediction loss was considered as a reward signal.

Furthermore, ML techniques can be employed to enhance the security, resiliency and robustness of IoT/mMTC systems. Compared with the conventional communication systems, IoT systems are usually more vulnerable to various security issues including intrusions, spoofing attacks, jamming, malwares, eavesdropping, Denial of Service (DoS) attacks and distributed DoS [27]. Due to the limited computation power, memory, battery resources and operating bandwidth, execution of computationally intensive and latency inducing tasks for security provisioning at the IoT devices becomes complicated. However, most of the legacy security solutions generate heavy communication and computation overhead for the IoT devices. In this regard, authors in [28] identified various IoT security threats and reviewed several ML-based techniques IoT security solutions including IoT authentication, secure offloading, access control, anti-jamming and malware detection schemes. The main security threats in IoT include DoS attackers, jamming, spoofing, man-in-the-middle attack, software attacks and privacy leakage.

Various techniques based on supervised, unsupervised and RL have been investigated to improve the security of IoT systems. For example, SVMs can be used in IoT devices to detect network intrusion [178] and spoofing attacks [30]. Also, an IoT device can utilize K-NNs and random forest classifiers to devise a malware-detection model [29]. In addition, Neural Networks (NNs) can be utilized to detect network intrusion and in this context, authors in [31] have provided a comprehensive review of ML and data mining techniques for intrusion detection in cyber analytic applications. Furthermore, IoT devices can utilize different unsupervised techniques including multivariate correlation analysis to detect DoS attacks and infinite Gaussian mixture model for physical layer authentication [28]. Moreover, various RL techniques such as Q-learning, Dyna-Q and deep Q-network can be utilized to enhance the security of IoT protocols against various attacks. As an example, authors in [32] showed that Q-learning based offloading policy can reduce the spoofing rate by 50% and jamming rate by 8% as compared to the conventional approach without learning.

In addition, to address the scalability and accuracy issues of classical ML techniques, emerging deep learning techniques such as Long Short-Term Memory (LSTM) networks seem promising. The LSTM networks not only reduce the burden of feature engineering over the classical ML but also are found to be resilient against adversaries with high detection accuracy [33]. Furthermore, unlike the classical ML, LSTM networks are suitable for training unstructured data-sets encountered in IoT applications and also can be utilized to recognize the repetitions of the attack patterns in the long sequence without depending on a defined window size. Moreover, design and implementation of robust attack detection systems for IoT devices becomes challenging due to several issues including latency sensitivity, and resource constraints. Although the remote cloud can address the issue of computational and caching resources, it suffers from the problem of high reaction time and needs expensive transmission links to transmit to/from the cloud. In this regard, authors in [33] utilized the self-learning capabilities of deep learning techniques to detect cyber-attacks by utilizing fog nodes as data and control processing centers in fog-to-things computing platforms. In addition, the recent article [34] utilized the Reservior Computing (RC) paradigm, which is considered promising to address the training difficulties of the conventional recurrent NNs, for the purpose of attack detection in smart grids. The accuracy of the proposed RC-based attack detection has been shown to be insensitive to attack variations including the number of compromised meters and the attack magnitude.

The emerging deep learning techniques enable the extraction of accurate information from raw IoT data in complex wireless networks. Also, due to multi-layer structure of DL techniques, they become advantageous for complex edge computing environments [35]. Furthermore, DL can enhance the privacy preservation in IoT systems since the intermediate data in deep learning usually have different symantics than that of the source data [36]. However, DL techniques demand for a significant amount of energy, battery and memory resources and their implementation in resource-constrained devices becomes challenging. Therefore, it is crucial to investigate suitable techniques to make DL techniques suitable for resource-constrained IoT devices. Some promising techniques in this direction include the following [35]: (i) network compression, (ii) approximate computing, and (iii) accelerators.

Network compression enables the conversion of a dense network to a sparse network, thus reducing the storage and computational capabilities of Deep Neural Networks (DNNs). In this regard, authors in [37] investigated the applicability of DL models on different hardware devices including Intel Edison used in wearables, Snapdragon 800 employed in some models of smartphones and Nvidia Tegra K1 used in IoT-enabled vehicles by taking measurements of energy consumption, time and memory footprint into account. Based on the energy usage measurement, the performance of Convolutional Neural Networks (CNNs) and DNNs was investigated in the aforementioned three different hardware platforms and it was shown that all the platforms were able to run the compressed models of DL. In the context of CNNs, DL models can be enhanced for resource-constrained IoT devices by replacing convolutional layers with the feed-forward layers wherever possible. Similarly, DL models with DNNs can be made implementable in resource-constrained devices by reducing the number of involved parameters, i.e., by removing the redundant parameters, and the models can be made more time-efficient by selecting a suitable activation function in DNNs.

Another enabling approach to make DL models implementable for IoT devices is to integrate DL models with approximate computing [38]. The benefit of approximate computing arises due to the fact that in many IoT applications, the acceptable range of ML results with the desired quality works without requiring the need of obtaining exact ML prediction results. Furthermore, another promising method to enable the implementation of DL models in IoT devices is to design particular accelerators in the form of hardware and circuits with the objective of optimizing memory footprint [39] and energy efficiency [40] of DL models.

In several IoT applications including energy usage prediction, demand side management and control, and non-intrusive activity detection, time-series methods can be utilized [42]. The LSTM is one of the popular DL architectures in the literature for addressing time-series problems and has been investigated for several IoT applications [41, 42]. An LSTM network consists of long short term memory blocks conprising of memory cell units which enable LSTM to remember the state values for the arbitrary period of time. In [41], authors employed an LSTM model to predict the working conditions of the power station equipments by analyzing the data collected from sensors. The performance of the LSTM model was found to be better as compared with that of the Autoregressive Integrated Moving Average (ARIMA) model in terms of the prediction accuracy. Furthermore, the contribution in [42] compared the performance of LSTM and ARIMA models in predicting the number of occupants at a given time and location by using WiFi networks in a smart building environment. Via numerical results, it was demonstrated that LSTMs can achieve low root mean square error than the ARIMA models but often require more training data and need to tune several input parameters. Moreover, the article [33] proposed an LSTM network for detecting distributed cyber-attack in edge IoT networks by utilizing the feature resilient feature of DL against morphing attacks. Also, authors identified and analyzed various critical threats and attacks for IoT devices and demonstrated the effectiveness of DL models over the classical ML models in addressing the attack issues in wireless IoT systems.

D. Q-learning for RACH Congestion Problem

The ML techniques can be utilized to address various non-conventional challenges of IoT/mMTC systems highlighted in Section I. In Section V, by considering RAN congestion problem as a use-case example, we discussed various existing solutions and emerging solutions including learning-assisted techniques. Herein, we employ Q-learning for RACH congestion problem as a use-case illustration of the application of ML techniques for RAN congestion problem in cellular IoT networks.

The main problem with the existing contention-based RA schemes is that the occurrence of collisions is unavoidable and the achievable RACH throughput in the presence of massive access requests/loads gets significantly reduced [11]. For example, the maximum throughput of the widely-used slotted ALOHA technique is e^{-1} (37%). Furthermore, because of low RACH throughput and the employed backoff strategies, the aggregated traffic including both the newly generated and the retransmitted exceeds the RACH channel capacity at a certain point, thus making the system unstable. Although Slotted ALOHA can work well with the conventional cellular/HTC traffic despite the instability issue, the support of M2M traffic becomes problematic due to infrequent and massive number of access requests, thus causing the problem of RACH overload.

Learning techniques can be employed at the MTC devices in order to enable them to learn to avoid concurrent transmissions during the RACH contention period without any assistance from the central entity. After a learning technique achieves its convergence, each MTC device can get a unique dedicated slot, thus avoiding collisions among their transmissions. In contrast to the conventional centralized Q-learning approach, for example, Q-learning based ACB control in [23], we focus on providing a framework for the distributed Q-learning mechanism in which each MTC device attempts to find its unique time slot for its transmission towards minimizing the RACH congestion. In the following, first, we present a framework for the RL with a single MTC device and then develop the formulation of Q-learning in the considered context.

The environment perceived by an MTC device can be usually described by a Markov Decision Process (MDP) and a finite MDP can be denoted by a tuple $< X, U, f, \rho >$, where X represents the finite set of environment states, U is the finite set of device actions, f denotes the state transition probability function and captures the environmental dynamics, and ρ is the reward function [64]. In this MDP modeling, a state parameter $x_t \in X$ indicates the characteristics of the environment at the tth time instance. At each time-step, the device can change its state by taking actions $u_t \in U$ and due to this action, the environmental state alters from the current state x_t to some other state x_{t+1} based on the employed transition probability function $f(x_t, u_t, x_{t+1})$. During this transition, the device receives an instantaneous reward $r_{t+1} \in R$ with the defined function ρ, i.e., $r_{t+1} = \rho(x_t, u_t, x_{t+1})$.

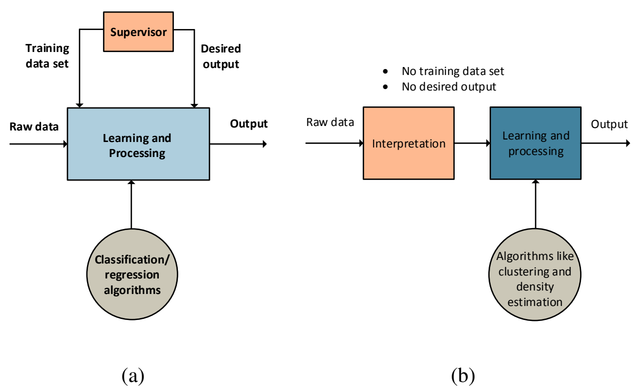

Fig. 7. Illustrations of the principles of supervised and unsupervised learning.

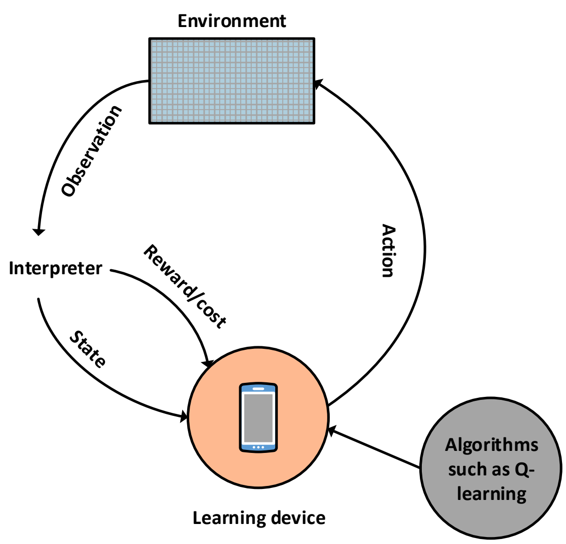

Fig. 8. Illustration of the principles of a reinforcement learning technique.

Given a state, the device chooses its action based on its policy π and the policy can be either stochastic or deterministic. Each time the device applies a policy, it accumulates some rewards from the environment, resulting in the return of $\sum_{l=0}^{L}\gamma^l r_{l+1}$, where $\gamma \in [0\ 1]$ is the discount factor which provides more weights on the immediate rewards and L denotes the length of one episode [65]. At each time-step, the learning device aims to maximize the expected discounted return in the long-term, i.e., long-term reward, given by [64, 65]: $R_t = E\left(\sum_{l=0}^{L}\gamma^l r_{t+l+1}\right)$, where E denotes the expectation operator and this is taken over probability state transitions, i.e., dynamics of the considered environment.

In this RL process, the learning device attempts to maximize its long-term performance, while only receiving feedback about its immediate action, i.e., one-step performance. To achieve this, the device needs to compute an optimal action-value function, known as a Q-function. Given a certain policy π, the expected return of a state-action pair, $Q^\pi(x,u)$ is given by

$$Q^\pi(x,u) = E\left(\sum_{l=0}^{L}\gamma^l r_{t+l+1}|x_t = x, u_t = u, \pi\right). \quad (8)$$

Subsequently, the optimal Q-function can be written as: $Q^*(x,u) = \max_\pi Q^\pi(x,u)$ and it satisfies the well-known Bellman optimality equation [64, 187].

One simplest way of choosing a future action by a device is to employ a greedy policy which selects the action with the highest Q-value at every state as follows: $\pi(x) = \arg\max_u Q(x,u)$.

Among different ML techniques listed in Fig. 6, Q-learning is simple to implement and uses a look-up table with the Q-values representing the utilities for state-action pairs. The utility of taking an action 'u' in a state 'x' is denoted by $Q(x,u)$ and can be calculated as the expected value of the sum of immediate reward and discounted utility of the resulting state after executing the action 'u'. In the Q-learning process, the current estimate of Q^* value, i.e., $Q_t(x_t,u_t)$ is updated by using the estimated samples computed by relating with the actual experience from the execution of the action, in the form of the pairs of subsequent states (x_t, x_{t+1}) and the rewards r_{t+1}. In this way, Q-learning involves an iterative procedure in which the learning agent implicity assumes an initial condition before making an update, i.e., the agent starts with an arbitrary fixed Q-value. Then, at each time instant, the agent chooses an action, observes a reward, and enters into a new state which may depend on both the selected action and the previous state, and subsequently the Q-value is updated. This iterative process continues till the Q-value becomes stable [187].

To provide an example framework for the application of Q-learning in an MTC scenario with N number of devices, we consider a frame-based slotted ALOHA scheme as in [186, 188], in which a frame is divided into K number of access slots. Each MTC node has individual Q values corresponding to every slots in the frame and these values are updated based on the outcomes of transmission, i.e., success or failure. At the start, all the MTC nodes can start with zero or random Q values, learn gradually via their transmissions, and then finally reach to the optimal transmission strategy after finding unique RA slots for their transmissions.

Let Q(i, k) indicate the preference of the i-th node to transmit a packet in the kth RA slot. After every data transmission, the new Q value is updated based on the previous Q value and the current reward. A positive reward can be assigned for each successful transmission while the reward becomes negative for each failed transmission. At each instance of transmission, the node selects a slot with the highest Q value and in case of two or more maximum values, a random selection approach can be applied [188].

Regarding the selection of learning rate in the Q-learning process, the higher the value, the faster will be the convergence of Q value towards the reward value. However, a small value of learning rate is usually preferred in practice to enhance the robustness with respect to infrequent collisions caused by channel variations [186]. Furthermore, although the learning rate is usually considered fixed in the existing works, it is typically time varying in nature, decreasing with time and each state-action pair may be associated with a different learning rate [64].

E. Exploration Strategies for Q-Learning

Q-learning aims at finding an optimal policy to select an action in the current state and one of the design aspects for the Q-learning algorithm is to balance the exploration-exploitation tradeoff. The exploitation is performed by executing one of the actions which maximizes $Q(x, u)$ whereas the exploration is carried out by randomly selecting an action to build a better estimation of the optimal Q-function.There are several strategies to create this balance and the widely-used three strategies are described below [189, 190].

1) **ϵ-greedy strategy**: This is the most commonly used exploration strategy in which the Q-learning algorithm utilizes a parameter $0 \leq \epsilon \leq 1$ to decide on the action to follow. The algorithm chooses the action with the highest Q-value in the current state with $(1 - \epsilon)$ probability and a random action with the probability ϵ. The value of the parameter ϵ can be varied over time as the learning progresses. The main drawback of this approach is that it treats all the possible actions equivalently during its exploration by choosing an action uniformly from the set of possible actions.
2) **Soft-max strategy**: To address the drawback of ϵ-greedy stage during the exploration phase, this soft-max strategy uses either a Gibbs or Boltzmann distribution, in which the learning device at a state x selects an action u with the probability of $\pi(x, u) = \frac{e^{\frac{Q(x,u)}{\tau}}}{\sum_u e^{\frac{Q(x,u)}{\tau}}}$, where $\tau > 0$ denotes the temperature parameter of the Boltzmann distribution and depicts how randomly the values are chosen. For example, $\tau = 0$ case represents no exploration at all while $T \to \infty$ case reflects the scenario in which the learning device chooses the action values almost with equal probability.
3) **Optimism in the face of uncertainty**: This approach encourages exploration by assigning higher initial values to the Q-function. However, the convergence time will be increased since the estimated Q-values can be quite bad estimates of the actual Q-value and these bad estimates may last longer during the learning process. Also, in the presence of dynamic uncertainty, this technique is not useful since the exploration mostly occurs at the beginning of the learning process.

F. Performance Enhancement of Q-Learning

In the following, we present several variants of Q-learning techniques which aim towards improving the performance of the ordinary Q-learning.

1) Collaborative Q-learning: In this subsection, we describe the need of collaborative Q-learning in ultra-dense IoT networks and review the related works. The traditional Q-learning algorithm which is based on single state-action may not be suitable for the multi-agent environment with multiple policies. In this regard, collaborative learning can be utilized by exploiting the overall/global objective/reward of the collective environment instead of the reward for a single learning device. Furthermore, the global reward can be utilized in addition to the individual reward to improve the performance of the existing learning schemes.

In a multi-agent environment like the case of ultra-dense IoT networks, if all agents keep mappings of their joint states and actions, this will require each learning device to maintain very large Q-value tables, whose sizes are exponential in the number of agents [191]. This becomes difficult even in the case of a single state case. For instance, when M number of learning devices play the repeated game with only two actions, the size of the table becomes 2^M. Therefore, it requires more state space, information space and action space. Furthermore, in the multi-agent case, state transitions, the instantaneous rewards and future expected return are based on the joint action of the multiple agents. In addition, instead of the individual policies, a joint policy formed by individual policies will impact the Q-function. Therefore, the design of state-action pairs or the optimal policy to maximize the overall reward becomes the joint/collaborative problem in a multi-agent environment, thus leading to the need of collaborative learning.

Some existing works have applied collaborative Q-learning in different settings [43, 191, 193, 194]. One way of realizing collaborative Q-learning is to estimate the belief of the opponent and the environment knowledge instead of the Q-value function in a way that the learning agent does not need to observe the opponents' reward and their Q learning parameters [191]. In this regard, the authors in [192] applied the collaborative Q-learning framework to optimize the waiting time in intelligent traffic control applications. Furthermore, the authors in [43] employed a collaborative ML technique at the edge computing devices with the objective of extracting the statistical relationships among the contextual information collected by the edge devices and constructing predictive models to maximize the communication efficiency. With this edge-centric learning approach, only the inferred knowledge can be transferred to the network instead of transmitting all the raw contextual information. Furthermore, the authors in [193] used collaborative Q-learning in finding an optimal path between any starting point and a target in a grid environment for a mobile robot. In contrast to the conventional approach of using a single Q-table, the work in [193] exploited the use of two Q-tables, i.e., one local Q table and another Master Q table.

Moreover, in a multi-agent environment, it is necessary for a learning agent to keep the track of its environment, as well as other agents' actions. In such a multi-agent environment, rather than considering individual actions of the agents as in the ordinary Q-learning, the joint actions need to be considered while devising a learning strategy. Instead of the Q-value used in the ordinary Q-learning, a Nash Q-value function should be defined [196], and the convergence of Q-learning algorithm with the Nash Q-value function becomes slower since the number of agents increases due to the resulted increase in the joint action set [195].

2) Situation-Aware Adaptive Q-learning: The collisions of the RA requests caused due to concurrent transmissions of multiple RA requests in one RACH sub-frame results in higher access delay since the devices have to retransmit their RA requests. Although the average access delay can be minimized by using a higher number of RACH sub-frames, the sub-frames available for data transmission will be reduced since the sub-frames allocated for RACH procedure can not be used for the data transmission purpose [197]. In this regard, it is crucial to balance the tradeoff between the radio resources allocated for RACH procedure and data transmission process. Furthermore, in many cases, the network usage pattern (the number of devices/users trying to connect to the network) is time varying in nature. In this context, a network should be capable of detecting the variation in the arrival rate of the access request and should adapt the number of sub-frames to be allocated for RACH accordingly [197]. In addition, it is important to balance a tradeoff between exploration and exploitation while adapting the Q-learning parameters to the dynamic uncertain environment [198].

Although Q-learning has been shown to converge and has been used in many fields including mechatronics control and robotics, it has some issues such as how to improve the convergence rate and to avoid the convergence in the local optimum [199]. To address these issues, three different parameters of the Q-learning technique, namely, learning rate α, discount rate γ and temperature parameter τ in Boltzmann distribution, should be dynamically adapted based on the dynamicity of the underlying learning environment. One of the widely used methods to adapt the learning parameters in various applications is fuzzy-logic based learning, which is briefly described in the following subsection along with the related literature.

3) Fuzzy-logic based Adaptive Q-learning: In the Q-learning process, the Q-values are usually stored in a look-up table but this storage process becomes infeasible in practice in the presence of a large number of state-action spaces and with the continuous state space [200]. Although Q-values can be stored by using feed-forward neural networks or self-organizing maps, the learning process becomes slower. In contrast, incorporation of Q-learning into fuzzy environments seems promising since fuzzy interference systems being universal approximations can be considered as good candidates to store Q-values and the prior knowledge can be provided to the fuzzy rules in order to significantly reduce the training part.

The implementation of Q-learning becomes impractical and even impossible in continuous state spaces [200]. In such cases, fuzzy-logic based approach helps to discretize the continuous state or action spaces into finite states by employing suitable fuzzy rules and also the speed of fuzzy-logic based Q-learning can be increased by incorporating the prior knowledge via fuzzy rules [201]. In other words, Fuzzy Q-learning discretizes continuous variables by using fuzzy labels and a fuzzy rule-based inference system is employed to find an action for these discretized states [202].

Other drawbacks with the ordinary RL are that it is difficult with the continuous states and behaviors in the real world environment due to discrete set of actions and spaces, and it becomes complex to learn the problem with multiple objectives [203]. To solve these issues, fuzzy-logic based rules can be employed to tune the learning parameters of the Q-learning technique towards making it more adaptive.

In the context of cooperative fuzzy Q-learning, authors in [202] utilized this learning approach to optimize the coverage and capacity of cellular networks by adapting the tilting of vertical antennas. The employed cooperative Fuzzy Q-learning mechanisms enable cooperation among the learning agents during the exploration phase and is fully distributed in the exploitation phase. The cooperation in the exploration phase is employed by utilizing the global reward of all the considered cells instead of the local rewards belonging to individual cells. This cooperation with the help of global reward helps to speed up the exploration phase of the Q-learning process while also allowing the learning agents to exploit the learned knowledge independently while selecting their actions. Furthermore, a self-learning cooperative strategy is developed in [199] by combining adaptive Q-learning with the fuzzy method for its application in robot soccer systems.

Moreover, the contribution in [204] recently proposed a fuzzy Q learning-based user centric backhaul-aware user association scheme in which the BSs broadcast their constraints and capabilities in terms of meeting the requirements of heterogeneous UEs in terms of the optimized bias factors. The employed fuzzy-logic based Q-learning scheme helps to dynamically adjust these bias factors based on network conditions and users' requirements in an automated and distributed manner.

Similarly, authors in [205] proposed a fuzzy Q-Learning based energy controller for a small cell powered by local renewable energy, local storage, and the smart grid to elongate the lifespan of the storage devices and to minimize the electricity expenditures of the mobile operators. The employed fuzzy Q-learning based controlled can be utilized without the prior knowledge of the mobile traffic demand, energy pricing and weather.

In order to avoid the inefficient and expensive manual tuning of cellular network parameters in 5G small cell networks, it is crucial to perform automatic configuration and optimization of the network parameters including the handover parameters. In this regard, authors in [206] employed a fuzzy logic controller based dynamic fuzzy Q-Learning algorithm for mobility robustness optimization in a heterogeneous network. With the proposed dynamic fuzzy Q-learning algorithm, the system learns necessary parameter values towards optimizing the call dropping ratio and handover ratio, and it has been shown that the Q-Learning algorithm can lower the handover ratio while keeping the call-dropping ratio at the lower level. Besides, by considering various parameters such as link quality, the available bandwidth, link quality, and relative vehicle movement, authors in [207] proposed a fuzzy constraint Q-learning algorithm for vehicular ad-hoc networks in order to evaluate the quality of a wireless link towards finding the optimal route.

4) Model-based Q-Learning: The main drawback of model-free learning is the convergence time. By predicting a model about the transition state probabilities, the performance of the learning techniques could be improved. In time-varying dynamic scenarios, it becomes advantageous to predict the environmental dynamics in the centralized entities such as cloud-center and to utilize the corresponding model to enhance the performance of distributed Q-learning at the resource-constrained IoT devices. In such a collaborative cloud-edge processing framework [21], the predicted model at the cloud-center can be communicated to the edge-side to improve the performance of distributed learning at the edge-side of the network.

Existing works have used model-based Q-learning in different applications such as robotic applications [208, 209] and wireless channel allocation [210].

G. Discussion on the conventional and learning-based solutions for RAN congestion

In Section V, we discussed various existing solutions for RAN congestion problem and some emerging solutions including learning-based solutions, distributed queueing, and SDN and virtualization-enabled techniques. In addition, we have presented the qualitative comparison of the main existing RA solutions for RAN congestion problem in Table VII in terms of the important performance metrics. In Section VI-A, we discussed the advantages of learning techniques in wireless communications along with various use-case examples from the existing literature, and in Section VI-C, we provided an overview of existing ML techniques along with their classification. Furthermore, in Section VI-B, we discussed various applications related to learning techniques in IoT/mMTC systems and in the later subsections, we presented a framework for the application of Q-learning for RACH congestion problem along with some performance enhancement techniques.

As highlighted in Table VI, there exist several schemes including back-off based scheme, ACB, EAB, and slotted random access to mitigate the RAN congestion in cellular networks. Most of these schemes are based on the adjustment of either the retransmission back-off value or an ACB parameter. Although the performance of these basic schemes in dense access networks can be enhanced with various approaches such as pull-based scheme, group-based RA scheme, code-expanded RA scheme and tree-based RA method, most of these techniques have been developed only in the context of the conventional cellular networks. The incorporation of IoT/mMTC devices in cellular networks brings various device-level and network-level challenges as highlighted earlier in Fig. 2. Furthermore, from the qualitative comparison presented in Table VII, it should be noted that a single RA solution is not perfect in terms of all the considered performance metrics and the selection of a technique for a particular application depends on the required trade-off among these metrics.

Due to highly bursty traffic and the massive number of contending devices to be managed, it becomes highly challenging to employ the conventional RA schemes in ultra-dense IoT/mMTC networks. Mainly, estimation of the number of devices and network status, and the adaptation of operating parameters such as transmission probability and access control parameter become difficult in ultra-dense dynamic IoT networks. As an example, the instability issue of the conventional slotted ALOHA RA scheme can be considered. This scheme effectively works for the case of H2H traffic due to the low dimensioning of the system and also due to regular pattern of the RACH access requests. However, the performance of this conventional slotted ALOHA RA scheme suffers while supporting MTC traffic due to the massive number of irregular and dynamic access requests, thus leading to the problem of RACH congestion [11]. Therefore, it is crucial to investigate suitable learning-assisted intelligent RA schemes which aim to address the issue of access network congestion and support massive connectivity in IoT/mMTC networks.

As described earlier in Section VI-A, learning-enabled techniques can assist towards the mitigation of RAN congestion problem in various ways. Some of the promising approaches include: (i) by learning unique time slots to be allocated for the contending MTC devices [11], (ii) by learning to adapt an access control parameter used in the ACB scheme [23], and (iii) by learning to associate MTC devices with suitable eNodeBs [24, 25]. As illustrated in [11] with the help of numerical results, Q-learning assisted RA scheme performs better than the conventional slotted ALOHA RA scheme in terms of the RACH throughput and average end-to-end delay. Furthermore, authors in [23] employed Q-learning algorithm to adaptively control the ACB factor to be assigned to MTC devices based on the previous experience obtained from the interaction with the environment and the performance of the proposed scheme. Via numerical results, it was shown that the success probability of accessing the RACH by MTC devices increases as the learning progresses over time.

VII. Summary of Lessons Learned, Research Challenges and Future Directions

In this section, we present the summary of lessons learned from this paper, existing research challenges and future directions under the following topics: (i) Cellular connectivity for mMTC/IoT Systems, (ii) Spectrum management for mMTC systems, (iii) Traffic characterization and modeling for mMTC systems, (iv) Random access schemes for ultra-dense IoT networks, (v) Distributed resource management in ultra-dense IoT networks, (vi) Device heterogeneity and grouping-based transmission schemes, (vii) ML applications in wireless IoT/mMTC systems, and (viii) Deep learning for emerging IoT applications.

A. Cellular Connectivity for mMTC/IoT Systems

Due to ever-increasing need to support the massive number of connected sensors and MTC devices, cellular operators have to face a number of unique device-level and network-level challenges in supporting MTC devices with the existing cellular networks. The main issues include diverse QoS requirements, RAN congestion, highly dynamic and sporadic MTC traffic, ultra-low device complexity, low battery lifetime, huge signalling overhead, small data packet transmission, network scalability, need for enhanced coverage, and efficient management of distributed computing, caching and communication resources. Also, MTC devices may cause harmful interference to the existing cellular users and may significantly degrade the system performance of LTE/LTE-A based cellular systems. Towards addressing these issues, some potential enablers identified in this paper include flexible waveform design, dynamic resource allocation techniques, advanced spectrum sharing techniques, clustering and data aggregation techniques, SDN and virtualization-based techniques, advanced RA schemes, constant envelope coded modulation schemes, CS-based MUD, MAC protocols with low signalling overhead, advanced transmission scheduling techniques, collaborative edge-cloud processing and energy-efficient techniques for the green IoT.

However, several challenges need to be addressed to support diverse QoS requirements (transmission reliability, latency, data rate and spectral efficiency) of MTC devices in wireless IoT systems due to mult-path fading, low reliability, co-channel interference and time-varying capacities. Also, most of the available physical layer-based capacity models do not capture the link-layer QoS requirements and the investigation of suitable QoS-based capacity models is necessary to characterize the performance of MTC systems. In this regard, this paper presented a mathematical framework to analyze the performance of MTC systems with the QoS support in terms of effective SNR, achievable rate and the estimated number of devices for the uplink access in Section IV.

Current protocols designed for cellular IoT such as NB-IoT and LTE-M are based on the assumption of the low-latency requirement. This requirement results in significant cost in terms of device price and system capacity [53]. Besides, the main issues involved in providing cellular connectivity to the low-power MTC devices include the device battery life, system capacity, coverage and cost. Due to low cost and low capability of MTC devices, a significant compromise in the link performance needs to be made resulting in the shrinkage of the coverage area. One way of compensating this coverage loss is to utilize an extended transmission time interval, which, however, will lead to the increase in the battery consumption. Therefore, it is crucial to balance the tradeoff between energy consumption and coverage expansion while designing transmission technologies for the MTC devices.

Furthermore, transmission schemes should be able to support a significantly large number of devices within a given bandwidth while ensuring higher battery efficiency. In this regard, the concept of effective bandwidth [53] could be utilized to achieve a good balance between the spectral efficiency for a given coverage and the corresponding transmission time. On one hand, the bandwidth allocated for a device should not be far less than the effective bandwidth to avoid the excessive transmission time. On the other hand, for a given sensitivity level of the device, it is preferred to have allocated bandwidth not more than the effective bandwidth in order to accommodate more number of devices in the saved bandwidth without significantly increasing the transmission time. Moreover, advanced transmission scheduling techniques and low signalling overhead MAC protocols need to be investigated to effectively support the massive number of MTC devices in the upcoming 5G and beyond cellular networks.

B. Spectrum Management for mMTC Systems

Since the available usable spectrum below 6 GHz is limited, it is crucial to investigate suitable spectrum sharing solutions for emerging 5G and beyond systems including eMBB, URLLC and mMTC. The most commonly discussed dynamic spectrum sharing solutions for 5G and beyond systems include interweave cognitive communications, underlay cognitive communications, carrier aggregation, Licensed Assisted Access (LAA), Licensed Shared Access (LSA) and Spectrum Access System (SAS) [93]. Another option is to explore higher frequency bands such as millimeter wave (mmWave) bands. However, due to propagation related and hardware imperfections issues in mmWave bands, it becomes challenging to operate mMTC devices in mmWave bands as they are constrained in terms of resources and are mostly deployed in indoor or underground environments where propagation issues could become problematic. Therefore, for mMTC applications, the frequency bands below 6 GHz is of more importance. In this spectrum region, there arises the issue of whether licensed or unlicensed band becomes suitable for mMTC applications as pointed out in the following.

In the spectrum region below 6 GHz, the operation in the unlicensed band by using carrier aggregation and LAA could be a cost-effective solution for mMTC applications since these frequencies are freely available to any device. However, uncontrolled interference conditions resulted from the free access from the massive number of devices and the lack of QoS guarantees may severely limit the ability of utilizing unlicensed bands for mMTC applications [4]. On the other hand, emerging techniques such as LSA and SAS provide better interference characterization due to the centralized control and seem more suitable for mMTC applications. Besides, some of the mMTC applications based on periodic data reporting such as smart metering can effectively work with the shared spectrum bands in a demand-based manner instead of exclusively assigning the portion of a licensed spectrum. Due to low bit-rate requirements of mMTC applications, the bandwidth requirement may not be significant, however, the exclusive licensed spectrum can be allocated to more demanding applications such as eMBB and URLLC. To this end, it is an important future research direction to investigate the feasibility of utilizing shared spectrum for mMTC applications.

In summary, one of the crucial challenges in mMTC systems is to address the issue of the scarcity of the available usable radio spectrum either by exploiting suitable spectrum sharing schemes or by exploring higher frequency bands such as mmWave bands. However, it may be challenging to operate mMTC devices in the mmWave bands due to propagation related issues and resource constraints of the devices. Out of licensed band and unlicensed bands below 6 GHz, unlicensed bands seems to be promising for the low-cost mMTC systems but the issue of uncontrolled interference may be problematic for QoS guarantee of some applications. Also, emerging dynamic spectrum sharing techniques such as LSA and SAS provide better interference characterization due to the centralized control and seem more suitable for mMTC applications.

C. Traffic Characterization and Modeling for mMTC Systems

Since traffic characteristics in IoT sensory networks usually depend on the application scenarios, traffic characterization is considered to be an important issue for the design and optimization of the network infrastructure. The mMTC networks may generate various types of traffic patterns such as PU, ED and streaming, and these traffic patterns may have different amplitudes, activation periods and starting times [73]. Besides, the data packets generated by MTC devices can be of varying sizes and bandwidth requirements. For example, the data packets generated by the temperature and humidity sensors are usually of small size in the order of bytes, whereas video monitoring devices can generate data sizes in the order of megabytes.

MTC devices have completely different QoS requirements than those of the conventional HTC devices, and also the nature of MTC traffic is significantly different than the HTC traffic. MTC devices usually have heterogeneous traffic patterns in terms of their starting times, amplitudes and duty cycles. Also, IoT/mMTC systems may need to support different heterogeneous applications including elastic, hard real-time, delay adaptive and rate-adaptive applications. However, existing cellular networks are mostly designed to support the conventional HTC traffic and a suitable characterization of mMTC traffic is crucial to support MTC devices. In this regard, the 3GPP has already defined the following two types of aggregated MTC traffic models: (i) uniform distribution over a duration T, and (ii) Beta distribution over T. Although such an aggregated traffic modeling is suitable to the scenarios with a huge number of devices and is simpler to realize, this approach can not capture the real traffic features at the source level and is less precise than the source traffic modeling. On the other hand, source traffic models although being more precise become complex for a large number of devices. Therefore, there arises a need to investigate suitable hybrid approaches such as coupled Markov modulated Poisson processes, which can combine the advantages of both the aggregated traffic modeling and source traffic modeling.

Furthermore, as compared to the existing wireless systems which are mainly designed to support the conventional HTC traffic having long data-packets, the transmission of short data-packets in mMTC/IoT systems creates new challenges since the size of metadata is no longer negligible and the classical law of large numbers does not become applicable. Therefore, it is a crucial research issue to design efficient techniques to tackle the challenges caused by short data-packet transmissions, and also to investigate suitable information theoretic principles to characterize the short data-packet transmission in IoT/mMTC systems. In this regard, several research works have already investigated different physical layer approaches to support small data-packet transmissions as highlighted in Table V. The main approaches include the optimization of pilot overhead, design of waveforms and air interface, modulation and coding schemes, NOMA scheme, minimization of core network signalling, joint encoding of grouped messages, autonomous transmission mode, receiver algorithms to enhance the reception quality and frequency diversity to enhance transmission reliability.

As highlighted in Section II-D, the 3GPP has defined two different models for the aggregated traffic of the mMTC systems. The first model based on uniform distribution is for the non-synchronized type of traffic whereas the second model based on Beta distribution suits more for the highly synchronized traffic. This aggregated traffic modeling is simpler but is less precise than the source traffic modeling which treats the traffic for each MTC device separately and is more complex. In this regard, investigating suitable low-complexity and precise traffic modeling is an important future research direction. Furthermore, the Transmission Control Protocol (TCP) employed at the transport layer of current LTE-A networks is not efficient for MTC traffic due to several issues associated with connection set up, congestion control, data buffering and real-time applications [73]. Therefore, it is crucial to develop an enhanced version of TCP over LTE/LTE-A to address various issues faced by the existing TCP for supporting the MTC traffic.

D. Random Access Schemes for Ultra-Dense IoT Networks

Regarding the RA schemes in ultra-dense cellular IoT/mMTC networks, investigating suitable contention-based schemes to concurrently support the massive number of MTC devices without affecting the performance of the existing cellular users is a crucial challenge since the available RBs are limited as compared to the massive number of access requests. In the existing LTE/LTE-A based cellular networks, the limitation in the number of available preambles may result in significantly high number of collisions among the access requests, and subsequently lead to the RACH congestion problem in the IoT access network. The widely-followed RA procedure in LTE/LTE-A based cellular networks follows a four stage message handshake procedure and this RA procedure may fail due to various reasons including the failure of preamble transmission, message 2 reception, message 3 transmission and message 4 reception. In the absence of access collisions, the absolute maximum capacity of an RACH channel in LTE-A networks is 10,800 preambles per seconds. However, because of the ALOHA type RA protocol and random backoffs, the practical performance becomes much lower than this maximum limit and the situation becomes further worse while supporting massive MTC devices.

In the above context, the four-stage message handshake procedure employed in the legacy LTE systems is not efficient to support MTC devices due to various reasons including the limited number of available preambles, need of additional downlink resources and huge signalling overhead in the short data-packet transmissions from MTC devices. Another limitation arises due to the hardware limitations of MTC devices in terms of their capacity to listen to a wideband PDCCH channel and in this regard, the 3GPP has proposed a modified RACH structure, i.e., EPDCCH with the narrow bandwidth of 1.4 MHz for low-cost MTC devices and each PRACH having dedicated NB EPDCCH. Furthermore, there exist two main categories of cellular IoT standards: (i) LTE-M, and (ii) NB-IoT, whose main differences were highlighted in Table IV.

As compared to the legacy LTE systems, the main changes incorporated in the LTE-M standard include the support of frequency hopping, and repetition procedures in the PRACH and PDCCH, MTC search spaces to reduce the number of decoding trials and three different DCI formats for uplink grant, downlink scheduling and paging in MTC devices. On the other hand, NB-IoT systems can operate in three different modes: (i) in-band operation, (ii) guard band operation and (iii) stand-alone operation. And, the main requirements of NB-IoT systems include low power consumption, low channel bandwidth, low cost for end devices, low deployment cost, extended coverage and support for massive connectivity.

Towards addressing RAN congestion problem, several schemes (listed in Table VI) have been investigated in the literature. The main existing RA based solutions include the back-off based scheme, ACB, EAB, cooperative ACB scheme, dynamic ACB, prioritized RA with dynamic ACB, dynamic resource allocation, slotted RA, RA resource separation, pull-based/paging-based scheme, group-based, code-expanded and tree-based RA schemes. Although there are recent advances in the areas of novel RA schemes towards supporting mMTC devices with the existing cellular networks, these schemes may not be sufficient to support the massive number of connections concurrently while satisfying their diverse QoS requirements. Some emerging techniques for future research include learning-based schemes, distributed queueing, and SDN and virtualization based techniques. Furthermore, grant-based random access and Compressive Sensing (CS)-based device activity detection for enabling the use of grant-free access scheme to minimize the delay of the existing contention-based random access schemes seem promising [213].

E. Distributed Resource Management in Ultra-Dense IoT Networks

In contrast to the connection-oriented design approach of the existing wireless networks while considering only the communication resources, future content-oriented networks are expected to utilize other resources such as computing and caching resources. In ultra-dense IoT networks, these communications, computing and caching resources are usually distributed across the different entities of the network including the devices, aggregator/eNodeBs and core-network/cloud-center. The coordination of these distributed resources to enhance the performance of ultra-dense IoT networks involving low-end MTC devices is an important research issue. Furthermore, in the emerging cloud-assisted IoT networks, it is advantageous to handle computationally intensive task at the cloud-center due to the availability of huge computing and storage capabilities while it becomes beneficial to handle delay-sensitive applications at the edge-side of the network. In this regard, it is an important future research direction to investigate suitable techniques for edge-cloud collaborative processing in various applications including dynamic spectrum sharing, event detection, context-aware resource allocation, live data analytics and security enhancement [21].

Moreover, various features of MTC device transmissions such as time-controlled, time tolerant, priority alarm message, infrequent transmission, group-based policing and addressing can be useful to utilize the distributed cache embedded in MTC devices. The caching capabilities distributed across heterogeneous IoT devices can enable the scheduling of sporadic transmissions from the MTC devices in order to significantly reduce the peak traffic in an IoT access network as demonstrated in [14]. This will subsequently reduce the demand for the radio resources at the peak time, thus significantly saving the radio resource cost for the telecommunication operators. In addition, distributed caching can be exploited to enable aggregate transmission for various other applications such as saving energy for low-power IoT devices and reducing signalling/protocol overhead for MTC device transmissions. Also, the physical-level cache embedded in low-end distributed MTC devices can be exploited to facilitate the implementation of a cross-layer based transmission scheduling in pushing data from the physical layer to the MAC layer at the suitable intervals.

In ultra-dense cellular IoT networks, one crucial challenge is to investigate the efficient coordination of distributed computing, communication and caching resources. Furthermore, the heterogeneity of MTC devices in terms of cache size, computing capability, battery power, latency and data rate requirements creates challenges in providing efficient QoS provisioning in ultra-dense IoT networks. One of the potential approaches to address some of the existing problems including power consumption and signalling congestion is to enable the group-based operation as highlighted in the following subsection.

F. Device Heterogeneity and Grouping-based Transmission Schemes

The heterogeneity of MTC devices in terms of different aspects such as computing capability, cache size, battery power, data rate and latency requirements becomes problematic for the efficient QoS provisioning in ultra-dense IoT networks. Furthermore, RAN congestion, high signalling overhead and power consumption are critical issues to be addressed in ultra-dense IoT networks. In this regard, grouping-based features of MTC transmissions such as group-based policing and group-based addressing [10] can be utilized to enable the group-based operation in the mMTC environment towards alleviating various problems such as signalling congestion and power consumption. Furthermore, by implementing a group-based access authentication technique, severe signalling congestion caused by the conventional independent access authentication scheme in the existing cellular systems can be avoided [211] and the security of emerging MTC applications can be significantly enhanced.

By grouping MTC devices on the basis of either service requirements or the physical locations of MTC devices, group-based data gathering, aggregation and reporting can be utilized in ultra-dense IoT networks including IEEE 802.11ah based systems [212]. In such group-based schemes, a group header/cluster head collects the access requests, uplink data packets, and device status information from the resource-constrained MTC devices belonging to that group, and then forwards the aggregated traffic to an eNodeB/aggregator [6].

Also, the downlink data packets and signalling messages can be relayed to the MTC devices by the group header, thus significantly reducing the radio resources required for direct communications between the devices and the eNodeB. Moreover, as the jittering constraints become challenging while transmitting small data packets from a large number of devices, this issue can be addressed by employing grouping-based resource scheduling which allocate radio resources to the specific groups having similar QoS characteristics [82]. However, the existing device grouping mechanisms are formed mostly in a traditional way by selecting the devices in a random fashion or in an uniform manner, and it is an important research direction to investigate efficient grouping mechanisms by exploiting the QoS characteristics and locations of heterogeneous MTC devices.

G. ML Applications in Wireless IoT/mMTC Systems

Emerging ML-assisted solutions can be employed to mitigate various problems in wireless systems and also to address the non-conventional challenges of IoT/mMTC systems. Mainly, the emerging ultra-dense IoT networks will need to observe the environmental variations, learn uncertainties, plan response actions and configure the network parameters effectively, and in this regard, the ML-assisted techniques can be significantly useful in acquiring the environmental parameters, in automating management and operational tasks, in classifying the involved use cases and in predicting future results/challenges. As listed in Table VIII, ML techniques find significance to address various issues in ultra-dense cellular systems including RACH congestion minimization, adaptive resource allocation, dynamic spectrum sharing, edge-side processing, traffic offloading, RAT selection, network traffic control, adaptation of transmission parameters, data analytics and provisioning of personalized services.

Existing ML techniques can be broadly categorized into supervised, unsupervised and reinforcement learning as depicted in Fig. 6. In Table IX, we provided the pros and cons of these learning techniques along with their applicability to IoT/mMTC environments. Supervised learning benefits from the full centralized control but requires the labelled data-sets and also needs the prior knowledge of a training data-set as well as the expected output. On the other hand, unsupervised learning techniques attempt to find hidden patterns in the data without the need of any labelled data, training data-set and expected output but the learning process becomes more complex since there is no clear guidance and the learning objective is of more subjective nature. The unsupervised learning is preferred over supervised learning in large and complex models having deep hierarchies where there exists a large gap between the input and output observations. As compared to the classical ML, DL techniques are highly flexible/configurable and can achieve much higher learning accuracy but involve a number of hyper-parameters and demand for highly capable GPU platforms.

While employing ML techniques in an mMTC environment, an important aspect to be considered is how to make ML techniques more practically realizable for the resource-constrained MTC devices. To achieve this objective, the following issues need to be considered. First, the convergence rate/learning time of the employed learning algorithm should be as small as possible and there may arise the issue of a local minimum. Since the required learning time may reduce the time for data transmission purpose, their trade-off should be designed properly. Second, in mMTC environments, the distributed implementation of the algorithm needs to be considered across multiple learning devices. Third, different parameters associated with the learning algorithms such as learning rate α, discount rate γ and exploration-exploitation tradeoff parameter ϵ should be adapted dynamically to enhance the performance of Q-learning algorithm in dynamic environments. Furthermore, there may arise the fairness issue while applying learning algorithms in a multi-agent environment since different devices may reach to convergence at different time intervals, thus creating different learning times for different devices. Moreover, the heterogeneity of MTC devices need to be considered in terms of different aspects such as learning capability, cache size, data rate and delay tolerance limit.

With regard to the application in IoT/mMTC systems, the RL technique seem to be more promising due to its simplicity of operation as well as the feasibility of distributed implementation. Furthermore, an RL technique utilizes the action-reward feedback by exploiting the interactions between learning agents and the underlying environment in model-free settings. One of the important applications where ML can be advantageous in IoT/mMTC systems is the security enhancement and privacy preservation. The resource-constrained IoT/mMTC devices may not be able to implement legacy security solutions since they demand for huge computation and communication capabilities. To this end, ML/DL-assisted security techniques including access control, secure offloading, IoT authentication and malware detection schemes can be significantly useful in addressing various IoT security threats including jamming, spoofing, DoS attacks, software attacks and privacy leakage. Also, DL schemes are suitable for privacy preservation since the intermediate data and source data in the DL architectures have different semantics. However, the implementation of DL techniques in mMTC systems is challenging since they demand for a significant amount of memory, energy and battery resources. In this regard, some potential enabling techniques include network compression, approximate computing and accelerators.

One of the promising RL techniques suitable for IoT/mMTC devices is Q-learning since its operation is purely algebraic and can be implemented in a distributed manner. In this regard, we provided a framework for the application of Q-learning for RACH congestion in Section VI-D as an illustration of the use-case example. With this approach, each MTC device attempts to find its unique time slot for its transmission towards minimizing the RACH congestion. In a distributed set-up, all MTC nodes start with zero or random Q values and learn gradually by utilizing the reward obtained based on the success or failure of their transmissions till they all come across unique RA slots for their transmissions. One of the design strategies for Q-learning is to balance the exploration-exploitation tradeoff and the widely-used three strategies include ϵ-greedy strategy, soft-max strategy and optimization in the face of uncertainty. Furthermore, the performance of Q-learning can be enhanced with several novel concepts including collaborative Q-learning, situation-aware adaptive Q-learning, fuzzy-logic based adaptive Q-learning and model-based Q-learning.

In the above context, future works should focus on addressing the aforementioned issues to employ the ML-assisted solutions towards enabling the incorporation of MTC devices in the upcoming 5G and beyond cellular networks. Various emerging techniques described in Section VI-F such as collaborative learning, situation-aware adaptive learning, fuzzy-logic based Q-learning and model-based learning could be exploited to enhance the performance of the ML techniques in ultra-dense IoT networks involving MTC devices. In summary, investigating low-complexity ML solutions which are feasible to implement in the resource-constrained MTC devices is one of the future research challenges. Also, tuning of different parameters of ML/DL algorithms and their convergence time to reach a steady solution are other important aspects to be considered. In the emerging DL algorithms, another important issue is how to balance the tradeoff between the depth of the network and the computational complexity.

H. Deep Learning for Emerging IoT Applications

In real-world IoT environments, one of the major issues is how to reliably extract the meaningful information out of the massive amount of unstructured/semi-structured IoT data obtained from a complex and noisy environment. Conventional ML learning techniques may fail in such complex and dynamic environments and deep learning can be considered as a promising solution [60]. Due to multi-layer structure of DL, it is considered as an effective approach for the edge-computing environment in order to accurately extract necessary information from the raw IoT sensor data [214]. Since it is possible to offload the parts of learning layers in the edge-side of the network and transfer the reduced intermediate data to the remote cloud-center, the DL model seems suitable for the emerging edge computing paradigms in IoT systems. Another interesting advantage of DL in edge-computing environment is that it can provide privacy preservation in transferring the intermediate data. In contrast to the traditional big data systems such as Spark or MapReduce in which intermediate data contains the user privacy information, the intermediate data generated in DL networks usually have different semantics than that in the original source data [60].

In practical IoT systems, the data is usually dynamic and unlabelled, and the conventional statistically trained models are not efficient to handle the large unlabelled and dynamic data-set. Furthermore, it is highly impractical to manually label all the IoT raw data. Due to this, the conventional supervised training-based learning techniques are not suitable for large-scale IoT/mMTC environments [215]. Moreover, the conventional cloud-based architecture requires the transfer of a huge amount of IoT data from the edge-devices to the cloud-center. To address these issues, the application of DL with collaborative cloud-edge/fog processing seems to be a promising future research direction [21, 215].

However, the application of DL in IoT systems faces the crucial challenge of meeting the computational requirements. The main associated issues include the high-speed training of large-scale IoT networks with the massive data-sets and embedding DL capability in low-power IoT devices [216]. This computational challenge caused due to the expected growth in the size of the data-sets and the algorithmic complexity of ML algorithms is demanding the need of improving the computational efficiency of existing computing platforms. Also, it is not feasible to offload all the data to the cloud-center for processing due to constraints on the bandwidth, privacy and battery life, and the computational efficiency at the device-side needs to be accelerated for DL applications. In this regard, some of the potential future solutions to improve computational efficiency of ML algorithms include deep-learning accelerators, approximate computing and post-CMOS device technologies [216].

Many IoT products have already used the ML techniques to acquire and analyze the environmental data. For example, Google's Nest Learning Thermostat uses ML algorithms to understand the patterns of its users' temperature schedules and preferences by utilizing the temperature data recorded in a structure way. However, unstructured multimedia data such as visual images and audio signals are difficult to learn by using the conventional ML techniques. In this regard, some emerging IoT devices are already using sophisticated DL techniques to capture and understand the complex environments [46]. As an example, the face-recognition security system from the Microsoft's Windows IoT team uses deep-learning technology to perform tasks such as unlocking a door by recognizing its users' faces.

Due to demanding real-time requirements of IoT applications in terms of latency and high cost of radio resources required in delivering information to the cloud-center, it is advantageous to implement DL techniques at the device-side. However, due to the limited computing power and low memory size of IoT devices, it is challenging to implement DL at the device-side. Therefore, most of the time, existing DL applications require third-party libraries and it may be difficult to migrate them to the IoT devices [46]. In this regard, it is an important future research direction to investigate suitable paradigms such as convolution neural networks based inference engine [46] to facilitate the implementation of DL at the IoT devices.

VIII. Conclusions

Future cellular IoT networks are expected to support the massive number of resource-constrained MTC devices while satisfying their diverse QoS requirements and will need to deal with several challenges for enhancing the access latency, scalability, connection reliability, energy efficiency and network throughput. To this end, this paper has discussed various challenges of mMTC systems such as QoS provisioning, mMTC traffic characterization, transmission scheduling with QoS support, small data packet transmission and RAN congestion, and has provided a detailed review on the existing studies attempting to address these issues. By considering machine learning as an important enabler to address some of these issues in ultra-dense cellular IoT networks, the paper has identified the potential advantages, research challenges and the application scenarios of ML-assisted solutions. Among potential ML techniques, the application of Q-learning in minimizing the RAN congestion has been presented along with some performance enhancement techniques. Finally, a summary of lessons learned, some important research issues and interesting directions for future research have been provided.

REFERENCES

[1] A. Al-Fuqaha, *et al.*, "Internet of things: A survey on enabling technologies, protocols, and applications," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 4, pp. 2347–2376, Fourthquarter 2015.

[2] H. Wang and A. O. Fapojuwo, "A survey of enabling technologies of low power and long range machine-to-machine communications," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 4, pp. 2621–2639, Fourthquarter 2017.

[3] ITU-R, "IMT vision - framework and overall objectives of the future development of IMT for 2020 and beyond," ITU-R Rec. M.2083-0, Sept. 2015.

[4] B. A. Jayawickrama, Y. He, E. Dutkiewicz, and M. D. Mueck, "Scalable spectrum access system for massive machine type communication," *IEEE Network*, vol. PP, no. 99, pp. 1–7, 2018.

[5] 3GPP, "Study on scenarios and requirements for next generation access technologies (Release 14)," Technical Report 38.913, 2017.

[6] F. Ghavimi and H. H. Chen, "M2M communications in 3GPP LTE/LTE-A networks: Architectures, service requirements, challenges, and applications," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 2, pp. 525–549, Secondquarter 2015.

[7] M. Koseoglu, "Performance analysis of small data transmission schemes for cellular M2M communications," in *Proc. 16th Annual Mediterranean Ad Hoc Networking Workshop*, June 2017, pp. 1–6.

[8] Z. Dawy, *et al.*, "Toward massive machine type cellular communications," *IEEE Wireless Communications*, vol. 24, no. 1, pp. 120–128, Feb. 2017.

[9] K. Wang, *et al.*, "Green Industrial Internet of Things Architecture: An Energy-Efficient Perspective," in *IEEE Commun. Mag.*, vol. 54, no. 12, pp. 48-54, Dec. 2016.

[10] 3GPP, "Service requirements for machine-type communications (MTC), TS 22.368, V11.5.0, June 2012.

[11] L. M. Bello, P. Mitchell, and D. Grace, "Application of Q-learning for RACH access to support M2M traffic over a cellular network," in *European Wireless 2014; 20th European Wireless Conference*, May 2014, pp. 1–6.

[12] G. Durisi, T. Koch, and P. Popovski, "Toward massive, ultrareliable, and low-latency wireless communication with short packets," *Proc. of the IEEE*, vol. 104, no. 9, pp. 1711–1726, Sept. 2016.

[13] S. Oh and J. W. Jang, "A scheme to smooth aggregated traffic from sensors with periodic reports," *MDPI Sensors*, vol. 17, no. 503, pp. 1–18, March 2017.

[14] S. K. Sharma and X. Wang, "Distributed caching enabled peak traffic reduction in ultra-dense IoT networks," *IEEE Commun. Letters*, vol. 22, no. 6, pp. 1252-1255, June 2018.

[15] C. Bockelmann, *et al.*, "Massive machine-type communications in 5G: physical and MAC-layer solutions," *IEEE Commun. Mag.*, vol. 54, no. 9, pp. 59–65, Sept. 2016.

[16] A. Imran, A. Zoha, and A. Abu-Dayya, "Challenges in 5G: how to empower SoN with big data for enabling 5G," *IEEE Network*, vol. 28, no. 6, pp. 27–33, Nov 2014.

[17] R. Li, *et al.*, "Intelligent 5G: When cellular networks meet artificial intelligence," *IEEE Wireless Commun.*, vol. 24, no. 5, pp. 175–183, Oct. 2017.

[18] R. C. Daniels and R. W. Heath, "An online learning framework for link adaptation in wireless networks," in *Proc. Info. Theory and Applications Workshop*, Feb. 2009, pp. 138–140.

[19] S. Yun and C. Caramanis, "Reinforcement learning for link adaptation in MIMO-OFDM wireless systems," in *Proc. IEEE GLOBECOM* , Dec. 2010, pp. 1–5.

[20] J. Wu, *et al.*, "Context-aware networking and communications: Part 1 [Guest Editorial]," in *IEEE Commun. Mag.*, vol. 52, no. 6, pp. 14-15, June 2014.

[21] S. K. Sharma and X. Wang, "Live data analytics with collaborative edge and cloud processing in wireless IoT networks," *IEEE Access*, vol. 5, pp. 4621–4635, March 2017.

[22] X. Wang, X. Li, and V. C. M. Leung, "Artificial intelligence-based techniques for emerging heterogeneous network: State of the arts, opportunities, and challenges," *IEEE Access*, vol. 3, pp. 1379–1391, 2015.

[23] J. Moon and Y. Lim, "Access control of MTC devices using reinforcement learning approach," in *Proc. Int. Conf. on Information Networking*, Jan. 2017, pp. 641–643.

[24] M. Hasan, E. Hossain, and D. Niyato, "Random access for machine-to-machine communication in LTE-advanced networks: issues and approaches," *IEEE Commun. Mag.*, vol. 51, no. 6, pp. 86–93, June 2013.

[25] A. H. Mohammed, A. S. Khwaja, A. Anpalagan, and I. Woungang, "Base station selection in M2M communication using Q-learning algorithm in LTE-A networks," in *Proc. IEEE 29th Int. Conf. Advanced Info. Networking and Applications*, March 2015, pp. 17–22.

[26] T. Park and W. Saad, "Resource allocation and coordination for critical messages using finite memory learning," in *Prof. IEEE Globecom Workshops (GC Wkshps)*, Dec. 2016, pp. 1–6.

[27] J. Zhou, *et al.*, "Security and Privacy for Cloud-Based IoT: Challenges," in *IEEE Commun. Mag.*, vol. 55, no. 1, pp. 26-33, Jan. 2017.

[28] L. Xiao, *et al.*, "IoT Security Techniques Based on Machine Learning: How Do IoT Devices Use AI to Enhance Security?," in *IEEE Signal Process. Mag.*, vol. 35, no. 5, pp. 41-49, Sept. 2018.

[29] F. A. Narudin, *et al.*, "Evaluation of machine learning classifiers for mobile malware detection," in *Soft Comput.*, vol. 20, no. 1, pp. 343–357, Jan. 2016.

[30] M. Ozay, *et al.*, "Machine Learning Methods for Attack Detection in the Smart Grid," in *IEEE Transactions on Neural Networks and Learning Systems*, vol. 27, no. 8, pp. 1773-1786, Aug. 2016.

[31] A. L. Buczak and E. Guven, "A Survey of Data Mining and Machine Learning Methods for Cyber Security Intrusion Detection," in *IEEE Communications Surveys & Tuts.*, vol. 18, no. 2, pp. 1153-1176, Secondquarter 2016.

[32] L. Xiao, *et al.*, "A Mobile Offloading Game Against Smart Attacks," in *IEEE Access*, vol. 4, pp. 2281-2291, 2016.

[33] A. Diro and N. Chilamkurti, "Leveraging LSTM Networks for Attack Detection in Fog-to-Things Communications," in *IEEE Commun. Mag.*, vol. 56, no. 9, pp. 124-130, Sept. 2018.

[34] K. Hamedani, *et al.*, "Reservoir Computing Meets Smart Grids: Attack Detection Using Delayed Feedback Networks," in *IEEE Trans. Industrial Informatics*, vol. 14, no. 2, pp. 734-743, Feb. 2018.

[35] M. Mohammadi, A. Al-Fuqaha, S. Sorour and M. Guizani, "Deep Learning for IoT Big Data and Streaming Analytics: A Survey," in *IEEE Commun. Surveys & Tut.*, vol. 20, no. 4, pp. 2923-2960, Fourthquarter 2018.

[36] Psychoula, *et al.*, "A Deep Learning Approach for Privacy Preservation in Assisted Living," in *Proc. IEEE Int. Conf. Pervasive Computing and Commun. Workshops (PerCom Workshops)*, Athens, 2018, pp. 710-715.

[37] N. D. Lane, *et al.*, "An early resource characterization of deep learning on wearables, smartphones and Internet-of-Things devices," in *Proc. Int. Workshop Internet Things Towards Appl.*, Seoul, South Korea: ACM, 2015, pp. 7–12.

[38] B. Moons, *et al.*, "Energy-efficient ConvNets through approximate computing," in *Proc. IEEE Winter Conf. Applications of Computer Vision (WACV)*, Lake Placid, NY, 2016, pp. 1-8.

[39] T. Chen, *et al.*, "DianNao: A small-footprint high-throughput accelerator for ubiquitous machine-learning," *ACM Sigplan Notices*, vol. 49, no. 4, pp. 269-284, 2014.

[40] Y.-H. Chen, *et al.*, "Eyeriss: An energy-efficient reconfigurable accelerator for deep convolutional neural networks," *IEEE J. Solid-State Circuits*, vol. 52, no. 1, pp. 127-138, Jan. 2017.

[41] W. Zhang, *et al.*, "LSTM-Based Analysis of Industrial IoT Equipment," in *IEEE Access*, vol. 6, pp. 23551-23560, 2018.

[42] B. Qolomany, A. Al-Fuqaha, D. Benhaddou and A. Gupta, "Role of Deep LSTM Neural Networks and Wi-Fi Networks in Support of Occupancy Prediction in Smart Buildings," in *Proc. IEEE Int. Conf. HPCC/SmartCity/DSS*, Bangkok, 2017, pp. 50-57.

[43] K. Portelli and C. Anagnostopoulos, "Leveraging edge computing through collaborative machine learning," in *Proc. 5th Int. Conf. on Future Internet of Things and Cloud Workshops (FiCloudW)*, Aug. 2017, pp. 164–169.

[44] X. Chen, *et al.*, "Energy-Efficiency Oriented Traffic Offloading in Wireless Networks: A Brief Survey and a Learning Approach for

Heterogeneous Cellular Networks," in *IEEE J. Sel. Areas Commun.*, vol. 33, no. 4, pp. 627-640, April 2015.

[45] 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.

[46] J. Tang, D. Sun, S. Liu, and J. L. Gaudiot, "Enabling deep learning on IoT devices," *Computer*, vol. 50, no. 10, pp. 92–96, 2017.

[47] M. Wang, *et al.*, "Cellular machine-type communications: physical challenges and solutions," *IEEE Wireless Commun.*, vol. 23, no. 2, pp. 126–135, April 2016.

[48] A. Rico-Alvarino, *et al.*, "An overview of 3GPP enhancements on machine to machine communications," *IEEE Commun. Magazine*, vol. 54, no. 6, pp. 14–21, June 2016.

[49] M. Elsaadany, A. Ali, and W. Hamouda, "Cellular LTE-A technologies for the future internet-of-things: Physical layer features and challenges," *IEEE Commun. Surveys Tuts.*, vol. PP, no. 99, pp. 1–1, 2017.

[50] A. Hoglund, *et al.*, "Overview of 3GPP release 14 enhanced NB-IoT," *IEEE Network*, vol. 31, no. 6, pp. 16–22, Nov. 2017.

[51] S. K. Sharma, *et al.*, "Physical layer aspects of wireless IoT," in *Proc. 2016 International Symp. on Wireless Commun. Systems (ISWCS)*, Poznan, Sept. 2016, pp. 304-308.

[52] A. Laya, L. Alonso, and J. Alonso-Zarate, "Is the random access channel of LTE and LTE-A suitable for M2M communications? a survey of alternatives," *IEEE Commun. Surveys Tuts.*, vol. 16, no. 1, pp. 4–16, Firstquarter 2014.

[53] W. Yang, *et al.*, "Narrowband wireless access for low-power massive internet of things: A bandwidth perspective," *IEEE Wireless Commun.*, vol. 24, no. 3, pp. 138–145, 2017.

[54] M. S. Ali, E. Hossain, and D. I. Kim, "LTE/LTE-A random access for massive machine-type communications in smart cities," *IEEE Commun. Mag.*, vol. 55, no. 1, pp. 76–83, Jan. 2017.

[55] E. Soltanmohammadi, K. Ghavami, and M. Naraghi-Pour, "A survey of traffic issues in machine-to-machine communications over LTE," *IEEE Internet of Things J.*, vol. 3, no. 6, pp. 865–884, Dec 2016.

[56] A. G. Gotsis, A. S. Lioumpas, and A. Alexiou, "M2M scheduling over LTE: Challenges and new perspectives," *IEEE Veh. Technol. Mag.*, vol. 7, no. 3, pp. 34–39, Sept. 2012.

[57] M. A. Mehaseb, Y. Gadallah, A. Elhamy, and H. Elhennawy, "Classification of LTE uplink scheduling techniques: An M2M perspective," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 2, pp. 1310–1335, Secondquarter 2016.

[58] M. Marjani, *et al.*, "Big IoT data analytics: Architecture, opportunities, and open research challenges," *IEEE Access*, vol. 5, pp. 5247–5261, March 2017.

[59] R. Atat, *et al.*, "Big Data Meet Cyber-Physical Systems: A Panoramic Survey," in *IEEE Access*, vol. 6, pp. 73603-73636, Nov. 2018.

[60] S. Verma, *et al.*, "A survey on network methodologies for real-time analytics of massive IoT data and open research issues," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 3, pp. 1457–1477, thirdquarter 2017.

[61] H. Elhammouti, *et al.*, "Self-organized connected objects: Rethinking QoS provisioning for IoT services," *IEEE Commun. Mag.*, vol. 55, no. 9, pp. 41–47, 2017.

[62] 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.

[63] O. B. Sezer, E. Dogdu, and A. M. Ozbayoglu, "Context-aware computing, learning, and big data in internet of things: A survey," *IEEE Internet of Things J.*, vol. 5, no. 1, pp. 1–27, Feb. 2018.

[64] L. Busoniu, *et al.*, "A comprehensive survey of multiagent reinforcement learning," *IEEE Trans. Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 38, no. 2, pp. 156–172, March 2008.

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

[66] R. Baldemair, *et al.*, "Ultra-dense networks in millimeter-wave frequencies," *IEEE Commun. Mag.*, vol. 53, no. 1, pp. 202–208, Jan. 2015.

[67] J. An, *et al.*, "Achieving Sustainable Ultra-Dense Heterogeneous Networks for 5G," in *IEEE Commun. Mag.*, vol. 55, no. 12, pp. 84-90, Dec. 2017.

[68] S. Chen, *et al.*, "Machine-to-machine communications in ultra-dense networks: A survey," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 3, pp. 1478–1503, thirdquarter 2017.

[69] M. Kamel, W. Hamouda, and A. Youssef, "Ultra-dense networks: A survey," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 4, pp. 2522–2545, Fourthquarter 2016.

[70] K. Smiljkovic, V. Atanasovski, and L. Gavrilovska, "Machine-to-machine traffic characterization: Models and case study on integration in LTE," in *Proc. 4th Int. Conf. Wireless Commun., Veh. Technol., Info. Th. and Aerospace Electronic Systems (VITAE)*, May 2014, pp. 1–5.

[71] X. Li, J. B. Rao, and H. Zhang, "Engineering machine-to-machine traffic in 5G," *IEEE Internet of Things J.*, vol. 3, no. 4, pp. 609–618, Aug 2016.

[72] N. Afrin, J. Brown, and J. Y. Khan, "An adaptive buffer based semi-persistent scheduling scheme for machine-to-machine communications over LTE," in *Proc. Eighth Int. Conf. Next Generation Mobile Apps, Services and Technol.*, Sept 2014, pp. 260–265.

[73] J. Kim, J. Lee, J. Kim, and J. Yun, "M2M service platforms: Survey, issues, and enabling technologies," *IEEE Commun. Surveys Tuts.*, vol. 16, no. 1, pp. 61–76, Firstquarter 2014.

[74] J. Fabini and T. Zseby, "M2M communication delay challenges: Application and measurement perspectives," in *Proc. IEEE Int. Instrumentation and Measurement Technol. Conf.*, May 2015, pp. 1859–1864.

[75] S. Y. Lien and K. C. Chen, "Massive access management for QoS guarantees in 3GPP machine-to-machine communications," *IEEE Commun. Letters*, vol. 15, no. 3, pp. 311–313, March 2011.

[76] Y. Saleem, N. Crespi, M. H. Rehmani and R. Copeland, "Internet of Things-aided Smart Grid: Technologies, Architectures, Applications, Prototypes, and Future Research Directions", CoRR, abs/1704.08977,2017, online: http://arxiv.org/abs/1704.08977.

[77] K. Zheng, *et al.*, "Radio resource allocation in LTE-advanced cellular networks with M2M communications," *IEEE Commun. Mag.*, vol. 50, no. 7, pp. 184–192, July 2012.

[78] M. Y. Cheng, *et al.*, "Overload control for machine-type-communications in LTE-advanced system," *IEEE Commun. Mag.*, vol. 50, no. 6, pp. 38–45, June 2012.

[79] R. Ratasuk, N. Mangalvedhe, D. Bhatoolaul, and A. Ghosh, "LTE-M evolution towards 5G massive MTC," in *Proc. IEEE Globecom Workshops (GC Wkshps)*, Dec. 2017, pp. 1–6.

[80] P. K. Wali and D. Das, "A novel access scheme for IoT communications in LTE-advanced network," in *Proc. IEEE Int. Conf. Advanced Networks and Telecommun. Systems (ANTS)*, Dec 2014, pp. 1–6.

[81] A. M. AlAdwani, A. Gawanmeh, and S. Nicolas, "A demand side management traffic shaping and scheduling algorithm," in *Proc. Sixth Asia Modelling Symp.*, May 2012, pp. 205–210.

[82] S. Y. Lien, K. C. Chen, and Y. Lin, "Toward ubiquitous massive accesses in 3GPP machine-to-machine communications," *IEEE Commun. Mag.*, vol. 49, no. 4, pp. 66–74, April 2011.

[83] Z. Li, *et al.*, "Enabling heterogeneous MMTC by energy-efficient and connectivity-aware clustering and routing," in *Proc. IEEE Globecom Workshops (GC Wkshps)*, Dec. 2017, pp. 1–6.

[84] 3GPP, "Cellular system support for ultralow complexity and low throughput Internet of Things (CIoT),", TR 45.820, V13.1.0, Nov. 2015.

[85] P. Andres-Maldonado, *et al.*, "Narrowband IoT data transmission procedures for massive machine-type communications," *IEEE Network*, vol. 31, no. 6, pp. 8–15, Nov. 2017.

[86] J. Huang, *et al.*,, "Optimizing M2M communications and quality of services in the IoT for sustainable smart cities," *IEEE Trans. on Sustainable Computing*, vol. PP, no. 99, pp. 1–1, 2017.

[87] M. Centenaro, *et al.*,, "Comparison of collision-free and contention-based radio access protocols for the internet of things," *IEEE Trans. Commun.*, vol. 65, no. 9, pp. 3832–3846, Sept. 2017.

[88] U. Devi, *et al.*, "SERA: A scheduling framework for M2M transmission in cellular networks," online, web: http://researcher.ibm.com/researcher/files/in-umamadev/M2M-long.pdf.

[89] Y. Medjahdi, *et al.*, "On the road to 5G: Comparative study of physical layer in MTC context," *IEEE Access*, vol. 5, pp. 26556–26581, 2017.

[90] 3GPP, "Study on RAN improvements for machine-type communications," Tech. Rep., TR 37.868, 2012.

[91] J. Choi, "On the adaptive determination of the number of preambles in RACH for MTC," *IEEE Commun. Letters*, vol. 20, no. 7, pp. 1385–1388, July 2016.

[92] S. J. Sung-Hyung Lee and J. Kim, "Dynamic resource allocation of random access for MTC devices," *ETRI Journal*, vol. 39, no. 4, pp. 546–557, Aug. 2017.

[93] S. K. Sharma, *et al.*, "Dynamic spectrum sharing in 5G wireless networks with full-duplex technology: Recent advances and research challenges," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 1, pp. 674–707, Firstquarter 2018.

[94] M. Li, *et al.*, "Random access and virtual resource allocation in software-defined cellular networks with machine-to-machine communications," *IEEE Trans. Veh. Technol.*, vol. 66, no. 7, pp. 6399–6414, July 2017.

[95] A. Samir, *et al.*, "Partial contention free random access protocol for M2M communications in LTE networks," *J. of Wireless Networking and Commun.*, vol. 6, no. 3, pp. 66–72, 2016.

[96] G. Wunder, C. Stefanovic, P. Popovski, and L. Thiele, "Compressive coded random access for massive MTC traffic in 5G systems," in *Proc. 49th Asilomar Conf. on Signals, Systems and Computers*, Nov. 2015, pp. 13–17.

[97] I. M. Delgado-Luque, *et al.*, "Evaluation of latency-aware scheduling techniques for M2M traffic over LTE," in *Proc. 20th European Signal Processing Conference (EUSIPCO)*, Bucharest, 2012, pp. 989-993.

[98] S. Ali, N. Rajatheva and W. Saad, "Fast Uplink Grant for Machine Type Communications: Challenges and Opportunities", CoRR, abs/1801.04953, 2018, online: http://arxiv.org/abs/1801.04953.

[99] S. Bhandari, S. K. Sharma and X. Wang, "Latency Minimization in Wireless IoT Using Prioritized Channel Access and Data Aggregation," in *Proc. IEEE Global Commun. Conf.*, Singapore, Dec. 2017, pp. 1-6.

[100] S. K. Sharma and X. Wang, "Cooperative sensing delay minimization in cloud-assisted DSA networks" in *Proc. IEEE 28th Annual Int. Symp. on Personal, Indoor, and Mobile Radio Commun. (PIMRC)*, Montreal, QC, Oct. 2017, pp. 1-6.

[101] T. P. C. de Andrade, C. A. Astudillo and N. L. S. da Fonseca, "Impact of M2M traffic on human-type communication users on the LTE uplink channel," in *Proc. 7th IEEE Latin-American Conf. Commun. (LATINCOM)*, Arequipa, 2015, pp. 1-6.

[102] F. Han, S. Zhao, L. Zhang and J. Wu, "Survey of Strategies for Switching Off Base Stations in Heterogeneous Networks for Greener 5G Systems," in *IEEE Access*, vol. 4, pp. 4959-4973, Aug. 2016.

[103] F. K. Shaikh, S. Zeadally and E. Exposito, "Enabling Technologies for Green Internet of Things," in *IEEE Systems Journal*, vol. 11, no. 2, pp. 983-994, June 2017.

[104] K. Wang, *et al.*, "A Survey on Energy Internet: Architecture, Approach, and Emerging Technologies," in *IEEE Systems J.*, vol. 12, no. 3, pp. 2403-2416, Sept. 2018.

[105] M. Laner, P. Svoboda, N. Nikaein, and M. Rupp, "Traffic models for machine type communications," in *Proc. The Tenth Int. Symp. Wireless Commun. Systems*, Aug. 2013, pp. 1–5.

[106] N. Nikaein, *et al.*, "Simple traffic modeling framework for machine type communication," in *Proc. The Tenth Int. Symp. Wireless Commun. Systems*, Aug 2013, pp. 1–5.

[107] A. M. AlAdwani, A. Gawanmeh, and S. Nicolas, "Traffic shaping and delay optimization in demand side management," in *Proc. 15th Int. Conf. on Computer Modelling and Simulation*, April 2013, pp. 722–727.

[108] A. Kumar, A. Abdelhadi, and C. Clancy, "A delay efficient multiclass packet scheduler for heterogeneous M2M uplink," in *Proc. IEEE MILCOM Conf.*, Nov 2016, pp. 313–318.

[109] ——, "An online delay efficient packet scheduler for M2M traffic in industrial automation," in *Proc. 2016 Annual IEEE Systems Conf.*, April 2016, pp. 1–6.

[110] Y. Zhao, B. Zhang, C. Li, and C. Chen, "On/off traffic shaping in the internet: Motivation, challenges, and solutions," *IEEE Network*, vol. 31, no. 2, pp. 48–57, March 2017.

[111] H. Chao, Y. Chen, J. Wu and H. Zhang, "Distribution Reshaping for Massive Access Control in Cellular Networks," in *Proc. IEEE 84th Veh. Technol. Conf. (VTC-Fall)*, Montreal, QC, 2016, pp. 1-5.

[112] J. S. Kim, S. Lee, and M. Y. Chung, "Efficient random-access scheme for massive connectivity in 3GPP low-cost machine-type communications," *IEEE Trans. Veh. Technol.*, vol. 66, no. 7, pp. 6280–6290, July 2017.

[113] S. Duan, *et al.*, "D-ACB: Adaptive congestion control algorithm for bursty M2M traffic in LTE networks," *IEEE Trans. Veh. Technol.*, vol. 65, no. 12, pp. 9847–9861, Dec. 2016.

[114] G. Y. Lin, S. R. Chang, and H. Y. Wei, "Estimation and adaptation for bursty LTE random access," *IEEE Trans. Veh. Technol.*, vol. 65, no. 4, pp. 2560–2577, April 2016.

[115] P. Osti, *et al.*, "Analysis of PDCCH performance for M2M traffic in LTE," *IEEE Trans. Veh. Technol.*, vol. 63, no. 9, pp. 4357–4371, Nov. 2014.

[116] A. Biral, *et al.*, "The challenges of M2M massive access in wireless cellular networks," *Digital Commun. and Networks*, vol. 1, no. 1, pp. 1–19, 2015.

[117] S. Y. Lien, *et al.*, "Cooperative access class barring for machine-to-machine communications," *IEEE Trans. Wireless Commun.*, vol. 11, no. 1, pp. 27–32, Jan. 2012.

[118] S. Duan, V. Shah-Mansouri, and V. W. S. Wong, "Dynamic access class barring for M2M communications in LTE networks," in *Proc. IEEE Global Commun. Conf.*, Dec. 2013, pp. 4747–4752.

[119] M. Koseoglu, "Lower bounds on the LTE-A average random access delay under massive M2M arrivals," *IEEE Trans. Commun.*, vol. 64, no. 5, pp. 2104–2115, May 2016.

[120] DIGI, "What are the difference between LTE-M and NB-IoT Cellular Protocols," available online: https://www.digi.com/, last accessed 15th March 2018.

[121] I. Lusky, "CAT-M1 vs NB-IoT-examining the real differences," available online: https://www.iot-now.com/, last accessed 16th March 2018.

[122] S. Andreev, *et al.*, "Efficient small data access for machine-type communications in LTE," in *Proc. IEEE Int. Conf. Commun. (ICC)*, June 2013, pp. 3569–3574.

[123] J. Hu, *et al.*, "Enhanced LTE physical downlink control channel design for machine-type communications," in *Proc. 7th Int. Conf. New Technologies, Mobility and Security (NTMS)*, July 2015, pp. 1–5.

[124] 3GPP, "General packet radio service (GPRS) enhancements for evolved universal terrestrial radio access network (E-UTRAN) access," TS 23.401 V14.3.0, March 2017.

[125] C. Yu, *et al.*, "Uplink scheduling and link adaptation for narrowband internet of things systems," *IEEE Access*, vol. 5, pp. 1724–1734, 2017.

[126] R. Boisguene, S. C. Tseng, C. W. Huang, and P. Lin, "A survey on NB-IoT downlink scheduling: Issues and potential solutions," in *Proc. 13th Int. Wireless Commun. and Mobile Computing Conf. (IWCMC)*, June 2017, pp. 547–551.

[127] Y. P. E. Wang, *et al.*, "A primer on 3GPP narrowband internet of things," *IEEE Commun. Mag.*, vol. 55, no. 3, pp. 117–123, March 2017.

[128] D. Wu and R. Negi, "Effective capacity: a wireless link model for support of quality of service," *IEEE Trans. Wireless Commun.*, vol. 2, no. 4, pp. 630–643, July 2003.

[129] F. Ghavimi, Y. W. Lu, and H. H. Chen, "Uplink scheduling and power allocation for M2M communications in SC-FDMA-based LTE-A networks with QoS guarantees," *IEEE Trans. Veh. Technol.*, vol. 66, no. 7, pp. 6160–6170, July 2017.

[130] H. G. Myung, J. Lim, and D. J. Goodman, "Single carrier FDMA for uplink wireless transmission," *IEEE Veh. Technol. Mag.*, vol. 1, no. 3, pp. 30–38, Sept. 2006.

[131] A. Aijaz, *et al.*, "Energy-efficient uplink resource allocation in LTE networks with M2M/H2H co-existence under statistical QoS guarantees," *IEEE Trans. Commun.*, vol. 62, no. 7, pp. 2353–2365, July 2014.

[132] C. W. Park, D. Hwang, and T. J. Lee, "Enhancement of IEEE 802.11ah MAC for M2M communications," *IEEE Commun. Letters*, vol. 18, no. 7, pp. 1151–1154, July 2014.

[133] N. Afrin, J. Brown, and J. Y. Khan, "Performance analysis of an enhanced delay sensitive LTE uplink scheduler for M2M traffic," in *Proc. 2013 Australasian Telecommun. Networks and Applications Conf.*, Nov. 2013, pp. 154–159.

[134] E. Dahlman, *et al.*, "5G wireless access: requirements and realization," *IEEE Commun. Mag.*, vol. 52, no. 12, pp. 42–47, Dec. 2014.

[135] N. A. Johansson, Y. P. E. Wang, E. Eriksson, and M. Hessler, "Radio access for ultra-reliable and low-latency 5G communications," in *Proc. IEEE Int. Conf. Commun. Workshop (ICCW)*, June 2015, pp. 1184–1189.

[136] G. Durisi, *et al.*, "Short-packet communications over multiple-antenna rayleigh-fading channels," *IEEE Trans. Commun.*, vol. 64, no. 2, pp. 618–629, Feb. 2016.

[137] M. Mousaei and B. Smida, "Optimizing pilot overhead for ultra-reliable short-packet transmission," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2017, pp. 1–5.

[138] F. Schaich, T. Wild, and Y. Chen, "Waveform contenders for 5G-suitability for short packet and low latency transmissions," in *Proc. IEEE 79th Veh. Technol. Conf. (VTC Spring)*, May 2014, pp. 1–5.

[139] D. S. Yoo, *et al.*, "Coding and modulation for short packet transmission," *IEEE Trans. Veh. Technol.*, vol. 59, no. 4, pp. 2104–2109, May 2010.

[140] X. Sun, *et al.*, "Short-packet downlink transmission with non-orthogonal multiple access," *IEEE Trans. Wireless Commun.*, pp. 1–1, 2018.

[141] D. Aziz, H. Bakker, A. Ambrosy, and Q. Liao, "Signalling minimization framework for short data packet transmission in 5G," in *Proc. IEEE 84th Veh. Technol. Conf. (VTC-Fall)*, Sept. 2016, pp. 1–6.

[142] K. F. Trillingsgaard and P. Popovski, "Downlink transmission of short packets: Framing and control information revisited," *IEEE Trans. Commun.*, vol. 65, no. 5, pp. 2048–2061, May 2017.

[143] K. Balachandran, J. H. Kang, K. Karakayali, and K. M. Rege, "Delay-tolerant autonomous transmissions for short packet communications," in *Proc. IEEE Globecom Workshops (GC Wkshps)*, Dec. 2016, pp. 1–6.

[144] B. Lee, *et al.*, "Packet structure and receiver design for low latency wireless communications with ultra-short packets," *IEEE Trans. Commun.*, vol. 66, no. 2, pp. 796–807, Feb. 2018.

[145] Q. Yang, *et al.*, "Wireless powered asynchronous backscatter networks with sporadic short packets: Performance analysis and optimization," *IEEE Internet of Things J.*, vol. 5, no. 2, pp. 984–997, April 2018.

[146] C. She, C. Yang, and T. Q. S. Quek, "Uplink transmission design with massive machine type devices in tactile internet," in *Proc. IEEE Globecom Workshops (GC Wkshps)*, Dec. 2016, pp. 1–6.

[147] A. H. E. Fawal, *et al.*, "RACH overload congestion mechanism for M2M communication in LTE-A: Issues and approaches," in *Proc. Int. Symp. on Networks, Computers and Commun. (ISNCC)*, May 2017, pp. 1–6.

[148] H. Wu, *et al.*, "FASA: Accelerated S-ALOHA using access history for event-driven M2M communications," *IEEE/ACM Trans. Network.*, vol. 21, no. 6, pp. 1904–1917, Dec. 2013.

[149] R. G. Cheng, *et al.*, "Modeling and analysis of an extended access barring algorithm for machine-type communications in LTE-A networks," *IEEE Trans. on Wireless Commun.*, vol. 14, no. 6, pp. 2956–2968, June 2015.

[150] J. P. Cheng, C. h. Lee, and T. M. Lin, "Prioritized random access with dynamic access barring for RAN overload in 3GPP LTE-A networks," in *Proc. IEEE GLOBECOM Workshops (GC Wkshps)*, Dec. 2011, pp. 368–372.

[151] T. M. Lin, *et al.*, "PRADA: Prioritized random access with dynamic access barring for MTC in 3GPP LTE-A networks," *IEEE Trans. Veh. Technol.*, vol. 63, no. 5, pp. 2467–2472, June 2014.

[152] G. Farhadi and A. Ito, "Group-based signaling and access control for cellular machine-to-machine communication," in *Proc. IEEE 78th Veh. Technol. Conf.*, Sept. 2013, pp. 1–6.

[153] T. H. Chuang, M. H. Tsai, and C. Y. Chuang, "Group-based uplink scheduling for machine-type communications in LTE-advanced networks," in *Proc. IEEE Int. Conf. Advanced Info. Networking and Applications Workshops*, March 2015, pp. 652–657.

[154] N. K. Pratas, H. Thomsen, C. Stefanović, and P. Popovski, "Code-expanded random access for machine-type communications," in *Proc. 2012 IEEE Globecom Workshops*, Dec 2012, pp. 1681–1686.

[155] H. S. Jang, *et al.*, "Spatial group based random access for M2M communications," *IEEE Communications Letters*, vol. 18, no. 6, pp. 961–964, June 2014.

[156] H. M. Gürsu, M. Vilgelm, W. Kellerer, and M. Reisslein, "Hybrid collision avoidance-tree resolution for M2M random access," *IEEE Trans. Aerospace and Electronic Systems*, vol. 53, no. 4, pp. 1974–1987, Aug. 2017.

[157] G. C. Madueno, C. Stefanovic, and P. Popovski, "Efficient LTE access with collision resolution for massive M2M communications," in *Proc. 2014 IEEE Globecom Workshops (GC Wkshps)*, Dec. 2014, pp. 1433–1438.

[158] A. J. E. M. Janssen and M. J. de Jong, "Analysis of contention tree algorithms," *IEEE Trans. Info. Th.*, vol. 46, no. 6, pp. 2163–2172, Sept. 2000.

[159] Y. Ruan, W. Wang, Z. Zhang, and V. K. N. Lau, "Delay-aware massive random access for machine-type communications via hierarchical stochastic learning," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2017, pp. 1–6.

[160] L. Alonso, R. Ferrus, and R. Agusti, "WLAN throughput improvement via distributed queuing MAC," *IEEE Commun. Letters*, vol. 9, no. 4, pp. 310–312, April 2005.

[161] E. Kartsakli, *et al.*, "Cross-layer enhancement for WLAN systems with heterogeneous traffic based on DQCA," *IEEE Commun. Mag.*, vol. 46, no. 6, pp. 60–66, June 2008.

[162] M. Li, *et al.*, "Energy-efficient M2M communications with mobile edge computing in virtualized cellular networks," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2017, pp. 1–6.

[163] L. Li, X. He, and H. Li, "Learning the spectrum using collaborative filtering in directional millimeter wave networks," in *Proc. IEEE GLOBECOM*, Dec. 2017, pp. 1–7.

[164] V. Rakovic and L. Gavrilovska, "Novel RAT selection mechanism based on hopfield neural networks," in *Proc. Int. Congress on Ultra Modern Telecommun. and Control Systems*, Oct. 2010, pp. 210–217.

[165] N. Kato, *et al.*, "The deep learning vision for heterogeneous network traffic control: Proposal, challenges, and future perspective," *IEEE Wireless Commun.*, vol. 24, no. 3, pp. 146–153, June 2017.

[166] Z. M. Fadlullah, *et al.*, "State-of-the-Art Deep Learning: Evolving Machine Intelligence Toward Tomorrow's Intelligent Network Traffic Control Systems," in *IEEE Commun. Surveys & Tut.*, vol. 19, no. 4, pp. 2432-2455, Fourthquarter 2017.

[167] J. Werbos, "Backpropagation through time: what it does and how to do it", in *Proc. of the IEEE*, vol. 78, no. 10, pp. 1550-1560, Oct. 1990.

[168] S. Hochreiter, "The vanishing gradient problem during learning recurrent neural nets and problem solutions," *Int. J. Uncertainty Fuzziness Knowl. Based Syst.*, vol. 6, no. 2, pp. 107-116, Apr. 1998.

[169] Y. Bengio, N. Boulanger-Lewandowski, and R. Pascanu, "Advances in optimizing recurrent networks," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, Vancouver, BC, Canada, May 2013, pp. 8624-8628.

[170] H. V. Hasselt, A. Guez and D. Silver, "Deep reinforcement learning with double Q-learning," in *Proc. the Thirtieth AAAI Conf. Artificial Intelligence (AAAI-16)*, pp. 2094-2100, 2015.

[171] R. C. Daniels, C. Caramanis, and R. W. Heath, "A supervised learning approach to adaptation in practical MIMO-OFDM wireless systems," in *Proc. IEEE GLOBECOM*, Nov. 2008, pp. 1–5.

[172] T. Desai and H. Shah, "Energy efficient link adaptation using machine learning techniques for wireless OFDM," in *Proc. Int. Conf. Inventive Computation Technologies (ICICT)*, vol. 3, Aug. 2016, pp. 1–4.

[173] S. K. Pulliyakode and S. Kalyani, "Reinforcement learning techniques for outer loop link adaptation in 4G/5G systems," *ArXiv e-prints*, 2017.

[174] M. Chen, *et al.*, "Data-driven computing and caching in 5G networks: Architecture and delay analysis," *IEEE Wireless Communications*, vol. 25, no. 1, pp. 70–75, Feb. 2018.

[175] Y. W. Blankenship, *et al.*, "Link error prediction methods for multicarrier systems," in *IEEE 60h Veh. Technol. Conf.*, vol. 6, Sept. 2004, pp. 4175–4179.

[176] Y. Park, *et al.*, "A new link adaptation method to mitigate SINR mismatch in ultra-dense small cell LTE networks," *IEEE Trans. Wireless Commun.*, vol. PP, no. 99, pp. 1–1, 2017.

[177] T. Park and W. Saad, "Learning with finite memory for machine type communication," in *Proc. Annual Conf. on Information Science and Systems (CISS)*, March 2016, pp. 608–613.

[178] M. A. Alsheikh, S. Lin, D. Niyato, and H. P. Tan, "Machine learning in wireless sensor networks: Algorithms, strategies, and applications," *IEEE Commun. Surveys Tuts.*, vol. 16, no. 4, pp. 1996–2018, Fourthquarter 2014.

[179] R. van Loon, "Machine Learning Explained: Understanding Supervised, Unsupervised, and Reinforcement Learning," https://www.datasciencecentral.com/profiles/blogs/machine-learning-explained-understanding-supervised-unsupervised, online; accessed 5th March 2018.

[180] J. Yoo and K. H. Johansson, "Semi-supervised learning for mobile robot localization using wireless signal strengths," in *Proc. Int. Conf. on Indoor Positioning and Indoor Navigation (IPIN)*, Sept 2017, pp. 1–8.

[181] K. Shah and M. Kumar, "Distributed independent reinforcement learning (DIRL) approach to resource management in wireless sensor networks," in *IEEE Int. Conf. Mobile Adhoc and Sensor Systems*, Oct. 2007, pp. 1–9.

[182] R. Sutton and A. Barto, "Reinforcement Learning: An Introduction", Cambridge, MA, USA: MIT Press, 1998.

[183] W. Wang, A. Kwasinski, D. Niyato and Z. Han, "A Survey on Applications of Model-Free Strategy Learning in Cognitive Wireless Networks," in *IEEE Commun. Surveys & Tut.*, vol. 18, no. 3, pp. 1717-1757, thirdquarter 2016.

[184] L. Mai, Nhu-Ngoc Dao, and M. Park, "Real-Time Task Assignment Approach Leveraging Reinforcement Learning with Evolution Strategies for Long-Term Latency Minimization in Fog Computing", in *Sensors*, MDPI, vol. 18, no. 2830, Aug. 2018.

[185] M. Chu, *et al.*, "Reinforcement Learning based Multi-Access Control and Battery Prediction with Energy Harvesting in IoT Systems," in *IEEE Internet of Things J.*, Sept. 2018.

[186] Y. Chu, *et al.*, "Application of reinforcement learning to medium access control for wireless sensor networks," *Engineering Applications of Artificial Intelligence*, pp. 23–32, Nov. 2015.

[187] C. J. Watkins and P. Dayan, "Technical note: Q-learning", *Machine Learning*, vol. 8, pp. 279-292, 1992.

[188] S. K. Sharma and X. Wang, "Collaborative Distributed Q-Learning for RACH Congestion Minimization in Cellular IoT Networks," in *IEEE Communications Letters*, vol. 23, no. 4, April 2019.

[189] A. D. Tijsma, M. M. Drugan, and M. A. Wiering, "Comparing exploration strategies for Q-learning in random stochastic mazes," in *IEEE Symp. Series on Computational Intelligence (SSCI)*, Dec. 2016, pp. 1–8.
[190] D. L. Pooole and Alan K. Mackworth, "Artificial Intelligence: Foundations of Computational Agents," in *Cambridge University Press*, Second Edition, 2017.
[191] C. g. Li, *et al.*, "Multi-intersections traffic signal intelligent control using collaborative Q-learning algorithm," in *Proc. Seventh Int. Conf. on Natural Computation*, vol. 1, July 2011, pp. 185–188.
[192] A. R. Rosyadi, T. A. B. Wirayuda, and S. Al-Faraby, "Intelligent traffic light control using collaborative Q-learning algorithms," in *Proc. 4th Int. Conf. Info. and Commun. Technol.*, May 2016, pp. 1–6.
[193] C. Lamini, Y. Fathi, and S. Benhlima, "Collaborative Q-learning path planning for autonomous robots based on holonic multi-agent system," in *Proc. 10th Int. Conf. on Intelligent Systems: Theories and Applications*, Oct. 2015, pp. 1–6.
[194] S. Kar, J. M. F. Moura, and H. V. Poor, "QD-learning: A collaborative distributed strategy for multi-agent reinforcement learning through consensus+innovations," *IEEE Trans. Signal Process.*, vol. 61, no. 7, pp. 1848–1862, April 2013.
[195] J. Ni, M. Liu, L. Ren, and S. X. Yang, "A multiagent Q-learning-based optimal allocation approach for urban water resource management system," *IEEE Trans. Automation Science and Engg.*, vol. 11, no. 1, pp. 204–214, Jan. 2014.
[196] J. Hu and M. P. Wellman, "NASH Q-learning for general-sum stochastic games," *J. Mach. Learning Res.*, vol. 4, no. 6, pp. 1039-1069, 2003.
[197] S. Choi, *et al.*, "Automatic configuration of random access channel parameters in LTE systems," in *Proc. IFIP Wireless Days (WD)*, Oct. 2011, pp. 1–6.
[198] M. Rahimiyan and H. R. Mashhadi, "An adaptive Q-learning algorithm developed for agent-based computational modeling of electricity market," *IEEE Trans. Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 40, no. 5, pp. 547–556, Sept. 2010.
[199] K.-S. Hwang, S.-W. Tan, and C.-C. Chen, "Cooperative strategy based on adaptive Q-learning for robot soccer systems," *IEEE Trans. Fuzzy Systems*, vol. 12, no. 4, pp. 569–576, Aug. 2004.
[200] P. Y. Glorennec and L. Jouffe, "Fuzzy Q-learning," in *Proc. Int. Fuzzy Systems Conf.*, vol. 2, Jul 1997, pp. 659–662.
[201] P. Jamshidi, *et al.*, "Self-learning cloud controllers: Fuzzy Q-learning for knowledge evolution," in *Proc. Int. Conf. Cloud and Autonomic Computing*, Sept 2015, pp. 208–211.
[202] M. N. ul Islam and A. Mitschele-Thiel, "Cooperative fuzzy Q-learning for self-organized coverage and capacity optimization," in *Proc. IEEE 23rd Int. Symp. (PIMRC)*, Sept. 2012, pp. 1406–1411.
[203] Y. Maeda, "Modified Q-learning method with fuzzy state division and adaptive rewards," in *Proc. IEEE Int. Conf. Fuzzy Systems*, vol. 2, 2002, pp. 1556–1561.
[204] F. Pervez, *et al.*, "Fuzzy Q-learning-based user-centric backhaul-aware user cell association scheme," in *Proc. 13th Int. Wireless Commun. and Mobile Computing Conf.*, June 2017, pp. 1840–1845.
[205] M. Mendil, *et al.*, "Fuzzy Q-learning based energy management of small cells powered by the smart grid," in *Proc. IEEE 27th Annual Int. Symp. PIMRC*, Sept. 2016, pp. 1–6.
[206] J. Wu, J. Liu, Z. Huang, and S. Zheng, "Dynamic fuzzy Q-learning for handover parameters optimization in 5G multi-tier networks," in *Proc. Int. Conf. Wireless Commun. Signal Process. (WCSP)*, Oct 2015, pp. 1–5.
[207] C. Wu, S. Ohzahata, and T. Kato, "Flexible, portable, and practicable solution for routing in vanets: A fuzzy constraint Q-learning approach," *IEEE Trans. Veh. Technol.*, vol. 62, no. 9, pp. 4251–4263, Nov. 2013.
[208] Z. Deng, *et al.*, "Combining model-based Q-learning with structural knowledge transfer for robot skill learning," *IEEE Trans. Cognitive and Developmental Systems*, pp. 1–1, 2017.
[209] A. Sharma, *et al.*, "Model based path planning using Q-learning," in *Proc. IEEE Int. Conf. Industrial Technology (ICIT)*, March 2017, pp. 837–842.
[210] E. S. El-Alfy, Y.-D. Yao, and H. Heffes, "A model-based Q-learning scheme for wireless channel allocation with prioritized handoff," in *Proc. IEEE GLOBECOM*, vol. 6, 2001, pp. 3668–3672.
[211] J. Cao, M. Ma and H. Li, "A group-based authentication and key agreement for MTC in LTE networks," in *Proc. IEEE Global Communications Conference (GLOBECOM)*, Anaheim, CA, 2012, pp. 1017-1022.
[212] S. Bhandari, S. K. Sharma and X. Wang, "Device Grouping for Fast and Efficient Channel Access in IEEE 802.11ah based IoT Networks," in *Proc. IEEE Int. Conf. Commun. (ICC) Workshops*, June 2018.
[213] L. Liu, *et al.*, "Sparse Signal Processing for Grant-Free Massive Connectivity: A Future Paradigm for Random Access Protocols in the Internet of Things," in *IEEE Signal Processing Magazine*, vol. 35, no. 5, pp. 88-99, Sept. 2018.
[214] H. Li, K. Ota, and M. Dong, "Learning IoT in edge: Deep learning for the internet of things with edge computing," *IEEE Network*, vol. 32, no. 1, pp. 96–101, Jan. 2018.
[215] M. Song, *et al.*, "In-Situ AI: Towards Autonomous and Incremental Deep Learning for IoT Systems," in *2018 Proc. IEEE Int. Symp. High Performance Computer Architecture (HPCA)*, Vienna, Feb. 2018, pp. 92-103.
[216] S. Venkataramani, K. Roy, and A. Raghunathan, "Efficient embedded learning for IoT devices," in *Proc. 21st Asia and South Pacific Design Automation Conference (ASP-DAC)*, Jan. 2016, pp. 308–311.

Shree Krishna Sharma (S'12–M'15-SM'18) is currently Research scientist at the SnT, University of Luxembourg. Prior to this, he worked as a Postdoctoral Fellow at the University of Western Ontario, Canada in the areas of 5G Wireless Communications and Internet of Things (IoT) systems. He also worked as a Research Associate at the SnT being involved in different European, national and ESA projects after receiving his PhD degree in Wireless Communications from the University of Luxembourg in 2014. In the past, he had held an industrial position as a Telecom Engineer with Nepal Telecom, and part-time and full-time teaching positions with three different universities in Nepal. His current research interests include 5G and beyond wireless, Internet of Things, machine learning, edge computing and optimization of distributed communications, computing and caching resources.

He has published more than 90 technical papers in scholarly journals and international conferences, and has over 1600 google scholar citations. He is a Senior Member of IEEE and is the recipient of several prestigious awards including "2018 EURASIP Best Journal Paper Award", "Best Paper Award" in CROWNCOM 2015 conference and "FNR Award for Outstanding PhD Thesis 2015" from the FNR, Luxembourg. He has been serving as a Reviewer for several international journals and conferences; as a TPC member for a number of international conferences including IEEE ICC, IEEE GLOBECOM, IEEE PIMRC, IEEE VTC and IEEE ISWCS; and an Associate Editor for IEEE Access journal. He organized a special session in IEEE PIMRC 2017 conference, worked as a Track Co-chair for IEEE VTC-fall 2018 conference, and has recently published an edited book on "Satellite Communications in the 5G Era" with the IET as a lead editor.

Xianbin Wang (S'98-M'99-SM'06-F'17) is a Professor and Tier 1 Canada Research Chair at Western University, Canada. He received his Ph.D. degree in electrical and computer engineering from National University of Singapore in 2001.

Prior to joining Western, he was with Communications Research Centre Canada (CRC) as a Research Scientist/Senior Research Scientist between July 2002 and Dec. 2007. From Jan. 2001 to July 2002, he was a system designer at STMicroelectronics. His current research interests include 5G technologies, Internet-of-Things, communications security, machine learning and locationing technologies. Dr. Wang has over 350 peer-reviewed journal and conference papers, in addition to 29 granted and pending patents and several standard contributions.

Dr. Wang is a Fellow of Canadian Academy of Engineering, a Fellow of IEEE and an IEEE Distinguished Lecturer. He has received many awards and recognitions, including Canada Research Chair, CRC President's Excellence Award, Canadian Federal Government Public Service Award, Ontario Early Researcher Award and six IEEE Best Paper Awards. He currently serves as an Editor/Associate Editor for IEEE Transactions on Communications, IEEE Transactions on Broadcasting, and IEEE Transactions on Vehicular Technology and He was also an Associate Editor for IEEE Transactions on Wireless Communications between 2007 and 2011, and IEEE Wireless Communications Letters between 2011 and 2016. Dr. Wang was involved in many IEEE conferences including GLOBECOM, ICC, VTC, PIMRC, WCNC and CWIT, in different roles such as symposium chair, tutorial instructor, track chair, session chair and TPC co-chair.