IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received September 3, 2020, accepted September 21, 2020, date of publication October 9, 2020, date of current version October 26, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3029903

Towards Energy Efficient 5G Networks Using Machine Learning: Taxonomy, Research Challenges, and Future Research Directions

AMNA MUGHEES, MOHAMMAD TAHIR, (Senior Member, IEEE), MUHAMMAD AMAN SHEIKH, AND ABDUL AHAD

Department of Computing and Information Systems, School of Science and Technology, Sunway University, Subang Jaya 47500, Malaysia

Corresponding authors: Mohammad Tahir (tahir@sunway.edu.my) and Amna Mughees (amna.m@sunway.edu.my)

ABSTRACT As the world pushes toward the use of greener technology and minimizes energy waste, energy efficiency in the wireless network has become more critical than ever. The next-generation networks, such as 5G, are being designed to improve energy efficiency and thus constitute a critical aspect of research and network design. The 5G network is expected to deliver a wide range of services that includes enhanced mobile broadband, massive machine-type communication and ultra-reliability, and low latency. To realize such a diverse set of requirement, 5G network has evolved as a multi-layer network that uses various technological advances to offer an extensive range of wireless services. Several technologies, such as software-defined networking, network function virtualization, edge computing, cloud computing, and small cells, are being integrated into the 5G networks to fulfill the need for diverse requirements. Such a complex network design is going to result in increased power consumption; therefore, energy efficiency becomes of utmost importance. To assist in the task of achieving energy efficiency in the network machine learning technique could play a significant role and hence gained significant interest from the research community. In this paper, we review the state-of-art application of machine learning techniques in the 5G network to enable energy efficiency at the access, edge, and core network. Based on the review, we present a taxonomy of machine learning applications in 5G networks for improving energy efficiency. We discuss several issues that can be solved using machine learning regarding energy efficiency in 5G networks. Finally, we discuss various challenges that need to be addressed to realize the full potential of machine learning to improve energy efficiency in the 5G networks. The survey presents a broad range of ideas related to machine learning in 5G that addresses the issue of energy efficiency in virtualization, resource optimization, power allocation, and incorporating enabling technologies of 5G can enhance energy efficiency.

INDEX TERMS 5G, energy efficiency, millimeter wave, machine learning, massive MIMO, SDN, NFV, CRAN, HetNet.

I. INTRODUCTION

Until the 4^{th} generation of mobile communication standard, the focus was to deliver high data rate. Over the past years, technologies such as the Internet of Things (IoT) have resulted in billions of connected devices and the generation of an enormous volume of data. It is expected that the traffic volume will increase exponentially and will become 1000 folds by 2020. Also, the number of connected devices will continue to increase exponentially. It is expected that there will be approximately 50 billion devices by 2021 [1]. Due to this, the focus has shifted to other design requirements to deliver a diverse set of service that includes:

- **Enhanced Mobile Broadband (eMBB):** This use case is in line with the previous generation use cases where the aim is to provide a higher data rate. 5G aims to provide 10 to 100x improvement over 4G and 4.5 networks, which is equivalent to 10Gbps.
- **Ultra-Reliable Low-Latency Communication (URRLC):** This use is geared to those mission-critical services that require extremely low error rate (high

The associate editor coordinating the review of this manuscript and approving it for publication was Tawfik Al-Hadhrami.

reliability) and low latency. These applications usually do not require a high data rate.

- **Massive Machine Type Communications (mMTC):** With the rise of the IoT, the ubiquity of devices has necessitated the development of connectivity standards that can support high device density with low power consumption. Usually, IoT devices operate on battery and are expected to last several years (10 years). The previous generations did not consider such scenarios.

Such requirements have resulted in the need to design a communication infrastructure that can readily adapt to changes. In this regard, 5G networks are being designed to provide pervasive networking, high data rates, coverage, reliability, and low latency. Meeting such diverse requirements has also resulted in increased ICT energy consumption. By 2025, the ICT industry itself could be responsible for 30% of power consumption globally, where data centers alone account for 3% of carbon emission [2]. In mobile networks, 80% of the total energy of cellular networks is consumed by base stations and has pivotal importance for energy efficiency improvements [3]. For instance, to improve the coverage and meet capacity requirements, a large number of small cells will be deployed. Small cells make the network denser, which leads to more energy consumption. According to the small cell forum, in 2024, all 4G small cells will be replaced by 5G small cells, reaching 13.1 million installations in 2025 [4]. Furthermore, massive MIMO also increases the power consumption due to more hardware components required for each BS [5]. Therefore, there is a need for efficient resource management and spectral sharing for improved energy efficiency.

Another factor of energy consumption of the network devices is the energy requirement that differ from peak hours to less load time. Base station related factors add up to the energy consumption every time a new function is added to the network. Such operations add cost to the Operating Expenses (OPEX) by adding dedicated hardware. This can be fixed by introducing virtualization of the infrastructure. As networks are expanding massively, the network functions are not easy to manage on dedicated devices and need a paradigm change in conventional management of the network. Network Function Virtualization (NFV) can handle such provisions by eliminating the need for hardware and implementing separate software-based functions [6]. This virtualization not only provides flexibility but also reduce OPEX and CAPEX cost. Different virtual machines can use a common node to implement NFV functions. For example, in the case of RAN, one virtual machine providing baseband processing and other virtual machines for core network user policies can use the single node. This reduction in hardware deployment through virtualization can lead to a more energy-efficient network. Furthermore, energy efficiency is directly affected by data rates; hence, a balance is required between energy consumption and QoS [7]. From a service provider's perspective, degradation in QoS is unacceptable. Therefore, the focus should be on maintaining the right level of energy efficiency so that the QoS does not affected.

Considering the energy constraints and versatile network requirements, traditional approaches are not enough for network optimization. In this regard, machine learning techniques are being used to let the system learn intelligently from data and optimize the overall operation of the network. For example, virtualization technology improves energy efficiency and resource utilization and can result in up to 50% of energy-saving [8]. To achieve energy-efficient virtualization and network optimization, machine learning can further improve energy-efficiency through load sharing and consolidation. Likewise, energy consumption in the data centers, which consume most of the energy, can be minimized by intelligent resource allocation and management through machine learning approaches.

Several machine learning approaches can be applied to improve the energy efficiency of 5G networks. In supervised learning, the model is trained on a set labeled data to predict optimal solutions. An example of a supervised learning application is massive MIMO for energy efficiency, in which channel estimation and detection are considered a problem because of a high number of antennas. Unlike supervised learning, unsupervised learning works on unlabeled input and is suitable for clustering and dimensionality reduction. For example, unsupervised learning can be used to cluster BS with similar behaviors for energy-efficient operation in varying load conditions. The use of reinforced learning approaches for energy-efficient solutions is suitable when little or no prior data is required for processing.

This paper aims to provide a comprehensive survey of recent advances in energy efficiency techniques at access, edge, and core network utilizing machine learning. Power allocation, resource optimization, pre-coding, and other energy-efficient techniques have been discussed in this paper with regards to 5G and energy efficiency.

A. MOTIVATION

Cellular technologies have seen gradual evolution from 1st generation to 5th (5G) for meeting the demand in terms of bandwidth, throughput, latency and jitter [9]. In 2017, there were around 8.4 billion connected devices, among them 2.7 billion were smartphone users. It is expected that the connected devices will reach 20.4 billion, with 3.5 billion smartphone users in 2020 [10]. With this tremendous increase in smartphone users, wearables, and IoT devices, providing high data-rates, coverage, and low latency is becoming a challenge. Furthermore, each generation gave rise to energy consumption due to the addition of hardware to support new applications and requirements. It is expected that 5G will give significant rise to this traditional trajectory of energy consumption. The need to support high data rates and a large number of devices are making these networks more hungry for energy. The energy consumption of 5G is four times than of 4G [11]. Currently, 0.5% of the entire world's energy is consumed by mobile network [12]. According to the Ericsson

Mobility Report [13], in 2025, the amount of user data will increase four-time compared to today's network. As a result, energy efficiency is a significant factor in 5G as compared to earlier generations. Several technologies are being integrated in the 5G network to realize the diverse set of services. These technologies include Software-Defined Network (SDN), Ultra-Dense Network (UDN), Network Function Virtualization (NFV), multi-access edge computing, and cloud computing. However, the integration of various technologies creates several challenges in terms of energy efficiency. For example, in Ultra Dense Network (UDN), even though energy consumption is decreased due to low transmission power but the increase in computational requirement results in higher energy consumption in a dense scenario. This increase in computation power will continue to increase over time [14]. Also, to serve growing demand, massive MIMO technology is used to serve denser sites. However, in massive MIMO, the balance among linearity and efficiency is critical. The measurements of power amplifiers directly affect the energy efficiency of the massive MIMO system. Making them linear escalates the cost, and non-linearity eventually affects energy efficiency. The demanding nature of future technologies requires appropriate hardware, efficient learning techniques (that can take energy efficient decisions intelligently), and new network design to break the energy curve.

Machine learning, in this regard, can deal with several challenges faced in 5G networks due to the integration of several new technologies in an energy-efficient manner. The motivation behind this article is to address the growing need for intelligent networks that take intelligent decisions to make the network energy efficient. The future generations of wireless communication and 5G networks are too diverse to make decisions based on pre-defined and fixed rules. The ability to interact with the environment and learn from the generated data makes it possible to design a network that adapts to improve the network energy efficiency. Furthermore, in 5G or even future wireless networks, machine learning techniques can help with various non-linear and non-convex problems that may arise due to 5G deployment and network design.

In terms of expenses, energy efficiency is also a significant cause of concern for mobile network operators. Besides economic importance for network operators, energy efficiency has ecological importance as well. SMARTer2020 published a report in 2015 that by the end of 2020, the carbon emission will reach 1.27GT (around 2.3% of global emission) [15].

B. COMPARISON WITH EXISTING LITERATURE

Various surveys has been conducted on 5G technologies [12], [16], energy efficiency [14], [17]–[19], and on enabling technologies such CRAN [20]–[23], SDN [21], NFV [6], UDN [24], massive MIMO [25], [26], HetNets [27] and MEC [28]–[30]. The surveys on energy efficiency is either limited to the energy harvesting techniques, system design [18], virtualization [21] or architecture [19]. Table 1 explains some of the research contributions and limitations of existing review on the issue of energy efficiency.

To the best of our knowledge, this survey is unique as it captures the application machine learning from a different perspective. The classification and review presented in this paper allow the researcher to understand the significance of various machine learning techniques in 5G for energy efficiency using the end-to-end approach.

C. NOVELTY AND CONTRIBUTION

This paper aims to provide a comprehensive survey on addressing energy efficiency in 5G network challenges encountered in enabling technologies such as mmWave, CRAN, massive MIMO, NFV, hetNets, small cells, and SDN using machine learning. Few studies have been made on energy efficiency in their respective enabling technologies in the literature. However, none of them categorized the network by covering whole network requirements from the core to the access network. This survey incorporates various machine learning-based energy-efficient techniques in the 5G network for researchers to benefit and explore further. Specifically, the main contributions are highlighted as follows:

- A detailed discussion on applying machine learning for improving energy efficiency focused on the enabling technologies for 5G.
- A review on the energy efficiency aspect of 5G technology using end to end layered approach that involves core network, access network, and edge network.
- Taxonomy of application of machine learning in 5G networks for energy efficiency found in the literature.
- Open issues and future research directions for achieving energy efficiency in the 5G ecosystem.

D. ARTICLE ORGANIZATION

The rest of the paper is organized, as shown in Figure 1. Section II provides a brief overview of the evolution of 5G enabling technologies, energy efficiency, and machine learning. Section III presents energy efficiency's needs, challenges, metrics, and projects. The taxonomy is discussed under section IV, underlining the importance of energy efficiency. Section V provides future directions and challenges to help researchers to continue research in this area. Section VI finishes the paper with a conclusion.

II. BRIEF OVERVIEW OF 5G, ENERGY EFFICIENCY & MACHINE LEARNING

This section gives a brief overview of 5G, energy efficiency and machine learning and we also bring forward the need for machine learning and its applicability in energy efficiency.

A. 5G AND ENABLING TECHNOLOGIES

Extensive coverage, reduced latency, greater bandwidth and higher data rates are the much-anticipated features of 5G. 3GPP initiated 5G studies with its estimated completion by 2020 in rel 16. As compared to 4G, 5G data rate requirements

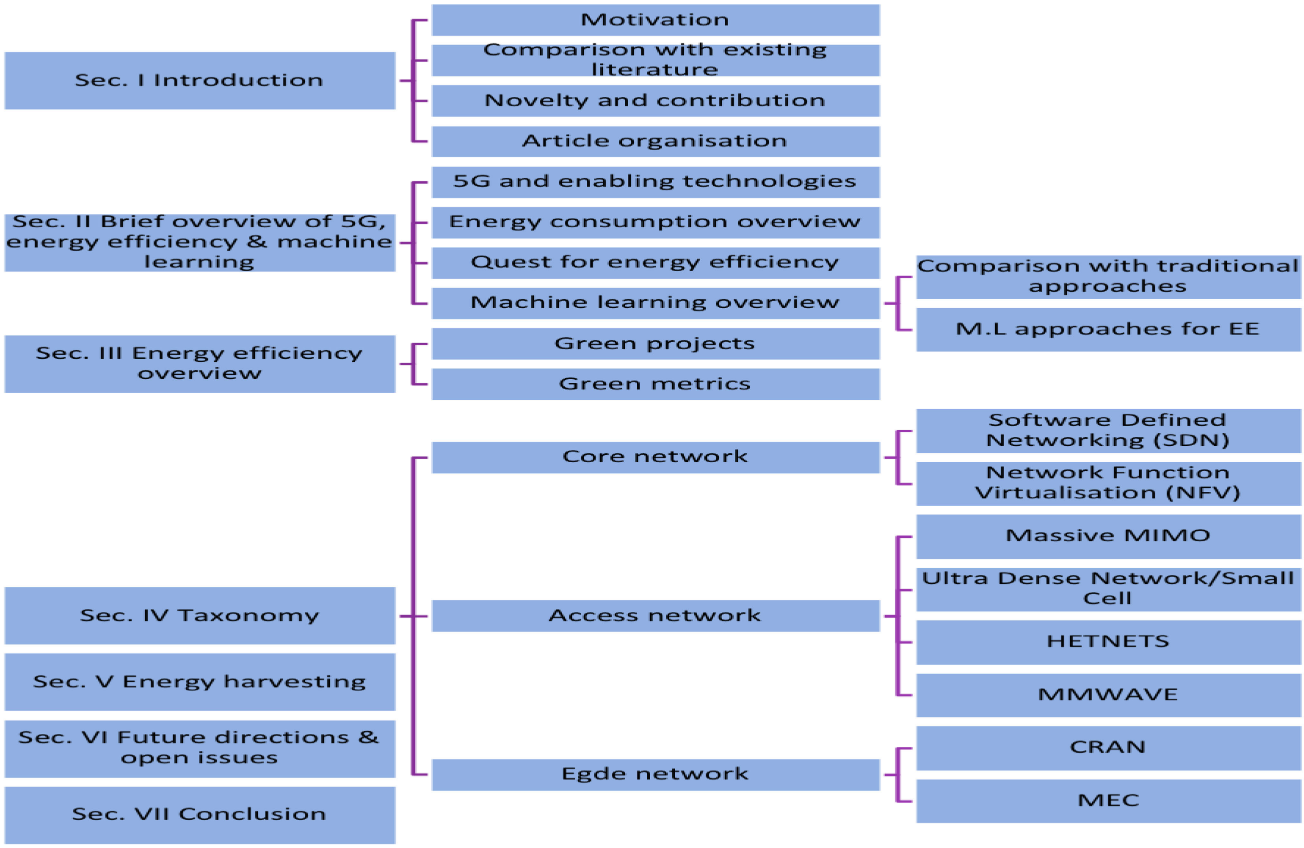


FIGURE 1. Outline of the paper.

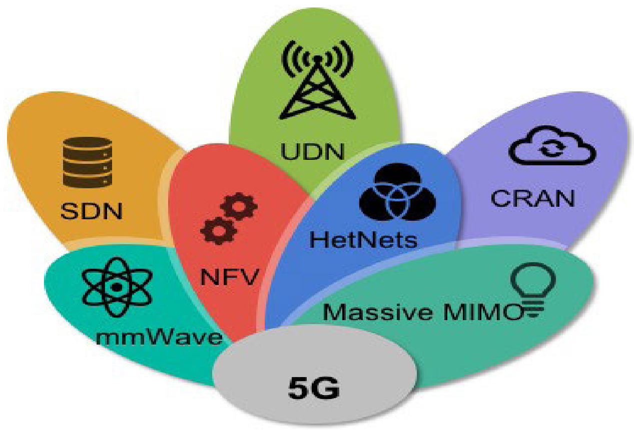


FIGURE 2. Outline of technologies analysed for energy efficiency in this survey.

are increased by ten folds and hence requires higher data rate and bandwidth. 5G aims to cater to the new spectrum utilization (C-band), which has higher frequencies to fulfil bandwidth needs. A summary of the performance difference between 4G and 5G is shown in Table 2 and the enabling technologies of 5G in Figure 2. Moreover, with the introduction of VANET, IoT, 5G assisted smart healthcare [32], real-time controlling of the machine and high data rate requires the integration of following enabling technologies:

- **Millimeter waves** fall under 30GHz to 300GHz, which provides more bandwidth to users. Larger bandwidth means a higher data transmission rate. However, at extremely high frequencies, attenuation increases, which means that mmWaves cannot be used for long-distance communication. However, these high frequencies work well for a short distance and are used in small cells.
- **Massive MIMO** is connecting multiple antennas to a single base station to provide improved spectrum utilization and data rate. Furthermore, it results in reduced interference due to efficient beamforming and spatial multiplexing. Despite several advantages, some issues need to be addressed, such as pilot contamination, channel correlation and interference management.
- **Heterogeneous Network (HetNet)** involves deploying different radio technologies and along with legacy systems to provide seamless coverage and capacity. Inter-tier and intra-tier interference, resource allocation, and optimization are the most significant bottlenecks to energy efficiency.
- **Ultra Dense Network** or dense deployment of small cells provide users with better coverage and throughput

TABLE 1. Existing recent surveys on energy efficiency.

Reference survey papers	Author Contribution	Limitations
[17]	This paper surveys energy efficiency issues at network level and radio access level. Energy conservation of 5G has been explored in terms of BS (architecture, switching techniques, interference reduction and deploying small cells), network (resource sharing and allocation), SDN level (energy monitoring and management) and machine learning.	The energy efficiency issues related to edge network part is not addressed, also all technologies are not addressed with respect to machine learning.
[18]	Author addressed the energy efficiency aspect of 5G technology with respect to resource allocation, network planning, energy harvesting, and hardware. The importance of network planning has been highlighted to achieve more energy efficiency, offloading techniques, dense deployment of the network, and network benefit-cost ratio to achieve QoS is also discussed.	The research is well summarized while highlighting the network planning and hardware issue that also affects energy efficiency. However, the latest trend of cloud and virtualization is not addressed that has significant benefits in energy efficiency.
[19]	In this paper, the author proposed a multi-tier architecture based energy-efficient method. This multi-layer architecture caters aspects starting from tier 1 planning, hardware solutions, spectrum sharing to tier 11, which is massive MIMO.	The proposed architecture and the survey of 5G technology are well defined with the focus on development and planning.
[31]	Author discussed energy efficiency aspects based on the increase in capacity. The study is based on spectrum efficiency, Spectrum reuse, and spectrum resources. The author discussed green techniques and how LLC RAN, and SAN can be achieved.	Green 5G technologies specifically are addressed in spectrum context but with no highlight of machine learning. Regardless of its importance in energy efficiency that machine learning can provide prominent results in large scale network problems.

TABLE 2. Performance difference between 4G and 5G (based on Verizon and 5G-ppp analysis).

Performance Criteria	4G	5G
Peak speed	1.4 Gigabit/s	10 Gigabit/s
Latency	40-50 milliseconds	<10milliseconds
Connectivity	10K-100K devices supported/mi^2	1 million devices supported/mi^2
Energy efficiency	90% more used energy/bit	90% less used energy/bit
Mobile data volume	1/100 Terabytes/s/Km^2	10 Terabytes/s/Km^2

[24]. Ultra-dense networks comprise of numerous small cells that can be deployed using the base station, relays, or Remote-Radio Heads, with all the functionalities of a conventional macrocell along with the benefit of low power consumption. Deployment of numerous small cells undoubtedly address the capacity and coverage issue but result in increase cost and management due to the deployment of a large number of base stations [33].

- **Software Defined Networking (SDN)** is one of the most crucial components that provides management facilities to large and high-speed networks by splitting the data plane and control plane. In the case of 5G network, SDN can orchestrate and control applications/services in a fine-grained and network-wide manner resulting in the more efficient management of the network.
- **Network Functions Virtualization (NFV)** decouples functions (firewall or encryption services) from allocated hardware into connectable blocks and shifts those functions to virtual switches, servers, or inexpensive hardware. Specialized hardware, used in networks are expensive and difficult to program to adapt to changing network requirements. Furthermore, the hardware has interoperability issues making the network less flexible. Therefore, by decoupling hardware from the associated network functions provides improved scalability and flexibility.
- **Cloud Radio Access Network (CRAN)** is a widely accepted paradigm to provide features like central processing, energy-efficient infrastructure, real-time computing, and improved spectral utilization. The three components Baseband Unit (BBU), Remote Radio Head (RRH), and Optical Transport Network (OTN), are used

to provide base-station functionalities, radiofrequency signalling, and their transmission to the cloud network. The use of densely deployed RRHs which are controlled by C-RAN enhances the scalability and improves network capacity.

- **MEC Mobile Edge Computing**, similar to CRAN technology, also aims to improve the RAN. CRAN focus on centralization and cloud services. In contrast, MEC aim towards decentralization by pushing the computation, processing and storage close to the user. MEC decreases the latency and reduces network congestion in the backhaul network. ETSI proposed the idea initially to resolve the network congestion issue by using the distributive computing approach. Certain features of MEC have been introduced in 4G as well.

B. ENERGY CONSUMPTION OVERVIEW

Cellular systems have made headway from purely analog to digital technology that provides ubiquitous connectivity. The focus of each generation was to provide higher data rate and capacity. The energy efficiency was not given significant consideration until 3G. According to a study on 2G and 3G power consumption [34], for a 15 minute time window, GSM consumes an average of 1.08kW to 1.20kW. Whereas, UMTS average power consumption for the same 15 minute time window was around 0.19kW to 0.22kW.

According to another study, 5G power consumption at peak hours is 1200W to 1400W, which is 300% to 350% greater than of 4G [35]. The power consumption varies significantly between peak and off-peak hours. To address this issue, researchers proposed to put the base station radios in sleep mode as the majority of the electricity consumption was due to base stations and RF transceivers (76% of total power consumption). The base station switching strategy is an efficient technique to save energy and improve energy efficiency. This behavior of turning the base station ON/OFF depends on the fluctuations of traffic patterns over time and space. China Mobile started using the same BS ON/OFF strategy from 2009 and can save around 36 million kWh [36]. Considering the potential of this BS sleep strategy, researchers started working on this technique to get more benefits. Considering the 5G network for this strategy, it becomes more challenging because of different enabling technologies and heterogeneity of the network. Other hurdles in the energy-efficient practices were site architecture, their distribution for coverage, the power consumption of electronic devices and cooling systems (24% of total power consumption). An estimation of consumed energy in ICT, data centers, carbon footprint, RBS and core network is shown in Table 3.

C. QUEST FOR ENERGY EFFICIENCY

In 1990, the Information & Computer Technology emerged, which led to work on electricity consumption. In the early times of ICT, electricity consumption was mainly divided into commercial and domestic use of wired and wireless devices [37]. If the increase in ICT consumption from 2010 to 2015 is compared, the power consumption of communication networks increased from 185TWh to 805Twh [38], that is an almost 31% increase. According to Ericsson, the energy consumption of 5G is much smaller than of 3GPP and 4G, which is also shown in Figure 3. Over the last decades, many reports highlighted the energy crunch of ICT sector. According to [39], it is estimated that in 2030, energy consumption will increase to 21%. Recent research by Ericsson [40] focused on environmental aspects and sustainability perspective of communication, where the target is to make ten times more energy-efficient than the networks in 2017.

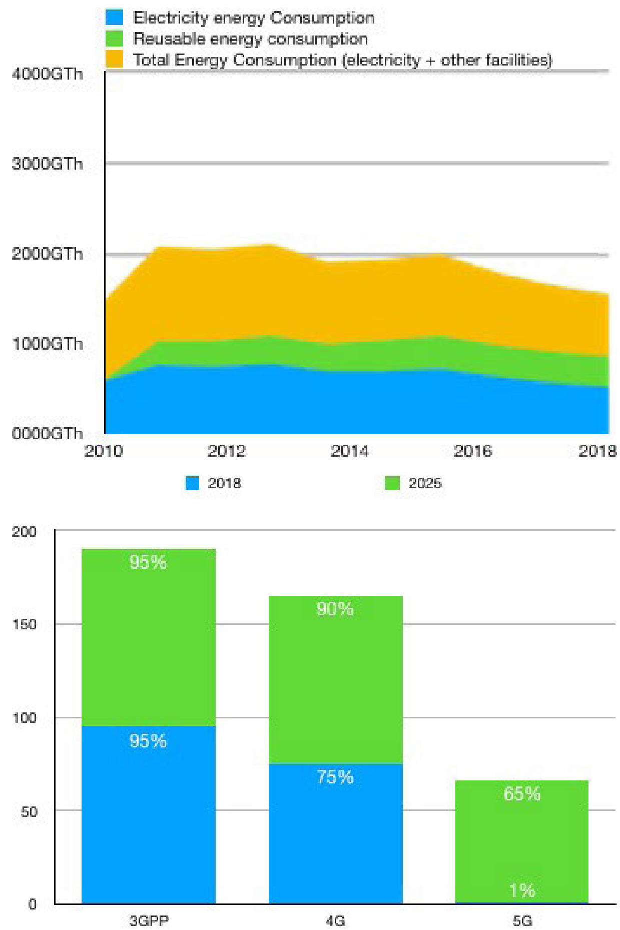


FIGURE 3. Energy usage of electricity, renewable sources and electricity plus other facilities of communication system (a) Ericsson [13] estimation for coverage of different technologies from the year 2010 to 2018 (b) Total energy consumption of 2018 and 2025.

In 3G, energy efficiency was in very early stages and was not a significant part of the research. The new modulation scheme, access schemes and channel coding required more power compared to 2G [41]. With the launch of CDMA, energy efficiency improved because of its efficient power control and resource utilization. Thus, researchers started working on how efficiently the power could be used in terms of data centers in 3G and base stations. In [42], power reuse factors for 3G were presented that discussed increasing energy efficiency by re-utilizing resources. For the optimized

TABLE 3. Estimation of the consumed energy in ICT, data centers, carbon footprint, radio base station and core network for the time period from 2001 to 2030.

Year	World's Power Consumption	Carbon Footprint in %	Carbon Footprint in Mto	RAN Electricity Consumption	OPEX	BS Density	Consumed Energy (ICT)
2005 (3G)	133602TWh	1.3%	86Mto	49TWh	low	4-5BS/ km^2	3.9%
2015 (4G)	21000TWh	1.5% - 3%	170Mto	77TWh	high	8-10BS/km^2	3.5%-7%
2020 (5G)	23000TWh	6%	3.5% - 235Mto	86TWh	high	40-50BS/km^2	3%

energy-efficient network researchers worked on dense low power networks [41], renewable energy supply [43], power management, power reuse and CDMA deployment. In Figure 4, the energy consumption comparison among different parts of the communication system is shown for the year 2013 and 2025.

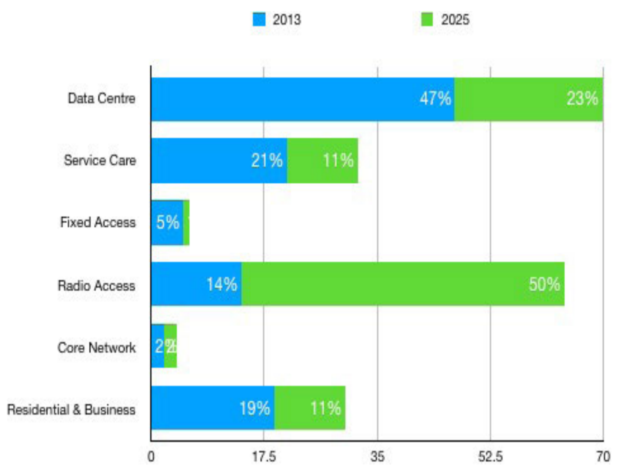


FIGURE 4. Energy consumption estimation in communication system in 2013 and 2025 [44].

In 4G, with the introduction of MIMO and OFDM, researchers aimed to explore both in terms of spectral efficiency and capacity. At that time, energy concerns were not given significant consideration [45]. Due to limitations, MIMO was replaced by multi-user MIMO, that provide much better results in terms of energy efficiency. OFDM work as a multi-user diversity aiming at spectral efficiency as well as efficient use of energy. According to [46] an efficient architecture was required in initial stages of 4G to make network energy efficient, the author also discussed green energy standardization, metrics and techniques required for 4G. Also, the consumed energy of ICT for the year 2011 was around 4.7% of the worlds energy consumption [47], [48]. Base stations consume approximately 80% of the total cellular network, whereas among this 70% is due to amplification and cooling purposes [48].

D. MACHINE LEARNING OVERVIEW

The engineering of developing intelligent programs (artificial intelligence) started in the 1950s. Machine Learning (which does not need any categorical programming to learn) started evolving in the mid-1980s and matured over time. Machine learning is the sub-field of Artificial Intelligence (A.I) and is further sub-categorized into Supervised, Unsupervised, and Reinforced Learning. Deep Learning is also a sub-field of machine learning which evolved in 2010 and can be classified as supervised, unsupervised, and reinforced. Recently, machine learning based approaches has been applied to many research fields for solving problems like resource management & allocation [49], power allocation [50], [51], cell sleeping [52], pre-coding [53], [54]. In this section, we look at different machine learning approaches used for the energy-efficient wireless network. A brief discussion on the benefits of applying machine learning based over the conventional approaches for improving energy efficiency in the 5G and beyond network is also presented.

1) COMPARISON WITH TRADITIONAL APPROACHES

The new wireless technology-based model needs high data rates and diverse applications that challenge traditional technology in learning and decision-making processes. Some of the M.L advantages over traditional approaches are as below:

- Machine learning can learn from its data, whereas the traditional techniques are mostly hard coded.
- Particularly in large scale problems, learning speed significantly improves.
- In conventional approaches, a new set of instructions needs to be coded for every new function.
- Machine learning has autonomous decision-making capability.
- Software development every-time for new applications is a costly project.

Besides its benefits, there are also some of the drawbacks associated with machine learning related to training. Machine learning integration for large scale processing, security, and how the application-level implementation is possible for research theories [55].

2) MACHINE LEARNING APPROACHES FOR ENERGY EFFICIENCY

Machine learning is further divided into supervised learning, reinforced learning, and unsupervised learning. There is also further classification on these techniques to be best utilized for particular problems. Machine learning techniques discussed in this paper are also presented in Table 4. Supervised learning is the best approach for channel-related problems such as channel estimation, its detection, and learning its behavior to take future predictions. This is because supervised learning produces the output from the collected data based on forgoing experiences. However, for the networks where the raised problems are unknown, reinforced learning is best to use, such as resource allocation and management. Reinforced learning has the potential to adjust its strategy to obtain the required results. It learns from the results systematically and improves the decisions further. Unsupervised learning is slightly different from supervised learning as it is better utilized for clustering and spectrum sensing problems in wireless networks. It functions on its own to learn the network and solve the problem and thus solve more complex problems compared to supervised learning. Figure 5 depicts machine learning classification and learning approaches that are frequently used in 5G enabling technologies and energy efficiency problems.

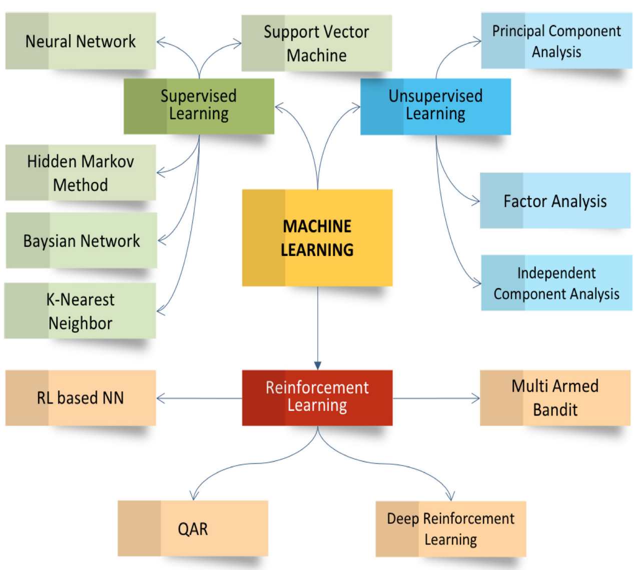


FIGURE 5. M.L classification and techniques used for energy efficiency.

III. ENERGY EFFICIENCY OVERVIEW

The goal of connecting billions of devices is non-sustainable in terms of both economic and environmental concerns. The rate at which network design demand is increasing, it will eventually lead to 1000 times more power consumption than today's network. This energy crunch lead researchers to set up a Green Touch Consortium [71] to research over the critical matter of green energy efficient network. According to [18], resource allocation, network planning and deployment, energy harvesting & transfer, and hardware solutions are the broad categories that can increase energy efficiency. According to the Shannon formula, with the increase in bandwidth, the energy consumption factor also rises [31]. Massive MIMO is a promising technology to deal with efficiency concepts in terms of both spectrum and energy. Multiple antennas attached to a BS can either sleep or turn off mode to increase energy efficiency. In [26], the authors worked on the trade-off between spectral efficiency and energy efficiency. The proposed work presented resource allocation to increase energy efficiency using the benefits of the Rayleigh fading channel model for massive MIMO. Authors in [16] worked not only on energy efficiency as well as on end to end delay. Besides spectrum efficiency, increase in bandwidth, deploying small cells, D2D/M2M communication, and ultra-dense networks, energy efficiency is another interlinked challenge that needs to be addressed. However, 5G is promising to decrease energy consumption by 90% [72]. According to [73], energy efficiency can be calculated as a ratio between the energy consumption of a system and Joules per bit capacity, which is an energy consumption ratio.

$$ECR = \frac{E^{sys}}{C^{sys}} \tag{1}$$

A. GREEN PROJECTS

The telecommunication sector is among the top energy consumers. Data centers, base stations, and core networks have the highest carbon footprint and energy consumption compared with overall ICT energy consumption. It is assumed that in 2030 almost 20% of the global CO2 emission will decrease [74]. Despite all the new IoT, architecture, and traffic growth, the concerns are to meet the minimum requirement of energy consumption. The need for green communication has led researchers to work on various projects to achieve dual benefits. Firstly, reducing the energy cost as it affects profit calculations directly. Secondly, by reducing the carbon footprint, that is also an alarming environmental aspect. For this purpose, many joint ventures and projects mentioned in Table 5 were initiated over the past years to cut the energy consumption factor. In [75] 5GrEEN is discussed, which took the initiative to highlight the need for energy efficiency in 5G. In 2010, the Green Touch consortium aimed to improve energy consumption to 90% at the end of 2020. 5G Infrastructure Association covering the private side of 5PPP mutually launched a 5G Infrastructure Evaluation Association Group in 2006 [76]. The target was to develop international standards, cooperation among 5G standards for the long run evaluation, and provide more secure internet. Some of the other research projects over past years are [77]–[81].

Many other 5G projects are also in line for automotive, vertical industries, and 5G long term evolution. In 2018, 5G-EVE, 5G-VINNI, and 5GENESIS started working on infrastructure enhancement aspects to create a foundation that eventually helps to mount end to end 5G. 5G SMART, 5GROWTH, and 5G-SOLUTIONS are the projects started

TABLE 4. M.L techniques discussed in this paper for energy efficiency.

5G technology	Machine Learning technique	Reference
SDN	Unsupervised Learning	[56] [57]
	Reinforced Learning	[56]
	Q-learning	[58]
	Neural Network	[58]
NFV	Supervised Learning	[59]
	Reinforced Neural Network	[59]
	Deep Learning	[60]
	Deep Reinforcement Learning	[61]
massive MIMO	Deep Learning	[50]
	Machine Learning	[54]
	Deep Neural Network	[53]
UDN	Reinforced Learning	[62]
	Neural Network	[62]
HetNets	Deep Reinforcement Learning	[49]
		[63]
		[64]
		[65]
		[66]
mmWave	Deep Learning	[54] [67]
	Deep Neural Network	[53]
CRAN	Machine Learning	[68]
	Deep Neural Network	[52]
MEC	Supervised Learning	[69]
	Deep Learning	[70]

in 2019 for smart energy, machine-based remote operations, architecture, and dynamic use of the network. In most of these projects, the work was on energy efficiency aspect in 5G with focus on load balancing. However, a lot of work is required for on-demand response modeling and service-level optimization specifically to the power side.

B. GREEN METRICS

The volume of the network is expanding by the factor of 10 every five years. Energy efficiency is now required as an important component and needs to be involved in every development aspect. From architecture level to deployment, network-level to facility-level green metrics roles are meaningful. Energy efficiency seems to be understandable if it can be measured. They are used to measure the consumed energy also for enhancing efficiency by comparing performance trade-offs. The following are the international standardization bodies to study the telecom equipment to enhance the energy efficiency globally [82]:

- International Telecommunication Union (ITU) has the main focus on energy consumption reduction, energy metrics and environmental protection & recycling. Their focus is also on Green House Gas (GHG) impact and how ICT can contribute to GHG.
- European Telecommunications Standards Institute (ETSI) is concerned with the life cycle of telecom network, telecom infrastructure, ICT equipment and aim towards the reduction in energy consumption. Its main focus is on power optimization, energy consumption, power feeding and assessment of ICT energy impact globally.
- Climate Change Standardization Landscape
- Alliance for Telecommunication Industry Solutions (ATIS) is a standard organization that provides the evolving ICT industry solution. It deals explicitly with telecommunication equipment energy and power consumption at different load levels.

Green metrics can be implemented at equipment, facility, and network-level to measure and enhance efficiency [83]. Some of the network-level metrics are Energy Consumption Rating (ECR) [75], Energy Efficiency Rate (EER) [84], Access Per Cycle (APC) [85], (ECG) [83], (EEER) are used to measure energy efficiency at network level, performance evaluation and other aspects related to network capacity and coverage. Power Usage Efficiency PUE and its subordinate metric Data Centre Efficiency (DCE) [86] is implemented on the facility level for power. Telecommunication Equipment Energy Efficiency Rating (TEEER) [87] and Telecommunication Energy Efficiency Ratio (TEER) energy metrics were developed by ATIS as equipment level metrics. Some other energy metrics are mentioned in Table 6.

IV. TAXONOMY

Energy efficiency has gained its importance in the design and operations of 5G networks. The energy efficiency covers the whole network from the radio access network, core network,

TABLE 5. List of green projects.

International Research Projects	Year	Objectives	Conducted research	Aim of EE gain
eNergy nEutral Wireless SEnsor Networks (NEWSENs)	2019 to 2021	To design an architecture that works on renewable energy that works on RF technologies.	RF energy & wireless sensors	Renewable energy for wireless sensors
Innovative ultra-BROadband ubiquitous Wireless communications through terahertz transceivers (iBROW)	2015 to 2018	To develop energy efficient and low-cost wireless communication platform which is capable to fulfil future requirements	Platforms providing connectivity between fiber optics and high-speed wireless communication.	Cost and energy efficient platforms development
scalaBle and grEen wireleSs coMmunications for a sustAinable netwoRked socieTy (BESMART)	2017 to 2019	A wireless network that is self sustainable, can share energy with other node for longer lifetime of network nodes and can configure itself by allocating efficient radio resource.	Distributed mobile networks	100% coverage in urban areas Reduction in energy cost Self configured network utilizing energy efficient resource allocation
MATILDA	2017 to 2019	To integrate 5G applications with demanding infrastructure and network functionalities	Smart Cities C-RAN Virtual Resources	Up to 70% reduction in energy consumption
A NOvel Radio Multiservice adaptive network Architecture for the 5G era (5G NORMA)	2015 to 2017	To develop a kind of network architecture that will cope the growing need of traffic because of heterogeneous networks	For 5G flexible BS, controllers that are software based and can be centrally connected Software enabled RAN	To increase energy efficiency by selecting multi service efficient option.
Green Radio Project	3 years	To redesign backhaul, efficient resource allocation and multi-hop routing	Base station and handsets of mobile data services	Power efficient Dynamic spectrum access
Green Machine Learning for 5G and Beyond Resource Optimisation	2021 to 2023	To delevelop green machine learning algorithms	Radio resource management	To lead the network towards intelligence and green communication
Mobile and wireless communications Enablers for the Twenty-twenty Information Society (METIS) II	2015 to 2017	Radio access network designing	Technology components	Integrating technologies for efficient 5G framework
5G Infrastructure Public Private Partnership (5G PPP)	2015 onwards	To evaluate IMT-2020 proposal	Network elements	To save up to 90% energy Advance privacy 1000 x more wireless area coverage
Green Touch	2010 to 2015	Improving EE 1000 times by 2020 by communication	Architecture and specification	It was assumed that energy factor will be cut down with a factor of 10 with 2010 baseline
ViruWind	2015 to 2018	For sustainable energy constraint	SDN NFV	Horizon 2020 & To use wind sector energy in cost reduction

TABLE 6. List of green metrics.

EE Metrics	Level	Targets	Features	Unit	Pros & Cons
Energy Consumption Rating (ECR)	Network level/Equipment level	Energy metric	A ratio is measured among maximum data throughput when the power is at peak	Watts/Gbps	no network load consideration
Energy Consumption Rating-Variable Load (ECR-VL)	Network level/Equipment level	Energy metric	Dynamic Power management	watt/bps	Works actively
Energy Efficiency Rate (EER)	Network level/Equipment level	Energy metric	Output data rate with respect to consumed power	bps/watt	Reciprocal of ECR
Telecommunication Energy Efficiency Ratio (TEER)	Equipment level	Energy & Power metric	Calculates energy and power efficiencies	x/watt (x depends on taken parameter)	Includes environmental tests also
Telecommunications Equipment Energy Efficiency Rating (TEEER)	Equipment level	Energy metric	Tests variable load efficiencies (total energy consumption as weighted sum)	$-log\frac{Gbps}{Watt}$	not able to work on all properties of system.
Normalized Power Consumption (NPC)	Equipment level	Power metric	Used for broadband wired access	mwatts/Mbps/km	can connect multiple subscribers
Power Usage Efficiency (PUE)	Facility level	Power metric	used to improve operational efficiency of data centers	Watt	works at macro level hence not able to assess individual level energy efficiency
Data Center infrastructure Efficiency (DCiE) & DCE	Facility level	Power metric	Inverse of PUE	Watt	located within IT devices to calculate the total output
Energy Proportionality Index (EPI)	Equipment level	Network devices	Measurement is on the basis of consumed energy at idle mode and maximum load	Percentage	EPI=(Emax-Eint)/PM * 100%
Key Performance Indicator of Energy (KPIEE)	Network level	Energy metric	Used for evaluation and testing	-	Significant practical approach
PI rural	Network level	Power metric	evaluates rural areas network performance	$km^2/Watt$	Only for rural areas
PI urban	Network level	Power metric	based on average busy hour traffic	users/Watt	For urban areas only

data centers, and technologies. In this section, we will discuss the taxonomy of 5G enabling technologies to improve energy efficiency using ML techniques. The proposed taxonomy is shown in Figure 5. Several approaches, including resource

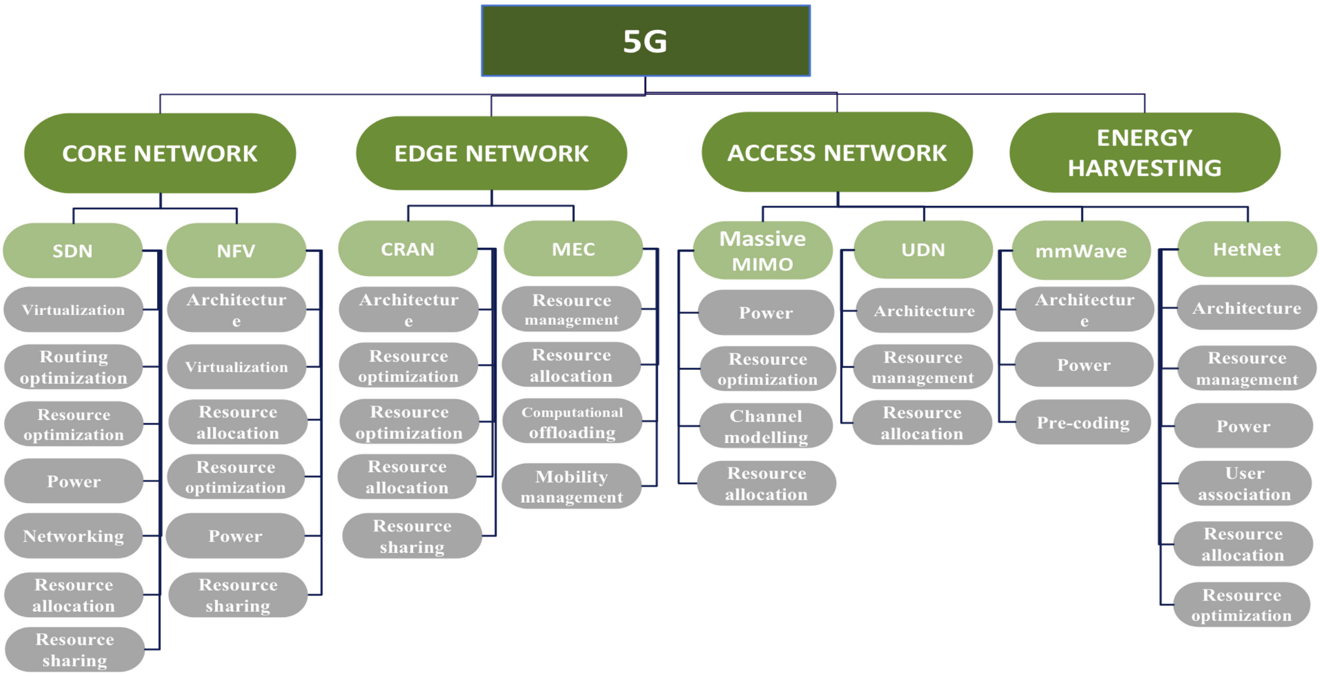


FIGURE 6. **Taxonomy of machine learning application for energy efficiency in 5G.**

management, resource sharing, bandwidth allocation, and power allocation, have been discussed to improve energy efficiency. The below section gives a detailed overview of energy efficiency in 5G and its solutions with the help of machine learning.

A. CORE NETWORK

1) SOFTWARE DEFINED NETWORKING (SDN)

5G networks are required to be more resilient and self-autonomous. The 5G infrastructure is based on Software Defined Networking (SDN). In this network architecture, it is possible to control the network centralized and intelligently to use software applications. All the communication between applications and services can be managed from the centralized center allowing dynamic adaption in real-time. Many of the ICT companies like Yahoo, Google, Facebook, and Cisco opted software-defined networking in their data centers and network equipment [88].

SDN separates the traditional vertical integration by detaching the data plane from the control plane, improving user experience by providing a higher data rate and low latency. Because of this separation, network switches start working as forwarding devices. A logically centralized controller controls the traffic replacing routers, switches, and traditional table forwarding format. These switches and controllers are connected via well-defined programmed interfaces. These application programming interfaces (API) are used to employ control by controller [89]. One of the widely use API is OpenFlow, and the famous controllers are NOX, POX, Beacon, Maestro, MUL, RISE, OpenDayLight, and NOX-MT [20]. This further helps in managing the forwarding plane and providing access to all other parts of the heterogeneous network. Some of the major advantages of SDN is intelligent networking, resource virtualization, and session management.

Other than several benefits, there are few issues in SDN that needs to be investigated further. One of the limitations is an overhead increase because of excessive requests to the controller. To solve the congestion problem, a framework is proposed based on low-cost load-balanced route management (L2RM) to monitor the burden of traffic in fat-tree DCN [90]. In the second phase, adaptive route modification (ARM) is triggered based on load. A dynamic polling system is adopted to update statuses to reduce overload. The ARM mechanism proposed works in two ways. Firstly, it helps switches to remain updated and remove old data to avoid overloading buffer. Secondly, it wakes up only when necessary, thus saving cost and energy. In terms of overloading, the proposed system is effective for the energy efficiency factor. Data centres utilize about 10% to 20% of power, and over-furnishing of the data centre with resources causes a lot of energy inefficiency. SDN is one method to reduce energy wastage and efficiently use power in peak hours and resultant traffic consolidation. In cloud computing, there is an agreement between cloud providers and organizations to ensure provided services and quality of service to customers. Because of overbooking, there are chances of service level agreement (SLA) violation. [91] offers a technique to improve energy efficiency based on the overbooking ratio, which is decided based on link information and investigating the correlation between VMs. When the scenario of overloading happens, VM shifts to another host to reduce SLA violation.

One approach to solve these issues is by making them intelligent enough to learn from their environment. SDN is very useful in realizing such smart solutions. Machine learning is an efficient way to work in conjunction with SDN to facilitate numerous challenges for optimization, organization, and network resources management. The latest computing technologies like TPU also can cope with the high computational requirements of machine learning. These specialized purpose processors like TPU and GPU have enough capacity to incorporate the use of machine learning techniques to give results within milliseconds [92], [93]. Most of the work done on SDN is on traffic, security, and routing. To the best of our knowledge, significant research is required to combine machine learning with SDN to improve energy efficiency. Here, we investigate machine learning techniques used in SDN to enhance performance and energy. In machine learning, the feature extraction approach is used to extract the most relatable data that incorporates feature learning (which helps differentiate different features from raw data) and feature reduction. The output depends on the features selected; the more complex features require more considerable training. More training means high computational power and memory.

Switches, ports, and active links consume a lot of power in any SDN. One way to conserve energy is by minimizing the power factor of these switches and links. Moreover, desirable network performance is achieved by changing the flow paths to get maximum throughput and minimum delay. To configure and control the network, the controller should have all the updated information related to the network. Based on this information, SDN can reconfigure the topologies. An energy-efficient routing based hybrid solution is proposed in [56]. A supervised and reinforcement learning framework named HyMER with a primary focus on energy efficiency and routing is discussed. In the first stage, supervised learning is used for feature reduction via PCA, training, and then testing. RL is used in the second phase, where Q-learning is used for network status components and links utility for dynamic routing based on repeated steps till destination. The proposed technique provides energy-efficient and maintains network performance as well. However, this technique relies on extensive training using historical data. If the training data is not sufficient, the output may be biased.

Another approach used for the efficient use of energy is combining SDN with machine learning [57]. It is implemented on the POX controller for traffic information and topology extraction. The principal component analysis (PCA) is used to reduce the feature size. The data with reduced features, along with topology, is fed to train the model. The proposed framework consists of three modules—traffic manager to store data of traffic flow and the status of topologies. Machine learning learns from historical data and draws graphs for traffic load. Linear regression is used to build a regression model to train data sets. In SDN, because of the iterative update of OpenFlow, the routing efficiency decrease. The purpose of routing techniques is to decrease energy factors, especially by minimizing packet delivery time.

Energy efficiency and routing schemes are interlinked with each other. Using neural network [58] developed a routing scheme that makes the controller work centralized to the data flow. In this technique, the data flow path can also be predicted, which helps to fulfil QoS requirements. A central controller controls data collection, neural network packet creation, training, routing, data processing, and rerouting. Moreover, in the data plane side, switches contribute to flow forwarding, NN creation, and route prediction. The Control plane also monitors network and topologies discovery. When the packet is received, it is first analyzed by the switch and then forwarded upon received request. Hop is predicted based on received NN data. With every hop headers of packets are changed accordingly. At the time of failure or overloaded network, a reroute request is generated. For intelligent routing, the neural network is trained by the controller based on collected data. One of the benefits of machine learning is its data-driven nature. As already mentioned, the SDN controller has the benefit of global network visibility, proving helpful in collecting data to feed machine learning. Not only this, with the help of machine learning, the configuration is possible in real-time.

SDN has been employed in transport networks [23], wireless sensor networks [94], network function utilization (NFV) [95], cloud radio access networks (C-RAN) [22], Internet of Things (IoT) [21] and edge computing because of its intrinsic strengths. Some other benefits that SDN are granular, security, centralized control, less operation cost, software-based traffic scanning, cloud level abstraction, and guaranteed QoS.

2) NETWORK FUNCTION VIRTUALIZATION (NFV)

Next-generation wireless networks are all about independent service-related functions. Consequently, virtualizing network services is an approach to minimize use of hardware. Network Functions Virtualization (NFV) relieves network operators from increasing OPEX costs by reducing the conventional purposed hardware, installation, and up-grading for new services. NFV has the biggest advantage in energy efficiency. Almost 30% of energy consumption can be decreased with its implementation into 5G architecture [17]. NFV benefits network operators in several ways, such as:

- No need for dedicated hardware also results in energy saving
- No location dependency
- It is assumed that there is no energy consumption when BBU will be in the idle condition in an absolute state.
- Improved operational efficiency & reduced cost
- Seamless and reliable interoperability with the latest technologies
- Real-time and potent virtualization

The short distance between the user and virtual machines can also save power because of shorter paths. There are several standardization efforts to increase the adoption of NFV. ETSI community ISG NFV is in the phase of release 4, working toward NFV evolution, automation, management, and orchestration [96]. Apart from ETSI other standardization bodies ONF, IRTF, IETF, OPNFV, ATIS, BBF, OVF and 3GPP are also contributing to NFV standardization [6]. Virtual functions are different from the logical system and are virtually installed on commodity hardware. These are like the blocks which can be used for numerous purposes. Virtual network functions (VNF) are virtualized tasks implemented by the NFV platform, providing security, load balancing, and other EPC functions.

NFV usually works on high-performance modes and utilizes the CPU optimally, mostly Dynamic Voltage and Frequency Scaling (DVFS) mode, which helps maintain energy efficiency. Energy consumption using NFV is similar to a dedicated CPU energy consumption at its high processing mode. In virtual environments, where the physical machines are used for virtual network function (VNF), deployment needs proper research to increase power consumption and inefficient resources utilization. Also, traffic processing never remains the same during the peak and average hours, which also leads to energy wastage. Idle servers consume the same amount of power but waste more than half of the energy because of no utilization. Machine learning is an effective way to reduce energy consumption by handling VNFs, especially on average traffic hours. An energy-efficient NFV based architecture on 5G [97] investigated the effect of active users in the network. The aim was to investigate the energy consumption. All mobile core entities (mobility management entity, serving gateway, packet data network gateway, and policy & charging rules function) are formulated in one virtual machine as core network virtual machine (CNVM). The RRH and BBU are decoupled, where the BBU is implemented in a virtual machine referred to as BBUVM. The only way to pass the traffic is through CNVM and BBUVM. With the focus on energy consumption, the architecture provides services across flexible administration. Results proved that up to 38% of consumed energy could be saved using the proposed approach.

To handle complex networks, NFV Management and Orchestration systems (MANO) are deployed to manage virtualized infrastructure, communication and network infrastructure, NFV entities, and their life cycles [98]. As discussed earlier, ETSI projects are in phase IV; however MANO framework can improve management and orchestration services for NFV. Most relevant projects are open source MANO [99] for resource orchestrator, openbaton [100] for, throughout service orchestration, Juju for VNFM, open stack tracker [99] for optimization and resource allocation and X-MANO [101] for sensitive information. With the help of SDN, many services can be deployed over the network by providing flexible VNFs. There can be many virtual functions such as firewalls, servers, storage units and load balancers which are conventionally defined as middleboxes [102]. For proper flow in an organized network, all virtual functions should be interconnected. This connectivity to provide services throughout is called service function chaining (SFC). These SFCs support multiple VNFs to provide traffic flow and services.

Proper resource optimization is a mandatory aspect of quality of service. Thus, resource estimation is a critical aspect of a smooth service that should be appropriately utilized. A semi-supervised machine learning-based resource demand novel model is proposed to avail the NFV environment characteristics to do the predictions [59]. Long short-term memory (LSTM) model, which is also a type of recurrent neural network (RNN), can use past and current learning data. After training, the data is further processed to remove ambiguities. The collected SFC data is then used to predict performance. The result shows that the proposed technique is giving improved results as compared to the simple LSTM technique. Another resource allocation technique in NFV is proposed using Deep Learning [60]. It identifies the network traffic by utilizing the timing characteristics.

B. ACCESS NETWORK

1) MASSIVE MIMO

Among several metrics, bandwidth efficiency is one of the important factors to be chosen for the next-generation network. The rapid increase in carbon emission and the growing power demand of communication networks resulted in enhanced energy efficiency metrics. For this, MIMO became significant due to energy-efficient capability and enhanced throughput. In massive MIMO, the concept of numerous base station deployment is the same as TDD operations like conventional MIMO. However, it does not require additional power for the transmission and bandwidth [25]. Multiple Input Multiple Output (MIMO) is not a new concept. It has been deployed in 4G with one BS support to eight antenna ports. Although it is an old concept, it was not deployed fully as conventional BS was considered more cost-effective and MIMO more complicated.

As MIMO concept enters into 5G, a larger number of antennas can be deployed, which is referred to as massive MIMO. Massive MIMO gives many advantages over MIMO such as increased throughput, enhanced spectral efficiency, increased signal to noise ratio, increased capacity, reduced latency, increased data rate, and energy efficiency [25]. Despite the earlier mentioned massive MIMO benefits, antennas placement is still an issue in massive MIMO. The basic rule to place an antenna with spacing is half the signal wavelength to provide no-correlation among antennas. Massive MIMO, with hundreds of channels at one BS, leads to increased spatial diversity. However, channel hardening results when the faded channel behaves as a non-fading channel [103]. The random interference is still there in massive MIMO, but it has little effect on communication. One way to achieve zero correlation is by decreasing the wavelength; the higher the frequencies, the lower are the chances for correlation.

Transferring more bits per Hertz bandwidth makes the network more spectrum efficient. However, another challenge is to make the network more energy efficient. This can be somehow possible with spatial modulation. Massive MIMO is much better than MIMO in higher bandwidth, enhanced energy efficiency, and spatial freedom. However, the pilot contamination problem occurs because of inter-user interference using the same reference signal and is an inherent problem. Because of the frequency limitation, the cells are bound to use the same frequencies blocks. The orthogonal pilot sequences lead to pilot contamination. The pilot contamination issue can occur in both ordinary BS and massive MIMO. However, it still gathered more attention in the case of massive MIMO because of the reuse of pilots. As the channel difference between conventional MIMO and massive MIMO is significant, a pilot contamination problem in any BS is reduced by switching among different pilots (among large pilot sequences).

In the case of massive MIMO, due to more active terminals and more reuse of pilots (as pilots do channels estimation), it is challenging to avoid pilot contamination. Whereas in conventional MIMO, it can be overcome as more the terminals, higher will be pilot contamination. Regular Pilot (RP) and Superimposed Pilot (SP) are the two most frequently used methods to suppress pilot contamination [104]. In RP, data and pilot sequences are transferred in a fragmented way while adjusting the pilots sequence. In contrast, SP is an old concept, where data symbols and pilot together instead of placing them in time or frequency. The superimposed pilot has also been advocated for real-time implementation through simulation in [105]. The proposed work advocates that the superimposed pilot has shown better results in hybrid systems. Uplink MIMO provides significant power saving because of the higher array gain. This is possible because of the coherent signal integration. In contrast, in the downlink, the beams are focused on a particular direction for users. Cell-free massive MIMO is a new concept. A large number of access points are deployed in a distributed manner to serve numerous users. These access points (AP) work in the same TDD and consist of single or numerous antennas. This concept gives high energy efficiency and spectral efficiency because of less inter-user interference even with furnishing many users at the same time-frequency because of the less distance among antennas [106]. The cell-free massive MIMO concept is similar to small cell deployment; the more significant difference is the deployment of many AP vs. single AP. Massive cell-free MIMO's energy efficiency factor depends on power allocation and consumption, channel estimation, and the selection of best access points. Although massive MIMO has matured [107], supporting both multi-user and massive MIMO. Several researches are being carried into spectral efficiency, pilot contamination/decontamination, power allocation factor, and energy efficiency.

A deep learning-based approach is used in [50] to let the system learn from its user equipment location to allocate downlink power. A massive MIMO network is considered on TDD for both user equipment and base station operations. The initial optimal powers are calculated using the Monte Carlo method, and the training part is done offline. The deep learning approach is used to let the network allocate power based on user location. It is proved that max production strategy for neural network is more advantageous in complex calculations than conventional approaches. When used together for power allocation, max-min and maximum production approach showed incompetence, which is then addressed through a different neural network using the LSTM layer. Although the simulation provided promising energy-efficient power allocation results, the massive MIMO scenario considered is not significant to prove its efficiency for the real-time environment. However, deep learning is a promising tool to solve the real-time high computational problem as they can learn iteratively from the environment. Another work proposed [53] on pre-coding integrates deep neural networks because of its capability to reduce the computational complexity. It utilizes structural information through the training stage. Distributed massive MIMO is also considered an energy-efficient way to allocate resources. Compared to conventional massive MIMO, its throughput, energy efficiency, and channel modelling in a complex environment are noticeable [108]. Also, the beamforming in massive MIMO results in improved energy efficiency on the targeted coverage area [109].

2) ULTRA DENSE NETWORK/DENSE SMALL CELL

Ultra-dense networks were required to fulfill the needs of those areas that are highly packed and require more cell deployment. There are three ways to enhance the capacity of the network (a) by enhancing the spectrum efficiency (b) broadening the bandwidth, and (c) by deploying more cells. The concept of dense deployment is found back in 4G, stuffing the same area with many cells. However, the cost factor and interference among those macrocells surfaced, which has more diminishing returns. The better idea was to move towards cells that provide more coverage to end-users and less deployment cost. Small cells (picocells, femtocells) provide coverage closer to the end-users and require less power, with almost 90% more capacity. Small cell deployment does not entirely negate the need for macrocell as its coverage area is too small compared to macrocells, which is why macrocells are still required to cover a large area. Ultra-dense small cell network extends the coverage on the benefits of low power consumption and deployment cost.

Apart from coverage area frequency reuse is also another factor of small cells. Small cells are divided into four types: (a) Picocells are mainly used to increase the capacity up to 100m and can be deployed indoors and outdoors. (b) Femtocells are also a type of small cell with the same characteristics as picocells, except the coverage is 10-30m. (c) Relays are the macro extension and need proper planning for indoor and outdoor deployment to reduce interference. Its coverage area is a little larger than of femtocells (up to 100m). (d) RRHs can only be deployed outdoor but with proper planning, as this is normally connected with BS through a wired connection or microwave links. The provided coverage is around up to 100m. The consumed power of small cells is same for all approximately 100mW (indoor) and 0.25W to 2W(outdoor) [24].

Deploying small cells and making the area dense does not solve all the problems. Some new issues like interference and more energy consumption also arise. To resolve such issues, the integration of different techniques are required for challenges to be addressed. Capacity is not imperatively dependent on dense deployment on cells, many other factors like interference, frequent handoffs, excessive energy consumption, and mobility. The focus of 5G is to use higher frequency ranges, and hence ultra-dense networks are considered an efficient feasible solution. It benefits in utilizing frequencies more productively, deploying small base cells densely (to cater to exploding traffic demands), and better energy consumption. Therefore, the need for an energy-efficient network became indispensable.

A three-layer learning solution is provided for dense small cell networks in [62], macro base stations and small base stations are deployed where power grid feed energy is used to Macro Base stations (MBSs) and energy harvesting techniques like solar cell provide power to Small Base stations (SBSs). SBS also has the feature of on/off to save energy. The proposed first layer takes decision locally at SBS by making the best use of resources. It is composed of the heuristically accelerated reinforcement learning approach. The second layer takes decisions at MBS and is made up of a multi-layer feedback neural network and is also responsible for the energy factor. This approach gave promising results in terms of radio resource management for self-organizing networks and energy efficiency.

3) HetNets

Initially, HetNets were introduced to increase spectral efficiency and capacity in LTE-advance. At that time, macrocell was used in the majority for large coverage perspective, whereas small cell was to fill in the gaps. The power consumption of macrocells is also noticeably large as compared to pico, femto, and microcells. With 5G, enhanced energy consumption, high data rates, and large coverage capacity become essential for the network. Dense deployment is the key to provide better user association and cell selection. However, many other limitations and challenges emerge with it that will be discussed further in this section. Although small cells are more power-efficient than macrocells and the HetNet are an optimal choice, but still, there are some hurdles:

- With the growing installation of macrocells and small cells, not only the installation cost increases but also the functioning power cost of towers and equipment increases.
- Coverage gaps
- interference among small cells and macrocells
- Increased OPEX

The first thing that directly affects energy efficiency or power consumption is the network's architecture, the number of nodes, and deployment. As mentioned above, with massive deployment of small cells, the coverage increases massively, but the cost of these deployments and maintenance also increases. The urgent need for enhanced spectral efficiency in 5G and energy efficiency both caught the eyes of researchers. In [110] the spectral efficiency, trade-off with energy efficiency is also discussed. However, multiple BS are deployed, focusing on HetNets deployment, and they consume maximum power even if the traffic is minimal, resulting in OPEX and environmental energy efficiency concerns. HetNets consists of small-cells and macrocells that are also differentiated based on power consumption but can also be managed by the same operator. In this scenario, resource management is required to utilize the same frequency so that the coverage is not affected. Another method is to use discontinuous bands individually to cell types to minimize interference. After the dense deployment and architecture solutions, the major problem is how the users will be allotted to BS cells. The problem of associating a user with a BS cell, termed as user association, also affects the network's performance. To solve the user association problem efficiently, more accurate network information is required. In a dynamic environment, there is a need to solve the problem more efficiently and intelligently. For this, machine learning is an emerging technology to solve such issues. In [64], the problem of energy efficiency is effectively solved in uplink HetNets along with user association optimization using deep reinforcement learning. For such non-linear problems, traditional methods of problem-solving are not enough. Deep reinforcement learning can solve decision making and resource allocation problems efficiently in real-time.

According to [27], enhanced spectral efficiency and load across base stations are important to avoid congestion and better user association. The user association decision solely depends on the quality of service, requests and requirements, priority, and availability of resources [111]. Among some of the previously used user association methods is by using the maximum SINR for the association. However, in the case of lots of user association to that particular base station, the performance degrades significantly. Several researchers also worked on user association and power allocation together [49], [63] [64] using deep reinforcement learning (DRL) and deep neural networks. According to [65], DRL is an efficient way to resolve complex issues. Another problem re-association, which is of the same importance as user association [66]. With the introduction of different cell sizes, user association becomes demanding. Channel conditions, bandwidth, load on the base station, and power consumption account for user association. Based on the transmitter and receiver characteristics, the available spectrum can be reused, thus leading HetNets to be more spectrum efficient. This reuse spectrum characteristic results in less power usage for both uplink and downlink, eventually making HetNets more energy efficient. Efficient resource allocation is essential for the energy efficiency of the network. Deep neural network can solve complex non-linear problems such as resource allocation, user association and resource management. A machine learning approach used for resource allocation in [112], works by rewarding the QoS for each femtocell and macrocell user. This helps to allocate power allocation and gain efficient energy more effectively as the environment changes dynamically. Another research work focused on resource allocation used Convolutions Neural Network (CNN), which also increased energy efficiency. The idea is to subdivide the resource allocation issue into classification and regression problems and achieve the energy efficiency decisions with low-level complexity [113].

Small cells, when deployed in the HetNets, use the same spectrum as of the microcell layer. Macrocells are the most power-consuming cells. Deploying small cells can result in spectrum reuse and reduce energy consumption. However, the interference is always there in small cells and microcells, even with the spectrum reuse strategy. The e-ICIC feature helps to mitigate this problem by allowing macrocells to reuse the almost underutilized spectrums. Thus small cell technology can help the network to cater data needs of several connected devices and massive data traffic. Although small cells can provide high data rates for communication, they are also prone to high energy consumption. It is essential to consider the energy consumption of BS and network benefits for operators, i.e., remaining profitable while consuming less energy.

4) mmWave

Most of these are using microwave frequency that is below 6GHz. As the number of devices is increasing expeditiously, these frequency ranges are becoming congested. Researchers are exploring new strategies to either use unused frequencies, new spectrum, or substitute technologies. One feasible solution to this problem is the use of a spectrum above 30GHz termed as millimeter wave. 30GHz to 300GHz is the under-utilized spectrum segment where 24GHz is used for microwave communications and is unlicensed, and 28GHz are put up for auction in 2019. Federal Communication Commission (FCC) promoted the auction of the high-band mmWave spectrum in 2019. The contiguous spectrum to be available for mmWave communication is 37GHz and 39GHz, which make 2400 MHz and an additional 1000 spectrum for 47GHz. In early 2020, 2400MHz of 5G spectrum were accessible for auction [114]. However, most of the research is already done on 28GHz, 71GHz to 76GHz, and 81GHz to 86GHz band.

However, over the past years, researchers have doubts about its sustainability. It has an extremely short range of wavelengths that is the reason it is best used for line of sight communication and has the benefit of faster data transfer. Because of its wavelength nature, it can easily get blocked by any obstacle. Various mmWave limitations are:

- Deterioration in mmWave signals can badly affect propagation.
- mmWave provides high data rates but also vulnerable (sensitive and easily affected by blockage). The sensitivity of mmWave to weather and especially rain causes severe attenuation and effect communication because the size of rain droplets and mmWave wavelengths are approximately the same [115].
- The major benefit of mmWave over other wireless communication systems is ten-fold enhancement in frequency to carry data. Not only this, because of less wavelength, but more antennas arrays can also be installed on transmitter and receiver base stations [116].

Hybrid precoding has become a significant research area. It helps mmWave take benefits from beamforming when combined with spatial multiplexing. Precoding schemes always have a significant impact on energy efficiency. Architecture, planning, and hardware are also essential to rule out the energy issues. However, energy coherent hardware deployment sometimes risks-high data rates. Researchers are working on hybrid coding (combining analog & digital precoding). The channel estimation is necessary for mmWave hybrid precoding, and that is a difficult task. Millimeter waves are questioned over its sustainability, scattering, and sensitivity. Channel changes over time due to reflection and scattering. With massive antennas installation because of massive MIMO integration with mmWave (at transmitter and receiver side), it is challenging to estimate channel due to significant computational requirements and complexity. Deep learning is considered a feasible solution to such high computational problems.

A deep learning-based approach is proposed by [53] for hybrid precoding to enhance precoding performance and spectral efficiency. A deep neural network-based framework is considered for training purposes and creates mapping relationships among multiple layers to initiate functions. The considered system is with one base station and Uniform Linear Array (ULA) antennas with no prior information on the links. The framework consists of six hidden layers of DNN to do the mapping, and then training is done for statistical information of mmWave. As precoding techniques have a direct effect on energy efficiency, the same deep learning approach for precoding can be used to enhance energy efficiency as well. Another energy-efficient hybrid coding technique for mmwave and massive MIMO is proposed [54] that uses machine learning. The hybrid precoder is generated through their sum rates that is always with the high probability. This scheme proved enhanced energy-efficient and sum-rate hybrid precoding architecture when compared to traditional approaches.

Base stations using mmWave are mostly equipped with large arrays of antennas that help overcome path loss, improve spectral efficiency, and increase capacity. Due to these massive arrays of antennas, energy efficiency becomes a concern. To fully utilize this installation, analog beamforming is used to increase energy efficiency. On the other hand, spectral efficiency effects will increase because of independent radio frequency chains in the case of digital beamforming. However, energy efficiency will be reduced. Both analog beamforming and hybrid beamforming designs are used in mmWave on either radiofrequency or intermediate frequencies [117].

A beamforming scheme using deep learning focused on training and design issues of baseband is proposed in [67]. In conventional mmWave communication, multiplex techniques (OMA, TDMA, OFDMA and CDMA) are used because of fewer users than available RF chains. As the user capacity increases with 5G, the conventional techniques are not enough. Non-orthogonal multiple access (NOMA) is utilized as it works in the power domain. MmWave and NOMA provide good results for high data rates. However, it requires additional energy consumption. MmWave has the properties to provide high bandwidth at a short distance, and MIMO provides coverage on a large scale. A lot of work has been done on this combination to utilize it for power control [118], pre-coding [53], [119], power allocation [51].

C. EDGE NETWORK

1) CRAN

RAN provides connectivity via radio connections between end-users and network. Every RAN innovation is dependent on architectural scarcity and end-user demands. BSS was standardized for GSM, and it provides radio mobility functions and is considered the core of 2G. A traditional BS performs based on two devices: Digital Unit (DU) and Radio Equipment Controller (REC), where DU deals with all frequency functions like modulation/demodulation, frequency amplification, A/D and D/A conversion, and frequency filtration. Baseband function is performed at REC like controlling and managing the base station.

As the end-users increased and GPRS evolved, this called for a network architecture up-gradation; GPRS BSS. GSM was working on a circuit switching pattern whereas, GPRS evolved with packet-switching technology. Both RAN architectures can work parallel, but interface modification was required along with the addition of the Packet Control Unit (PCU). GERAN was carried out to cater GSM and EDGE network growing data needs and is the EDGE's access network. UTRAN was developed for UMTS and consisted of more than one RN, several transceivers, and one controller (RNC). It was modified from previous RAN architecture to focus on enhanced data rates. UTRAN was terrain dependent divided on the bases of the population in urban, rural, and suburban. With the LTE standardization, E-UTRAN was developed, which is different from all previous RAN architectures because of no centralized controller. There are two interfaces used in E-UTRAN that are X2-interface and S1-interface. All kinds of information transfer, mobility functions, load balancing, and interference coordination are dealt with by X2-interface. S1-interface is divided into further user plane and control plane. In UTRAN, the radio processing unit is segregated from signal processing units, and the same continued in E-UTRAN. In the previous 2G RAN architecture, all frequency-related functions like amplification (analog to digital and digital to analog) conversion and control, transport, and baseband functions were performed at the transceiver base station. However, with architectural change and data needs, it was difficult to perform all these functions in one place. This lead to the development of D-RAN. In D-RAN, Remote Radio Unit (RRU) and Baseband Unit (BBU) replaced REC and DU respectively. According to [120], D-RAN is considered as the efficient RAN architecture for 3G and 4G networks.

Due to increased data needs, a new RAN concept has been developed for massive data, which also caters to the interference problem. With the advancement in RAN architecture, user plane is segregated from the control plane, and the SDN switch is used to exchange the user data messages from the RAN controller and the other part from the control-based interface. This segregation makes RAN more versatile to accommodate different NFV and SDN features important in 5G networks like MIMO, service chaining, and network slicing. Cloud/Centralized RAN refers to the concept of consolidating all data at one point; cloud and is considered as a cheaper alternative of OPEX and CAPEX (capital and operative cost) [121]. A generalized CRAN consists of three main components: (i) RRH, (ii) BBU pool, and (iii) a front-haul network to provide connectivity between them. The BBU pool has several software-defined BBUs with centralized processors to optimize radio resources. RRH provides signal coverage and uses the uplinks/downlinks to transfer RF signals from UE to BBU and BBU to UE. To be more precise, it amplifies radio frequencies, does the filtration and conversion from analog to digital, and vice versa. RRH is distributed in a cost-efficient manner, and this is one reason why C-RAN is considered energy efficient.

C-RAN gives the edge to fulfil 5G technologies vision by rendering advanced network architecture and support features such as enhanced performance, high capacity, increased flexibility, energy efficiency, and minimized front-haul network cost [122]. As mentioned, to cater to large data rates and resource management, small cell technology is deployed massively with CRAN to get the heterogeneous CRAN concept [123]. H-CRAN is a new concept to merge both CRAN and heterogeneous to get more benefits for resource allocation. H-CRANs major benefit is exploiting the RRH benefits in providing high data rate capacity, QoS, and energy efficiency. High power nodes (micro BS, macro BS) are the more significant consumer of power than low power nodes (pico BS, Femto BS). All digital processing units reside in the BBU pool. H-CRAN is more energy efficient as it can gather data. The analysis process can take place on run-time, which is possible because its architecture is centralized. More dense deployment of low power nodes incurs acute interference. Spectral efficiency degradation and the energy consumption is achievable by suppressing interference. There are also other conventional techniques for energy consumption minimization, like switching off several BS, which are not in use. However, they are not feasible in every network because of the tight coupling among data services and convergence.

Convergence is the primary factor affecting energy efficiency.

Most of the energy problems have been solved with the C-RAN development (centralization of baseband operations), which aided network in terms of less deployment of BS. However, further research is required on the energy efficiency issue at resource utilization and allocation, power in C-RAN, and computational complexity. An H-CRAN resource allocation scheme is proposed based using machine learning to improve energy efficiency and QoS interference for the H-CRAN downlink [68]. The proposed scheme works by learning information online, and the allocation is performed on the assigned controller. Based on enhanced spectrum partitioning, user traffic is prioritized and differentiated with location to be used as the initial learner to feed the machine. Power is allocated based on a single controller connected to BBU, which also needs network state information to take further actions that guarantee energy efficiency. Another work on resource allocation scheme for downlink H-CRAN using machine learning is proposed [124]. The target was also on energy efficiency, fulfilling the service quality alleviating inter-tier interference. H-CRAN can support multiple technologies and bring efficiency and resource allocation factors with no need to rebuild the transport network. Energy efficiency is an essential factor in the wireless network directly linked with traffic demand and load. Radio resource index is also a factor to affect energy consumption, and bandwidth and energy consumption reciprocity are severe.

As mentioned earlier, the cell sleeping concept is also used to minimize power consumption. Thus, enabling way along with beamforming by using deep neural networks is proved to enhance energy efficiency through simulation work [52] significantly. The problem of joint cell sleeping has been focused on CRAN. A deep neural network-based framework is developed to solve the optimization problem. When the RRH is not selected for transmission, it is sent to sleep mode, as maximum power can be saved by putting more cells to sleep mode. DNN can learn from its inputs and output to give the optimal results used in this proposed work to achieve time, accuracy, and energy efficiency. The significant decrease in power consumption is because of BBU placement in data centers, as the RRH has minimal power consumption. BBU work is dependent on user traffic, data load, and demand. Hence managing BBU operation can result in enhanced energy. For real-time scenarios, more time-efficient algorithms are required to solve the problem. Machine learning can be applied to the NP-hard problem (splitting DUs, diverse requirements nodes, and densely deployed high-power nodes) to minimize the computational time and power consumption.

2) MEC

Mobile Edge Computing (MEC) emerged as an essential technology implementation for 4G and can be quickly adopted for the 5G network. It intelligently merges network conditions, location, and radio information to serve users efficiently [125]. In both 4G MEC and 5G MEC, consumers can decide the place of MEC installations. Because of the same deployment level, the transition between 4G MEC and 5G MEC is somewhat easy because of the same resources utilization, old management techniques utilization, and easy interaction with the control plane. However, integrating MEC with NFV and SDN will enhance its flexibility and services. This flexible nature of MEC will eventually help to accomplish URLLC by achieving the edge cloud milestone [126]. Although new mobile devices are equipped with high-speed processing units, they may not be able to handle complex processing. Also, battery consumption constraint restricts users from using computationally intensive applications. This lead to the development of Mobile Cloud Computing (MCC). In MCC, the end-user gets the advantage of centralized clouds (CC) storage resources and computing. MCC has centralized deployment but high latency, jitter & distance to the user equipment, and ample storage and computational power. On the contrary, MEC is deployed in a distributed manner and has little jitter, latency, and distance to the user equipment and limited storage and computational power.

Among other MEC advantages, computational offloading is one of them. Computational offloading gives the edge of energy consumption, response time, and performance [30]. In [127], three use cases have been discussed for MEC: (i) Consumer-oriented services (ii) Operator and third party services (iii) Network performance and QoE improvement services. Consumer-oriented service use case benefit end-users the most because of the computational offloading. Low latency applications like online gaming and some virtual & augmented reality get more benefits from MEC. In the second use case of operator and third-party services, MEC is utilized for IoT as a gateway to deliver the services. The third use case is used to enhance network performance. MEC can provide real-time information, and this helps to improve QoE and can enable the coordination between backhaul network and radio.

Optimization of offloading selections and resource allocation plays a vital role in enhancing energy efficiency. It has both the cloud computing facility and location & radio information. According to [128] proposed technique, offloading decisions in MEC for energy consumption reduction worked on an accurate channel state information. However, for dynamic channels, accurate channel state information is hard to achieve. For these dynamic systems, Reinforced Learning (RL) can be incorporated. In [70], RL based theme is used to enhance energy efficiency. Specific state, reward, and action have been described to utilize DRL features fully. The proposed framework is used for multi-user equipment computational offloading. Markov Decision Process (MDP) has been used in [69] to improve the service migration process. The focus is to migrate the service based on distance from the source to the UE. MEC is also different from MCC in terms of limited radio resources, storage, and computational resources. Because of these limitations, offloading actions sometimes prove expensive. Thus a proper offloading technique is required. A computational offloading framework is defined because of the different network conditions in [129] to improve offloading expenses. A pre-calculated offloading solution has been employed to take the recurrent offloading decision. Deep Reinforcement Learning is an excellent approach to control complex and high dimension problems for MEC. Moreover, by investigating deep connections for MEC, can decide for resource allocation and computational offloading intelligently.

V. ENERGY HARVESTING

Algorithms and protocols are designed to increase the energy efficiency of wireless networks. Researchers are also working on the means of energy that are amply available in the environment. Renewable energy is one way to provide power to network devices. Energy harvesting is the utilization of ambient energy from external sources. These energy sources can be thermal, solar, wind, kinetic, radiation, and magnetic. The obtained energy can be further stored or directly utilized for wireless devices. Some ways to harvest energy for wireless communication are as follow:

- Natural energy harvesting Natural resources like solar, wind, and water are used to harvest energy. These natural resources are stochastic, resulting in power fluctuations.
- Coupling techniques The two coupling techniques are used to harvest energy: inductive coupling and magnetic coupling. Both coupling techniques are used for the short ranges as they are dependent upon distance and coupling coefficient.
- Wireless Power Transfer (WPT) In WTP, radio frequency signals are used to harvest energy. These electromagnetic radiation is being harvested over the air to utilize the energy which would be otherwise neglected. The above-discussed randomness of energy (from natural resources) is covered in radio frequency energy harvesting. The radio-frequency range between 300GHz to 3kHz is used for harvesting energy [130].

For self-sustained network design, maintaining the energy flow and balancing fluctuations is very important, which can cause damage to the devices and service disruption. The near field energy generation has 80% of success, whereas for far distance RF energy harvesting method is used that also requires additional equipment like antennas and rectifier circuits [131]. Another way is to efficiently utilize the interference signals, which also do not affect the system performance is using interference signals as energy harvesting [132]. The EH is best suitable for portable devices and cannot work on the plugin power strategy and for those who are power-hungry.

Radiofrequency signals are considered a more efficient energy harvesting technique rather than solar and wind energy harvesting resources. In 5G communication, mmWave can engage a large number of antennas arrays due to its smaller wavelength and cell shrinking feature. It is a good candidate for future energy harvesting [133]. The growing 5G technology is leading by the massive deployment of small cells to increase the capacity and energy efficiency of HetNets. Small cell features can be further enhanced by exploiting the energy harvesting concept.

A distributed Q-Learning approach is used in [134], for small cells. The solar energy is taken as a reference as it helps to offload BS at day time. The Markov decision method is used to make decisions for each agent. For the growing needs and challenges, power is also a constraint for the machine to machine communication. The combination of machine learning and energy harvesting techniques, together with cognitive radio, can outperform in energy efficiency aspects. Cognitive machine to machine devices (CM2M) consume a lot of energy, and replacing the battery often is tedious. Thus, researchers started incorporating energy harvesting with CM2M. This integration helps to increase both spectrum as well as energy efficiency. EH-M2M uses cellular users energy, but it can also harvest energy from ambient sources, which ultimately helps to increase battery lifespan for devices. Machine type communication also holds challenges other than power, such as network control, resource allocation, and scheduling. A resource allocation technique proposed in [135] uses spectrum reusing scenarios to enhance energy efficiency in EH-CM2M.

Another resource allocation strategy for EH-CM2M networks is proposed in [136] that uses the deep reinforcement learning approach to enhance energy efficiency. An alternative solution to the M2M energy issue is to shift traffic to the device to device communication. D2D communication also provides feasibility to communicate with each other. In the case of EH-D2D, the energy is harvested from nearby access points [137]. Mostly D2D devices are data-hungry, and that the reason RF harvesting can offset supplementary energy to these devices [138]. Many researchers working on EH-D2D have studied resource allocation techniques in terms of power and resource allocation. The research in energy harvesting for D2D communication is far from mature and requires significant research efforts.

VI. FUTURE DIRECTIONS AND OPEN ISSUES

Radio interface components are the primary reason behind the energy efficiency factor, as 80% of the wireless systems are mainly composed of base station transceivers. Reducing energy consumption is the simplest way to gain green networks [123]. The goal of 5G is to increase spectral efficiency, ubiquitous coverage, and minimize latency. This can be possible by updating and reconstructing network architecture (.i.e virtualization) and advances in radio access network technologies (.i.e, massive MIMO). It will also maximize system performance and increased energy efficiency. Although a lot of research and experimental work has been done on virtualization and softwarization of the network, there is still more research required to overcome issues related to hardware design and deployment, service chaining, energy efficiency, policies, and virtual functions. In this section, we highlight some of the open issues and challenges associated with it.

1) **Combining technologies:** Small cells are great for more dense users, whereas massive MIMO is efficient for less dense environments. Massive MIMO implementation is different depending upon the dense nature of the targeted area. As the 5G network is an amalgamation of diverse technologies, combining these technologies can prove energy-efficient 5G design. Massive MIMO is less energy efficient as compared to small cell networks. However, when the consumed power of the active antennas circuit is less than switched off antennas, it gives higher energy efficiency values. Massive MIMO and mmWave can be combined to achieve less power consumed architecture because of its partially connectivity nature. However, further research should be focused on the dynamic installment of the architecture. The network's energy efficiency factor is highly dependent on the ratio of computational power and transmission power. Due to the dynamic nature of the 5G network, these power values do not remain the same all the time. Hence, the relationship between computation power and transmission power should also be studied to get overall network energy efficiency.
2) **Harvesting real-time benefits:** SDN controller has the benefit to program controller vigorously. As the central controller is detached from the data plane, it can monitor the network in real-time, which can do data configuration and monitoring. Thus integrating machine learning to SDN gives more benefits to real-time networks. MEC is also an excellent option to harvest real-time benefits. Currently, MEC manages mobility for end-users under simplistic scenarios. However, in the future, handling several nodes while effectively controlling the virtual machine migration is a challenge for guaranteed service delivery.
3) **Softwarization & virtualization:** Both SDN and NFV can be incorporated in a network together. They serve each other irrespective of their contrary nature. SDN can provide programmable benefits to NFV in the form of connectivity among virtual network functions. NFV helps SDN by providing virtualization of network function. Moreover, to cope with growing user demands and energy constraints, there is much space to research MEC's incorporation with SDN/NFV. For the future, research can be done on installing content and application both at the consumer side can reduced energy usage and OPEX as computational offloading benefits energy efficiency.
4) **Machine learning and data relationship:** Keeping in view the advantages of machine learning to solve complex problems to improve performance and less complicated implementation shows its feasibility over traditional algorithmic approaches. The bigger advantage that machine learning has is its learning nature from the environment. However, a severe lack of available data sets for research purposes and securing data from networks is difficult. Even after data acquisition, the model needs to be trained. Before training, all data must be aligned, debugged, and cleaned of all the biased values that will also require a lot of processing. Future researchers need to work on the trade-off between efficient machine learning for wireless networks and how models can be simplified. Especially for those areas where energy efficiency is a critical aspect.
5) **Reinforced learning in real-time environment:** Reinforced learning is a good approach to use in real-time environments because of its weight assignment based on learning. Another important benefit of reinforcement learning is that it can work well even with no sample or I/O data. It can learn iterative from its environment, giving rewards and responses. However, in intricate state space, it is not adaptive. This is because of the exponentially large storage space, making it difficult to search for data in the huge database. Further research is required to solve the means of storing statistical data, as inputs in the form of a vector in the traditional approach make it difficult.
6) **Collaboration & Discovery** MEC infrastructure is deployed at the consumer end. Because of this user-side deployment, proper communication is required among different network providers. This highlights the need for a proper protocol for collaboration to access network despite different deployment places. To harvest low latency and energy-efficient MEC benefits, a well-equipped discovery system is needed for future MEC framework to avoid unnecessary computations. Machine learning can also play a significant role in the automatic monitoring and synchronization of resources.
7) **Front-haul dependency** Due to the growing user data requirements, the front-haul bandwidth requirements are also growing. And because of the high cost of front-haul deployment, not only expanded infrastructure and increased OPEX & CAPEX will become a challenge. Also, it will lead to decreased energy efficiency. For these issues, fronthaul requires low latency and vast capacity networks to cater to the current capacity issue.
8) **D2D communication:** In device to device communication, energy efficiency is impacted due to frequent device discovery operations. It is highly dependent on protocols that force devices to listen and exchange discovery messages frequently. More research work is required in the future to mitigate the energy issue from these frequent discovery issues.
9) **Need of energy harvesting:** In the future, dense deployed BS enabled with the computational capabilities will consume a lot of energy. Hence harvesting energy resources should be used to energize MEC servers. This renewable energy resources will also help to sustain the network longer.
10) **Performance:** Applications that rely on extensive hardware resources or demand for low latency are at the risk of more performance degradation. This happens

because of their virtualized nature. As a consequence of this excessive virtualization nature, improving energy efficiency becomes a challenge.

11) **RAN and need for intelligence in algorithms:** Offloading in radio access networks provides the benefits of lower latency, improved QoS & QoE, and automatic network selection. However, for a heterogeneous environment, QoE is still a challenge. In the future, for better energy trade-off and less computational complexity, there is a need for intelligent learning techniques like machine learning, which saves energy by reducing recurrent information traffic.

VII. CONCLUSION

5G is a diverse network that will enable a variety of services with the help of several enabling technologies. The main drivers are virtualization, softwarization, new RANs, and backhaul strategies. All the enablers for 5G will help deliver extremely low latency rates, provide high throughput, and support massive connectivity simultaneously. Furthermore, the need for increased network capacity, geographical coverage, and increasing traffic demands require network densification. All such improvements to support diverse use cases will eventually lead to more energy consumption compared to past generations. This is not sustainable from an environmental and business perspective. The need for the energy-efficient network is adopted worldwide because of both economic and environmental concerns. A lot of research has been done on improving energy efficiency in the 5G networks from the past few years. Due to the autonomous decision-making capabilities and benefit of learning from its environment, there has been growing interest in using machine learning techniques to solve energy efficiency at various 5G network levels.

In this paper, we surveyed the state of the art literature to address the energy efficiency issue in the 5G network and the need for intelligent learning. For this purpose, we proposed a taxonomy where we categorized the 5G network into three main parts; access, edge, and the core. The enabling technologies under the provided taxonomy are discussed by addressing machine learning importance in improving energy efficiency. In conclusion, machine learning holds the ability to mitigate energy efficiency issues and improve performance in future networks and under unpredictable network conditions. If appropriately implemented, machine learning has the potential to optimize the operation of a 5G network while at the same time, improving energy efficiency. However, there are still several open challenges that need to be addressed to build highly energy-efficient networks. For this, we highlighted some of the key challenges that need to be thoroughly investigated and provided future research direction for the same.

REFERENCES

[1] N. Panwar, S. Sharma, and A. K. Singh, "A survey on 5G: The next generation of mobile communication," *Phys. Commun.*, vol. 18, pp. 64–84, Mar. 2016.

[2] A. Fonseca, R. Kazman, and P. Lago, "A manifesto for energy-aware software," *IEEE Softw.*, vol. 36, no. 6, pp. 79–82, Nov. 2019.

[3] P. Lahdekorpi, M. Hronec, P. Jolma, and J. Moilanen, "Energy efficiency of 5G mobile networks with base station sleep modes," in *Proc. IEEE Conf. Standards for Commun. Netw. (CSCN)*, Sep. 2017, pp. 163–168.

[4] D. D. Johnson. (Oct. 2018). *The 5G Dilemma: More Base Stations, More Antennas-Less Energy?* [Online]. Available: https://spectrum.ieee.org/energywise/telecom/wireless/will-increased-energy-consumption-be-the-achilles-heel-of-5g-networks

[5] S. Rajoria, A. Trivedi, and W. W. Godfrey, "A comprehensive survey: Small cell meets massive MIMO," *Phys. Commun.*, vol. 26, pp. 40–49, Feb. 2018.

[6] B. Yi, X. Wang, K. Li, S. K. Das, and M. Huang, "A comprehensive survey of network function virtualization," *Comput. Netw.*, vol. 133, pp. 212–262, Mar. 2018.

[7] E. Björnson and E. G. Larsson, "How energy-efficient can a wireless communication system become?" in *Proc. 52nd Asilomar Conf. Signals, Syst., Comput.*, Oct. 2018, pp. 1252–1256.

[8] M. Zakarya and L. Gillam, "Energy efficient computing, clusters, grids and clouds: A taxonomy and survey," *Sustain. Comput. Informat. Syst.*, vol. 14, pp. 13–33, Jun. 2017.

[9] R. Yadav, "Challenges and evolution of next generations wireless communication," in *Proc. Int. MultiConf. Eng. Comput. Scientists*, vol. 2, 2017, pp. 619–623.

[10] J. Vidal, "Tsunami of data could consume one fifth of global electricity by 2025," Climate Home News, 2017, vol. 11.

[11] C.-L. I, S. Han, and S. Bian, "Energy-efficient 5G for a greener future," *Nature Electron.*, vol. 3, no. 4, pp. 182–184, Apr. 2020.

[12] M. A. Al-Namari, A. M. Mansoor, and M. Y. I. Idris, "A brief survey on 5G wireless mobile network," *Int. J. Adv. Comput. Sci. Appl.*, vol. 8, no. 11, pp. 52–59, 2017.

[13] (Jun. 2020). *Ericsson Mobility Report*. [Online]. Available: https://www.ericsson.com/49da93/assets/local/mobility-report/documents/2020/june2020-ericsson-mobility-report.pdf

[14] K. Demestichas, E. Adamopoulou, and M. Choraś, "5G communications: Energy efficiency," *Mobile Inf. Syst.*, vol. 2017, Apr. 2017, Art. no. 5121302.

[15] (Jun. 2020). *Ericsson Mobility Report*. [Online]. Available: https://www.ericsson.com/en/mobility-report/reports/june-2020

[16] S. Zhang, X. Cai, W. Zhou, and Y. Wang, "Green 5G enabling technologies: An overview," *IET Commun.*, vol. 13, no. 2, pp. 135–143, Jan. 2019.

[17] M. Usama and M. Erol-Kantarci, "A survey on recent trends and open issues in energy efficiency of 5G," *Sensors*, vol. 19, no. 14, p. 3126, Jul. 2019.

[18] S. Buzzi, C.-L. I, T. E. Klein, H. V. Poor, C. Yang, and A. Zappone, "A survey of energy-efficient techniques for 5G networks and challenges ahead," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 4, pp. 697–709, Apr. 2016.

[19] I. B. Sofi and A. Gupta, "A survey on energy efficient 5G green network with a planned multi-tier architecture," *J. Netw. Comput. Appl.*, vol. 118, pp. 1–28, Sep. 2018.

[20] Y. Jarraya, T. Madi, and M. Debbabi, "A survey and a layered taxonomy of software-defined networking," *IEEE Commun. Surveys Tuts.*, vol. 16, no. 4, pp. 1955–1980, 4th Quart., 2014.

[21] S. Bera, S. Misra, and A. V. Vasilakos, "Software-defined networking for Internet of Things: A survey," *IEEE Internet Things J.*, vol. 4, no. 6, pp. 1994–2008, Dec. 2017.

[22] F. Z. Morais, C. André da Costa, A. M. Alberti, C. B. Both, and R. da Rosa Righi, "When SDN meets C-RAN: A survey exploring multi-point coordination, interference, and performance," *J. Netw. Comput. Appl.*, vol. 162, Jul. 2020, Art. no. 102655.

[23] R. Alvizu, G. Maier, N. Kukreja, A. Pattavina, R. Morro, A. Capello, and C. Cavazzoni, "Comprehensive survey on T-SDN: Software-defined networking for transport networks," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 4, pp. 2232–2283, 4th Quart., 2017.

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

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

[26] A. Kassaw, D. Hailemariam, and A. M. Zoubir, "Review of energy efficient resource allocation techniques in massive mimo system," in *Proc. Int. Conf. Inf. Commun. Technol. Converg. (ICTC)*, 2018, pp. 237–242.

[27] N. Sapountzis, T. Spyropoulos, N. Nikaein, and U. Salim, "User association in HetNets: Impact of traffic differentiation and backhaul limitations," *IEEE/ACM Trans. Netw.*, vol. 25, no. 6, pp. 3396–3410, Dec. 2017.

[28] W. Z. Khan, E. Ahmed, S. Hakak, I. Yaqoob, and A. Ahmed, "Edge computing: A survey," *Future Gener. Comput. Syst.*, vol. 97, pp. 219–235, Aug. 2019.

[29] S. Wang, X. Zhang, Y. Zhang, L. Wang, J. Yang, and W. Wang, "A survey on mobile edge networks: Convergence of computing, caching and communications," *IEEE Access*, vol. 5, pp. 6757–6779, 2017.

[30] N. Abbas, Y. Zhang, A. Taherkordi, and T. Skeie, "Mobile edge computing: A survey," *IEEE Internet Things J.*, vol. 5, no. 1, pp. 450–465, Feb. 2018.

[31] A. Bohli and R. Bouallegue, "How to meet increased capacities by future green 5G networks: A survey," *IEEE Access*, vol. 7, pp. 42220–42237, 2019.

[32] A. Ahad, M. Tahir, M. Aman Sheikh, K. I. Ahmed, A. Mughees, and A. Numani, "Technologies trend towards 5G network for smart health-care using IoT: A review," *Sensors*, vol. 20, no. 14, p. 4047, Jul. 2020.

[33] M. G. Kibria, K. Nguyen, G. P. Villardi, K. Ishizu, and F. Kojima, "Next generation new radio small cell enhancement: Architectural options, functionality and performance aspects," *IEEE Wireless Commun.*, vol. 25, no. 4, pp. 120–128, Aug. 2018.

[34] V. Mancuso and S. Alouf, "Reducing costs and pollution in cellular networks," *IEEE Commun. Mag.*, vol. 49, no. 8, pp. 63–71, Aug. 2011.

[35] Huawei. *5G Power Whitepaper*. pp. 4–9. Accessed: Apr. 2020. [Online]. Available: https://carrier.huawei.com/~/media/CNBGV2/download/products/network-energy/5G-Telecom-Energy-Target-Network-White-Paper.pdf

[36] M. Feng, S. Mao, and T. Jiang, "Base station ON-OFF switching in 5G wireless networks: Approaches and challenges," *IEEE Wireless Commun.*, vol. 24, no. 4, pp. 46–54, Aug. 2017.

[37] B. Aebischer and L. M. Hilty, "The energy demand of ICT: A historical perspective and current methodological challenges," in *ICT Innovations for Sustainability* (Advances in Intelligent Systems and Computing), vol. 310, L. Hilty and B. Aebischer, Eds. Cham, Switzerland: Springer, 2015, pp. 71–103, doi: 10.1007/978-3-319-09228-7_4.

[38] J. Malmodin and D. Lundén, "The energy and carbon footprint of the global ict and e&m sectors 2010–2015," *Sustainability*, vol. 10, no. 9, p. 3027, 2018.

[39] A. Andrae and T. Edler, "On global electricity usage of communication technology: Trends to 2030," *Challenges*, vol. 6, no. 1, pp. 117–157, Apr. 2015.

[40] (Mar. 2020). *Sustainability and Corporate Responsibility Report 2019*. [Online]. Available: https://www.ericsson.com/49353d/assets/local/about-ericsson/sustainability-and-corporate-responsibility/documents/2020/2019-sustainability-and-corporate-responsibility-report.pdf

[41] G. Koutitas and P. Demestichas, "A review of energy efficiency in telecommunication networks," *Telfor J.*, vol. 2, no. 1, pp. 2–7, 2010.

[42] E. Hwang, K. J. Kim, J. J. Son, and B. D. Choi, "The power-saving mechanism with periodic traffic indications in the IEEE 802.16e/m," *IEEE Trans. Veh. Technol.*, vol. 59, no. 1, pp. 319–334, Jan. 2010.

[43] H. A. H. Hassan, L. Nuaymi, and A. Pelov, "Renewable energy in cellular networks: A survey," in *Proc. IEEE Online Conf. Green Commun. (OnlineGreenComm)*, Oct. 2013, pp. 1–7.

[44] J. Lorincz, A. Capone, and J. Wu, "Greener, energy-efficient and sustainable networks: State-of-the-art and new trends," *Sensors*, vol. 19, no. 22, pp. 48–64, Nov. 2019, doi: 10.3390/s19224864.

[45] D. Feng, C. Jiang, G. Lim, L. J. Cimini, G. Feng, and G. Ye Li, "A survey of energy-efficient wireless communications," *IEEE Commun. Surveys Tuts.*, vol. 15, no. 1, pp. 167–178, 1st Quart., 2013.

[46] X. Wang, A. V. Vasilakos, M. Chen, Y. Liu, and T. T. Kwon, "A survey of green mobile networks: Opportunities and challenges," *Mobile Netw. Appl.*, vol. 17, no. 1, pp. 4–20, Feb. 2012.

[47] T. Chen, Y. Yang, H. Zhang, H. Kim, and K. Horneman, "Network energy saving technologies for green wireless access networks," *IEEE Wireless Commun.*, vol. 18, no. 5, pp. 30–38, Oct. 2011.

[48] A. Fehske, G. Fettweis, J. Malmodin, and G. Biczok, "The global footprint of mobile communications: The ecological and economic perspective," *IEEE Commun. Mag.*, vol. 49, no. 8, pp. 55–62, Aug. 2011.

[49] N. Zhao, Y.-C. Liang, D. Niyato, Y. Pei, M. Wu, and Y. Jiang, "Deep reinforcement learning for user association and resource allocation in heterogeneous cellular networks," *IEEE Trans. Wireless Commun.*, vol. 18, no. 11, pp. 5141–5152, Nov. 2019.

[50] L. Sanguinetti, A. Zappone, and M. Debbah, "Deep learning power allocation in massive MIMO," in *Proc. 52nd Asilomar Conf. Signals, Syst., Comput.*, Oct. 2018, pp. 1257–1261.

[51] M. Zeng, W. Hao, O. A. Dobre, and H. V. Poor, "Energy-efficient power allocation in uplink mmWave massive MIMO with NOMA," *IEEE Trans. Veh. Technol.*, vol. 68, no. 3, pp. 3000–3004, Mar. 2019.

[52] G. Du, L. Wang, Q. Liao, and H. Hu, "Deep neural network based cell sleeping control and beamforming optimization in cloud-RAN," in *Proc. IEEE 90th Veh. Technol. Conf. (VTC-Fall)*, Sep. 2019, pp. 1–5.

[53] H. Huang, Y. Song, J. Yang, G. Gui, and F. Adachi, "Deep-learning-based millimeter-wave massive MIMO for hybrid precoding," *IEEE Trans. Veh. Technol.*, vol. 68, no. 3, pp. 3027–3032, Mar. 2019.

[54] X. Gao, L. Dai, Y. Sun, S. Han, and I. Chih-Lin, "Machine learning inspired energy-efficient hybrid precoding for mmWave massive MIMO systems," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2017, pp. 1–6.

[55] J. Qiu, Q. Wu, G. Ding, Y. Xu, and S. Feng, "A survey of machine learning for big data processing," *EURASIP J. Adv. Signal Process.*, vol. 2016, no. 1, p. 67, 2016.

[56] B. G. Assefa and O. Ozkasap, "HyMER: A hybrid machine learning framework for energy efficient routing in SDN," vol. 2, Aug. 2019, *arXiv:1909.08074*. [Online]. Available: https://ui.adsabs.harvard.edu/abs/2019arXiv190908074G

[57] B. G. Assefa and O. Ozkasap, "MER-SDN: Machine learning framework for traffic aware energy efficient routing in SDN," in *Proc. IEEE 16th Int. Conf. Dependable, Autonomic Secure Comput., 16th Int. Conf. Pervasive Intell. Comput., 4th Int. Conf. Big Data Intell. Comput. Cyber Sci. Technol. Congr.(DASC/PiCom/DataCom/CyberSciTech)*, Aug. 2018, pp. 974–980.

[58] C. Zhang, X. Wang, F. Li, and M. Huang, "NNIRSS: Neural network-based intelligent routing scheme for SDN," *Neural Comput. Appl.*, vol. 31, no. 10, pp. 6189–6205, 2019.

[59] H.-G. Kim, D.-Y. Lee, S.-Y. Jeong, H. Choi, J.-H. Yoo, and J. W.-K. Hong, "Machine learning-based method for prediction of virtual network function resource demands," in *Proc. IEEE Conf. Netw. Softwarization (NetSoft)*, Jun. 2019, pp. 405–413.

[60] J. Xu, J. Wang, Q. Qi, H. Sun, and B. He, "IARA: An intelligent application-aware VNF for network resource allocation with deep learning," in *Proc. 15th Annu. IEEE Int. Conf. Sens., Commun., Netw. (SECON)*, Jun. 2018, pp. 1–3.

[61] Y. Zhang, J. Yao, and H. Guan, "Intelligent cloud resource management with deep reinforcement learning," *IEEE Cloud Comput.*, vol. 4, no. 6, pp. 60–69, Nov. 2017.

[62] M. Miozzo and P. Dini, "Layered learning radio resource management for energy harvesting small base stations," in *Proc. IEEE 87th Veh. Technol. Conf. (VTC Spring)*, Jun. 2018, pp. 1–6.

[63] N. Zhao, X. He, M. Wu, P. Fan, M. Fan, and C. Tian, "Deep Q-network for user association in heterogeneous cellular networks," in *Complex, Intelligent, and Software Intensive Systems*, L. Barolli, N. Javaid, M. Ikeda, and M. Takizawa, Eds. Cham, Switzerland: Springer, 2018, pp. 398–407, doi: 10.1007/978-3-319-93659-8_35.

[64] H. Ding, F. Zhao, J. Tian, D. Li, and H. Zhang, "A deep reinforcement learning for user association and power control in heterogeneous networks," *Ad Hoc Netw.*, vol. 102, May 2020, Art. no. 102069.

[65] H. Ramazanali, A. Mesodiakaki, A. Vinel, and C. Verikoukis, "Survey of user association in 5G HetNets," in *Proc. 8th IEEE Latin-Amer. Conf. Commun. (LATINCOM)*, Nov. 2016, pp. 1–6.

[66] Q. Kong and B. Wang, "User association for green heterogeneous cellular networks with hybrid energy supplies," in *Proc. 6th Int. Conf. Wireless Commun. Signal Process. (WCSP)*, Oct. 2014, pp. 1–6.

[67] A. Alkhateeb, S. Alex, P. Varkey, Y. Li, Q. Qu, and D. Tujkovic, "Deep learning coordinated beamforming for highly-mobile millimeter wave systems," *IEEE Access*, vol. 6, pp. 37328–37348, 2018.

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

[69] A. Ksentini, T. Taleb, and M. Chen, "A Markov decision process-based service migration procedure for follow me cloud," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2014, pp. 1350–1354.

[70] J. Li, H. Gao, T. Lv, and Y. Lu, "Deep reinforcement learning based computation offloading and resource allocation for MEC," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Apr. 2018, pp. 1–6.

[71] *The Greentouch Project*. Accessed: Mar. 22, 2016. [Online]. Available: http://www.greentouch.org

[72] P. Demestichas, A. Georgakopoulos, K. Tsagkaris, and S. Kotrotsos, "Intelligent 5G networks: Managing 5G WirelessMobile broadband," *IEEE Veh. Technol. Mag.*, vol. 10, no. 3, pp. 41–50, Sep. 2015.

[73] I. Humar, X. Ge, L. Xiang, M. Jo, M. Chen, and J. Zhang, "Rethinking energy efficiency models of cellular networks with embodied energy," *IEEE Netw.*, vol. 25, no. 2, pp. 40–49, Mar. 2011.

[74] Z. S. Andreopoulou, "Green informatics: ICT for green and sustainability," *Agrárinformatika/J. Agricult. Informat.*, vol. 3, no. 2, pp. 1–8, Jan. 2013.

[75] K. Samdanis, P. Rost, A. Maeder, M. Meo, and C. Verikoukis, *Green Communications: Principles, Concepts and Practice*. Hoboken, NJ, USA: Wiley, 2015.

[76] (2020). *Final Evaluation Report from the 5G Infrastructure Association on IMT-2020 Proposal IMT-2020*. [Online]. Available: https://5g-ppp.eu/wp-content/uploads/2020/02/5G-IA-Final-Evaluation-Report-3GPP-1.pdf

[77] D. Aziz, K. Kusume, O. Queseth, H. Tullberg, M. Fallgren, M. Schellmann, M. Uusitalo, and M. Maternia, "Deliverable D8.4 METIS final project report," Version 1, Final Rep. ICT-317669-METIS/D8.4, Apr. 2015. [Online]. Available: https://metis2020.com/wp-content/uploads/deliverables/METIS_D8.4_v1.pdf

[78] (2019). *European Annual Journal 2019*. [Online]. Available: https://bscw.5g-ppp.eu/pub/bscw.cgi/d302069/Euro_5G_PPP_Annual_Journal_2019-web.pdf

[79] M. A. Imran *et al.* (2012). *Energy Aware Radio and Network Technologies INFSO-ICT-247733 EARTH, Version 1*. [Online]. Available: https://cordis.europa.eu/docs/projects/cnect/3/247733/080/deliverables/001-EARTHWP6D62b.pdf

[80] M. Aubree, D. Marquet, S. Le Masson, H. Louahlia, A. Chehade, J. David, and F. Van Goethem, "OPERA-Net 2 project—An environmental global approach for radio access networks-achievements for off-grid systems," in *Proc. IEEE 36th Int. Telecommun. Energy Conf. (INTELEC)*, Oct. 2014, pp. 1–7.

[81] *Enhanced Multicarrier Techniques for Professional Ad-Hoc and Cell-Based Communications(EMPhAtiC)*. Accessed: Jun. 2020. [Online]. Available: http://www.ict-emphatic.eu/images/deliverables/deliverable_d4.1_final.pdf

[82] *Energy Efficiency Testing of Various Telecom Equipment and Green Telecom*. Accessed: Jun. 2020. [Online]. Available: http://tec.gov.in/pdf/Studypaper/final%20LATEST%20ENERGY%20EFFICIENCY22MARCH.pdf

[83] R. Mahapatra, Y. Nijsure, G. Kaddoum, N. Ul Hassan, and C. Yuen, "Energy efficiency tradeoff mechanism towards wireless green communication: A survey," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 1, pp. 686–705, 1st Quart., 2016.

[84] H. Hamdoun, P. Loskot, T. O'Farrell, and J. He, "Survey and applications of standardized energy metrics to mobile networks," *Ann. Telecommun.-Annales des Télécommunications*, vol. 67, nos. 3–4, pp. 113–123, Apr. 2012.

[85] D. Wang and X.-H. Sun, "APC: A novel memory metric and measurement methodology for modern memory systems," *IEEE Trans. Comput.*, vol. 63, no. 7, pp. 1626–1639, Jul. 2014.

[86] C. Belady, A. Rawson, J. Pfleuger, and T. Cader, "Green grid data center power efficiency metrics: PUE and DCIE," Green Grid, Beaverton, OR, USA, Tech. Rep., 2008. Accessed: Aug. 2020. [Online]. Available: https://www.missioncriticalmagazine.com/ext/resources/MC/Home/Files/PDFs/TGG_Data_Center_Power_Efficiency_Metrics_PUE_and_DCiE.pdf

[87] A. Aligrudic and M. Pejanovic-Djurisic, "Energy efficiency metrics for heterogenous wireless cellular networks," in *Proc. Wireless Telecommun. Symp.*, 2014, pp. 1–4.

[88] D. Kreutz, F. M. V. Ramos, P. E. Verissimo, C. E. Rothenberg, S. Azodolmolky, and S. Uhlig, "Software-defined networking: A comprehensive survey," *Proc. IEEE*, vol. 103, no. 1, pp. 14–76, Jan. 2015.

[89] M. Tahir, M. H. Habaebi, M. Dabbagh, A. Mughees, A. Ahad, and K. I. Ahmed, "A review on application of blockchain in 5G and beyond networks: Taxonomy, field-trials, challenges and opportunities," *IEEE Access*, vol. 8, pp. 115876–115904, 2020.

[90] Y.-C. Wang and S.-Y. You, "An efficient route management framework for load balance and overhead reduction in SDN-based data center networks," *IEEE Trans. Netw. Service Manage.*, vol. 15, no. 4, pp. 1422–1434, Dec. 2018.

[91] J. Son, A. V. Dastjerdi, R. N. Calheiros, and R. Buyya, "SLA-aware and energy-efficient dynamic overbooking in sdn-based cloud data centers," *IEEE Trans. Sustain. Comput.*, vol. 2, no. 2, pp. 76–89, May 2017.

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

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

[94] K. M. Modieginyane, B. B. Letswamotse, R. Malekian, and A. M. Abu-Mahfouz, "Software defined wireless sensor networks application opportunities for efficient network management: A survey," *Comput. Electr. Eng.*, vol. 66, pp. 274–287, Feb. 2018.

[95] C. Bouras, A. Kollia, and A. Papazois, "SDN & NFV in 5G: Advancements and challenges," in *Proc. 20th Conf. Innov. Clouds, Internet Netw. (ICIN)*, Mar. 2017, pp. 107–111.

[96] *Network Functions Virtualisation (NFV)*. Accessed: Jul. 2020. [Online]. Available: https://www.etsi.org/technologies/nfv

[97] A. N. Al-Quzweeni, A. Q. Lawey, T. E. Elgorashi, and J. M. Elmirghani, "Optimized energy aware 5G network function virtualization," *IEEE Access*, vol. 7, pp. 44939–44958, 2019.

[98] F. Z. Yousaf, M. Bredel, S. Schaller, and F. Schneider, "NFV and SDN—key technology enablers for 5G networks," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 11, pp. 2468–2478, 2017.

[99] (2012). *OpenStack Tacker*. [Online]. Available: https://wiki.openstack.org/wiki/Tacker

[100] (2016). *OpenBaton*. [Online]. Available: https://openbaton.github.io/

[101] A. Francescon, G. Baggio, R. Fedrizzi, E. Orsini, and R. Riggio, "X-MANO: An open-source platform for cross–domain management and orchestration," in *Proc. IEEE Conf. Netw. Softwarization (NetSoft)*, Jul. 2017, pp. 1–6.

[102] D. Bhamare, R. Jain, M. Samaka, and A. Erbad, "A survey on service function chaining," *J. Netw. Comput. Appl.*, vol. 75, pp. 138–155, Nov. 2016.

[103] B. M. Hochwald, T. L. Marzetta, and V. Tarokh, "Multiple-antenna channel hardening and its implications for rate feedback and scheduling," *IEEE Trans. Inf. Theory*, vol. 50, no. 9, pp. 1893–1909, Sep. 2004.

[104] D. Verenzuela, E. Björnson, and L. Sanguinetti, "Spectral and energy efficiency of superimposed pilots in uplink massive MIMO," *IEEE Trans. Wireless Commun.*, vol. 17, no. 11, pp. 7099–7115, Nov. 2018.

[105] K. Upadhya, S. A. Vorobyov, and M. Vehkapera, "Superimposed pilots are superior for mitigating pilot contamination in massive MIMO," *IEEE Trans. Signal Process.*, vol. 65, no. 11, pp. 2917–2932, Jun. 2017.

[106] H. Q. Ngo, L.-N. Tran, T. Q. Duong, M. Matthaiou, and E. G. Larsson, "On the total energy efficiency of cell-free massive MIMO," *IEEE Trans. Green Commun. Netw.*, vol. 2, no. 1, pp. 25–39, Mar. 2018.

[107] (2018). *Going Massive With MIMO*. [Online]. Available: https://www.ericsson.com/en/news/2018/1/massive-mimo-highlights

[108] K. N. R. S. V. Prasad, E. Hossain, and V. K. Bhargava, "Machine learning methods for RSS-based user positioning in distributed massive MIMO," *IEEE Trans. Wireless Commun.*, vol. 17, no. 12, pp. 8402–8417, Dec. 2018.

[109] T. Maksymyuk, J. Gazda, O. Yaremko, and D. Nevinskiy, "Deep learning based massive MIMO beamforming for 5G mobile network," in *Proc. IEEE 4th Int. Symp. Wireless Syst. Int. Conf. Intell. Data Acquisition Adv. Comput. Syst. (IDAACS-SWS)*, Sep. 2018, pp. 241–244.

[110] Y. Hao, Q. Ni, H. Li, and S. Hou, "On the energy and spectral efficiency tradeoff in massive MIMO-enabled HetNets with capacity-constrained backhaul links," *IEEE Trans. Commun.*, vol. 65, no. 11, pp. 4720–4733, Nov. 2017.

[111] D. Liu, L. Wang, Y. Chen, M. Elkashlan, K.-K. Wong, R. Schober, and L. Hanzo, "User association in 5G networks: A survey and an outlook," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 2, pp. 1018–1044, 2nd Quart., 2016.

[112] R. Amiri, H. Mehrpouyan, L. Fridman, R. K. Mallik, A. Nallanathan, and D. Matolak, "A machine learning approach for power allocation in HetNets considering QoS," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2018, pp. 1–7.

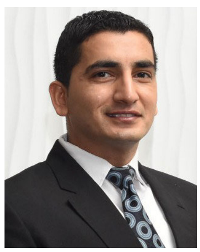
[113] D. Xu, X. Chen, C. Wu, S. Zhang, S. Xu, and S. Cao, "Energy-efficient subchannel and power allocation for HetNets based on convolutional neural network," in *Proc. IEEE 89th Veh. Technol. Conf. (VTC-Spring)*, Apr. 2019, pp. 1–5.

[114] B. Fletcher. (Mar. 2020). *Register for CBRS Week on May 18-21 Join Us for This Free Digital Event Series Exploring The Growing CBRS Ecosystem 5G FCC Wraps up Third Millimeter Wave 5G Spectrum Auction*. [Online]. Available: https://www.fiercewireless.com/5g/fcc-wraps-up-third-millimeter-wave-5g-spectrum-auction

[115] G. N. Kamga and S. Aïssa, "Wireless power transfer in mmWave massive MIMO systems with/without rain attenuation," *IEEE Trans. Commun.*, vol. 67, no. 1, pp. 176–189, Jan. 2019.
[116] O. E. Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. W. Heath, "Spatially sparse precoding in millimeter wave MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 3, pp. 1499–1513, Mar. 2014.
[117] S. Dutta, C. N. Barati, D. Ramirez, A. Dhananjay, J. F. Buckwalter, and S. Rangan, "A case for digital beamforming at mmWave," *IEEE Trans. Wireless Commun.*, vol. 19, no. 2, pp. 756–770, Feb. 2020.
[118] M. Alonzo, S. Buzzi, A. Zappone, and C. D'Elia, "Energy-efficient power control in cell-free and user-centric massive MIMO at millimeter wave," *IEEE Trans. Green Commun. Netw.*, vol. 3, no. 3, pp. 651–663, Sep. 2019.
[119] X. Li and A. Alkhateeb, "Deep learning for direct hybrid precoding in millimeter wave massive MIMO systems," 2019, *arXiv:1905.13212*. [Online]. Available: http://arxiv.org/abs/1905.13212
[120] I. A. Alimi, A. L. Teixeira, and P. P. Monteiro, "Toward an efficient C-RAN optical fronthaul for the future networks: A tutorial on technologies, requirements, challenges, and solutions," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 1, pp. 708–769, 1st Quart., 2018.
[121] I. Neokosmidis, T. Rokkas, P. Paglierani, C. Meani, K. M. Nasr, K. Moessner, and P. S. Khodashenas, "Assessment of capex and opex for media services in cloud enabled 5G networks," in *Proc. CTTE-FITCE Smart Cities Inf. Commun. Technol. (CTTE-FITCE)*, Sep. 2019, pp. 1–6.
[122] D. Pliatsios, P. Sarigiannidis, S. Goudos, and G. K. Karagiannidis, "Realizing 5G vision through cloud RAN: Technologies, challenges, and trends," *EURASIP J. Wireless Commun. Netw.*, vol. 2018, no. 1, p. 136, Dec. 2018.
[123] M. Peng, Y. Li, J. Jiang, J. Li, and C. Wang, "Heterogeneous cloud radio access networks: A new perspective for enhancing spectral and energy efficiencies," *IEEE Wireless Commun.*, vol. 21, no. 6, pp. 126–135, Dec. 2014.
[124] A. Alabbasi, X. Wang, and C. Cavdar, "Optimal processing allocation to minimize energy and bandwidth consumption in hybrid CRAN," *IEEE Trans. Green Commun. Netw.*, vol. 2, no. 2, pp. 545–555, Jun. 2018.
[125] E. Schiller, N. Nikaein, E. Kalogeiton, M. Gasparyan, and T. Braun, "CDS-MEC: NFV/SDN-based application management for MEC in 5G systems," *Comput. Netw.*, vol. 135, pp. 96–107, Apr. 2018.
[126] S. Kekki, W. Featherstone, Y. Fang, P. Kuure, A. Li, A. Ranjan, D. Purkayastha, F. Jiangping, D. Frydman, and G. Verin, "MEC in 5G networks," ETSI, Sophia Antipolis, France, White Paper, 2018, vol. 28, pp. 1–28.
[127] *Mobile Edge Computing (MEC); Technical Requirements*, document ETSI GS MEC 002, 2016.
[128] S. Barbarossa, S. Sardellitti, and P. Di Lorenzo, "Communicating while computing: Distributed mobile cloud computing over 5G heterogeneous networks," *IEEE Signal Process. Mag.*, vol. 31, no. 6, pp. 45–55, Nov. 2014.
[129] S. Yu, X. Wang, and R. Langar, "Computation offloading for mobile edge computing: A deep learning approach," in *Proc. IEEE 28th Annu. Int. Symp. Pers., Indoor, Mobile Radio Commun. (PIMRC)*, Oct. 2017, pp. 1–6.
[130] X. Lu, P. Wang, D. Niyato, D. I. Kim, and Z. Han, "Wireless networks with RF energy harvesting: A contemporary survey," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 2, pp. 757–789, 2nd Quart., 2015.
[131] S. Kim, R. Vyas, J. Bito, K. Niotaki, A. Collado, A. Georgiadis, and M. M. Tentzeris, "Ambient RF energy-harvesting technologies for self-sustainable standalone wireless sensor platforms," *Proc. IEEE*, vol. 102, no. 11, pp. 1649–1666, Nov. 2014.
[132] M. Imran, L. U. Khan, I. Yaqoob, E. Ahmed, M. A. Qureshi, and A. Ahmed, "Energy harvesting in 5G networks: Taxonomy, requirements, challenges, and future directions," 2019, *arXiv:1910.00785*. [Online]. Available: http://arxiv.org/abs/1910.00785
[133] T. X. Tran, W. Wang, S. Luo, and K. C. Teh, "Nonlinear energy harvesting for millimeter wave networks with large-scale antennas," *IEEE Trans. Veh. Technol.*, vol. 67, no. 10, pp. 9488–9498, Oct. 2018.
[134] M. Miozzo, L. Giupponi, M. Rossi, and P. Dini, "Distributed Q-learning for energy harvesting heterogeneous networks," in *Proc. IEEE Int. Conf. Commun. Workshop (ICCW)*, Jun. 2015, pp. 2006–2011.
[135] Z. Zhou, C. Zhang, J. Wang, B. Gu, S. Mumtaz, J. Rodriguez, and X. Zhao, "Energy-efficient resource allocation for energy harvesting-based cognitive machine-to-machine communications," *IEEE Trans. Cognit. Commun. Netw.*, vol. 5, no. 3, pp. 595–607, Sep. 2019.
[136] Y.-H. Xu, Y.-B. Tian, P. K. Searyoh, G. Yu, and Y.-T. Yong, "Deep reinforcement learning-based resource allocation strategy for energy harvesting-powered cognitive machine-to-machine networks," *Comput. Commun.*, vol. 160, pp. 706–717, Jul. 2020.
[137] L. Jiang, H. Tian, Z. Xing, K. Wang, K. Zhang, S. Maharjan, S. Gjessing, and Y. Zhang, "Social-aware energy harvesting device-to-device communications in 5G networks," *IEEE Wireless Commun.*, vol. 23, no. 4, pp. 20–27, Aug. 2016.
[138] R. Atat, L. Liu, N. Mastronarde, and Y. Yi, "Energy harvesting-based D2D-assisted machine-type communications," *IEEE Trans. Commun.*, vol. 65, no. 3, pp. 1289–1302, Mar. 2017.

AMNA MUGHEES received the M.S. degree in electrical engineering from the Department of Electrical and Computer Engineering, COMSATS University, Pakistan, in 2014. She is currently pursuing the Ph.D. degree with the Department of Computing and Information Systems, Sunway University, Malaysia. Her research interests include 5G, wireless networks, handover, the Internet of Things, and artificial intelligence.

MOHAMMAD TAHIR (Senior Member, IEEE) received the M.Sc. and Ph.D. degrees in electrical and computer engineering from the Department of Electrical and Computer Engineering, International Islamic University Malaysia, in 2011 and 2016, respectively. Prior to joining academics, he has worked with the Research and Development Division of industry for seven years, on several projects related to the Internet of Things and cognitive radio. His research interests include 5G, the Internet of Things, game theory for wireless networks, wireless security, blockchain, and autonomic computing.



MUHAMMAD AMAN SHEIKH is currently a Lecturer with the School of Science and Technology, Sunway University. Previously, he was a Senior Scientist in the Research and Development sector, and was involved in enormous projects for PETRONAS. As an active researcher, he has published extensively and established global collaborations by working with international collaborators and securing international grants. He is a Fellow of the Institute of Electrical and Electronic Engineers, Malaysia (MIEEE) and a Registered Engineer with the Board of Engineers, since 2010. He was awarded the Graduate Assistant Scholarship during his M.Sc. and Ph.D. degrees in electrical and electronics engineering in 2013 and 2018, respectively. He also received prestigious awards during his Ph.D. candidature, including the Best Graduate Assistant Award and the High Achiever Award.

ABDUL AHAD received the M.Sc. degree in computer science from the University of Swat, Pakistan, in 2014, and the M.S. degree in computer science from the Virtual University of Pakistan, in 2017. He is currently pursuing the Ph.D. degree in computer science with Sunway University, Malaysia. His research interests include 5G, the Internet of Things, wireless networks, wireless body area networks (WBANs), and artificial intelligence.