Paper

Energy Efficiency in 5G Communications – Conventional to Machine Learning Approaches

Muhammad Khalil Shahid, Filmon Debretsion, Aman Eyob, Irfan Ahmed, and Tarig Faisal

Higher Colleges of Technology, Abu Dhabi, United Arab Emirates

https://doi.org/10.26636/jtit.2020.146820

Abstract—Demand for wireless and mobile data is increasing along with development of virtual reality (VR), augmented reality (AR), mixed reality (MR), and extended reality (ER) applications. In order to handle ultra-high data exchange rates while offering low latency levels, fifth generation (5G) networks have been proposed. Energy efficiency is one of the key objectives of 5G networks. The notion is defined as the ratio of throughput and total power consumption, and is measured using the number of transmission bits per Joule. In this paper, we review state-of-the-art techniques ensuring good energy efficiency in 5G wireless networks. We cover the base-station on/off technique, simultaneous wireless information and power transfer, small cells, coexistence of long term evolution (LTE) and 5G, signal processing algorithms, and the latest machine learning techniques. Finally, a comparison of a few recent research papers focusing on energy-efficient hybrid beamforming designs in massive multiple-input multiple-output (MIMO) systems is presented. Results show that machine learning-based designs may replace best performing conventional techniques thanks to a reduced complexity machine learning encoder.

***Keywords*—5G, energy efficiency, wireless networks.**

1. Introduction

Conventional fuels used for power generation, heating and transport have contributed to an 80% increase in greenhouse gas emissions compared to 1970. According to projections pertaining to 2040, global energy demand is expected to increase even further, by 30%, with the pace of growth being even faster in developing countries [1]. 5G networks would inevitably be responsible for an increase in the amount of energy used by consumers, therefore contributing to climate change. As the amount of space within the wave spectrum in which consumer devices may operate is increased by the use of millimeter waves, energy usage grows as well, leading to faster global warming. 3GPP standards, including those related to 5G networks, aim to increase capacity and coverage of the system, with energy efficiency gains considered at architectural and functional level [2]. Ensuring that hardware is capable of working within extended operating condition ranges (temperature and humidity levels prevailing in rooms in which equipment is located) may lead to a decrease in the amounts of power consumed by air conditioning systems. Small cells used to provide 5G connectivity are claimed to be energy efficient and powered in a sustainable way. However, maintenance- and production-related issues may cause considerable cost implications [1]. Deployment of 5G systems is also expected to improve energy efficiency (EE) of the entire industry as a whole, as the cost of energy per bit of data transferred is, in 5G, equal to one tenth of the level experienced in 4G [2]. However, base stations still remain energy-hungry locations of the network, due to the foreseeable increase in traffic that is expected to grow by several thousand percent.

A few papers exist that focus on analyzing EE of 5G networks. Report [3] surveys various optimization techniques, the game theory and machine learning approaches that have been proposed for enhancing power allocation to downlink and uplink channels. Other energy-saving approaches are described therein as well. In paper [1], some of the significant examples discussed include deployment of newer radio resource control (RRC) for context signaling and for reducing the number of redundant state changes. Utilization of advanced clustering and caching techniques on the radio access network (RAN) side has been highly valued for improving latency requested by a group of users and for eliminating the factor of clogging the network by a huge number of requests for the same data. Commercial resource sharing between different operators offers encouraging results in terms of reduced deployment costs and good data rates, while ensuring minimum interference. In a paper [4] a detailed discussion of the various advantages and disadvantages of green and energy efficiency techniques is presented, contributing to understanding the ways in which green radio architecture may be used in 5G and future mobile networks, and presenting the challenges that will be encountered in the process. In this paper, three most promising green solutions are analyzed. Extreme mobile broadband (xMBB) is a service characterized by high data rates, low latency communication (LLC), and extreme coverage. Its spectrum resources include lower bands, and new higher bands with large contiguous bandwidth, (license + LSA + LAA). Its target values are defined by the peak data rate of up to 20 Gbps for downlink and 10 Gbps for uplink. Massive machine-type communications (mMTC) is another type of service that offers the following features: dense mobile networks, wide-area coverage and deep penetration. It relies on lower band frequencies and its spectrum resources include licensed shared access (LSA) and licensed-assisted access (LAA). The target connectivity density value equals 1 million devices per square kilometer. The third service offers ultra-reliable and low-latency communication links, such as V2X. Its spectrum resources are based on lower bands, exclusive licenses, and its target user latency quals 0.5 ms. A typical 5G network is shown in Fig. 1. Mathematically, energy efficiency of the base station side is defined as:

$$EE = \frac{R}{\eta P_t + P_c} \quad [\text{bits/Joule}] \,, \tag{1}$$

where, R is the average overall data rate in bits per second, η is the reciprocal of the transmit power (amplifier efficiency), P_t is the transmission power, and P_c is total power dissipated in the transmitter circuit.

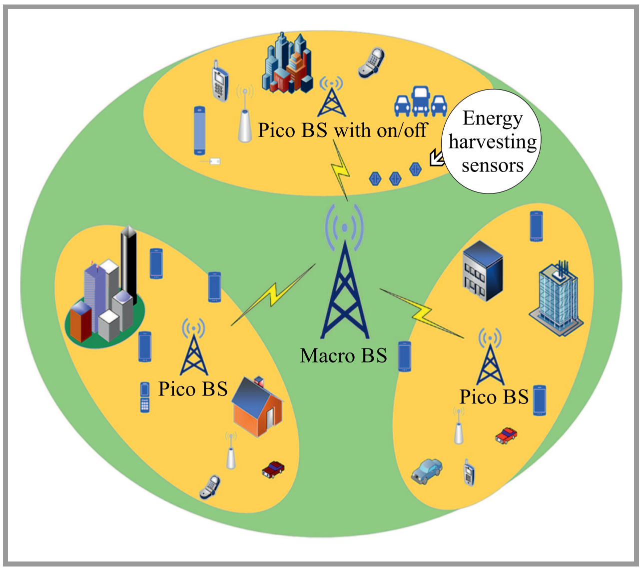

Fig. 1. Model of a multiuser massive MIMO downlink system.

The rest of the paper is organized as follows: in Section 2, on/off techniques are presented and energy harvesting is described. Section 3 presents heterogeneous networks and energy efficiency-related considerations. Energy-efficient physical layer hardware designs are reviewed in Section 4, and machine learning techniques ensuring EE are presented in Section 5.

2. Energy Efficiency Using On/Off Techniques and Energy Harvesting

Authors in [5] propose three independent energy efficiency optimization solutions to minimize energy consumption by either forcing idle base stations to go to sleep, or by dynamically adjusting the signal range of base stations through software-defined networking. This means that the stream table of the base stations is reconfigured to modify the connections between users and base stations in order to free as high a number of base stations that is feasible under specific circumstances. Secondly, the maximum transmission time required for given content to be downloaded from the server is minimized in order to let the base station to go to rest. Finally, at times a stronger signal may be used by base stations in situations in which weaker signal might be sufficient to cover the needs of all customers. Therefore, the power level of such a station may be adjusted to save energy.

In [6], the authors proposed a strategy for dense femtocell deployment based on sleep mode and hybrid access policies. In most studies, a femtocell base station awakens from rest when its clients are active. However, in this work, the femtocells with dynamic clients may continue to sleep if the reallocation of their clients is feasible based on hybrid access to their neighbors, with throughput enhancements ensured simultaneously. Simulations indicate that ab increase in femtocell density boosts the number of clusters formed and femtocells remaining in the sleep mode, hence improving energy efficiency.

Paper [7] describes the efficiency of harvesting energy through radio frequency (RF). Two mathematical simulations are conducted as part of the study, verifying the energy efficiency to training interval ratio and checking how EE is affected by the change in block size and in the number of users. Simulations show that EE is low for large Tt values due to the amount of energy dissipated, which means that the factor that is of most significance for EE deteriorates. Larger amounts of energy may be harvested with more receivers (users). This is because one of the users is decoding information while others harvest energy.

In paper [8], the authors propose an iterative hybrid analog-digital beamforming scheme for simultaneous wireless information and power transfer (SWIPT) for MIMO systems with limited RF chains at the base station. Compared to fully-digital SWIPT, the proposed scheme is reported to be a better solution for energy harvesting with significant gains in total power consumption.

3. Energy Efficient Heterogeneous Networks

A potential solution to improve energy efficiency of 5G heterogeneous networks consists in offloading traffic from macro cells to small cells. Paper [9] presents a scheme conditionally offloading traffic from macro cells to picocells and femtocells under the condition that the cell system load of all cells is maintained below a certain threshold. The proposed scheme is developed using an online reinforcement learning methodology capable of surmising that other macro cells are pursuing offloading strategies, which reduces the base station information exchange overhead. The study found that cell load is a key factor that impacts interference, small cell traffic congestion, and energy efficiency. As network load or cell load increase, energy efficiency, interference, and quality of service (QoS) decrease.

In [10], an energy efficiency comparison has been performed between single-tier, two-tier, and three-tier networks by using stochastic geometry tools. Base stations of all tiers are positioned in accordance with the Poisson point process, and they are not quite the same as one another in terms of their size, density, and transmitting power. Simulation results show that energy efficiency of a heterogeneous network is enhanced by increasing the number of pico and femtocells. However, the growing number of pico and femtocell base stations increases co-channel interference, since pico and femtocells are located inside macro cells.

Multi-connectivity is a feature of 5G technology that allows users to connect and consume resources from multiple base stations concurrently, potentially using different radio interfaces, i.e. Evolved-LTE and new radio. Authors in [11] propose new algorithms for multi-connectivity and compare them with single connectivity scenarios to examine how multi-connectivity is capable of improving reliability and the system's overall energy efficiency. The paper shows that at low speeds, multi-connectivity offers little improvement in terms of radio link failures. However, at high speeds, a reduction of up to 50% is observed. The newly proposed secondary cell selection algorithms presented in [11] improved energy efficiency of the network by up to 20% at low speeds.

The proposal [12] is to optimize power using spectrum sharing in the next generation networks (NGNs) in order to achieve high spectrum and energy efficiency for both the primary and the secondary system, without the involvement of a secondary transmitter. When comparing the performance of the proposed model with the opportunistic spectrum sharing model and other popular resource allocation algorithms, the efficiency of the proposed scheme turns out to be superior. Spectrum sharing is gaining popularity as it allows to expand the spectrum and boost power efficiency of the network. Spectrum sharing allows for cooperative or simultaneous use of limited radio frequency resources by a number of independent users within a particular geographical area. Spectrum sharing may help effectively use white spaces or underutilized portions of the spectrum. Also, there are various power allocation strategies for optimum resource block allocation in the spectrum sharing process. Optimum resource block allocation strategies aim to optimize the spectrum and energy efficiency of the system, simultaneously increasing the quality of service for both primary and secondary receivers.

4. Energy Efficient Physical Layer Hardware Design

Paper [13] aims to reduce both hardware complexity and power consumption in hybrid analog-digital beamforming systems. Reduced-dimension training sequence designs and transmit precoder designs are considered jointly. Paper [19] discusses the minimum number of RF chains required and trade-offs with the bandwidth of the transmitted signal needed in order to realize any given fully digital precoder relying on hybrid analog/digital precoding (HADP) in wide-band mm-Wave systems. The authors of [10] compare different hybrid beamforming architectures and optimize the number of antennas in order to maximize EE. The combination of phase shifters and switches has been shown to be superior to the conventional phase shifter-only architectures in terms of spectral and energy efficiencies. In the paper [14], the authors propose wireless power transfer (WPT) based algorithms for the hybrid beamforming design to achieve optimized performance of the fully digital beamforming architecture. In HB, the number of RF chains is in general significantly lower than that of transmit antennas. It is shown that for a general point-to-point MIMO WPT system, optimized performance may be achieved as long as the number of RF chains at the energy transmitter is not lower than twice the number of sub-bands used or twice the number of channel paths. For MISO WPT, the required number of RF chains may be reduced even further to equal the number of channel paths only.

Paper [15] compares spectral and energy efficiency of hybrid beamforming and digital beamforming systems, taking into account the effects of channel estimation, transmitter impairments, and multiple simultaneous users for a wideband multipath model. Considering the model of the quantization error at each antenna, better EE and achievable rates are obtained with 5-bit ADC resolution at high SNR.

The digital beamforming model with low resolution ADCs outperforms hybrid beamforming in terms of spectral efficiency and energy efficiency for single-user and multi-user scenarios with multipath propagation [15], [16]. Low resolution ADC DBF is resistant to small automatic gain control (AGC) imperfections. Energy efficiency and spectrum efficiency (SE) of wireless communication systems may be significantly enhanced by large-scale antenna systems. These systems are deployed in hybrid digital and analog BF structures. The analysis of this kind of structures presented in [17] provides an insight into optimal analog and digital BF designs, EE-SE relationship at the green point, and the impact of the number of transceivers on the green EE point, i.e. the point with the highest EE along the EE-SE curve. In paper [18], the authors compare the energy efficiency maximization problem for three different types of precoding scenarios, namely zero-forcing (ZF), general beamforming and conjugate-beamforming, in a multiuser downlink distributed antenna system with a hybrid energy supply. The proposed algorithm, based on joint power allocation and energy cooperation, allows to maximize EE and shows that ZF achieves better EE with low noise variance, while conjugate beamforming performs better in high-noise scenarios. Paper [8] proposes different hybrid beamforming schemes in which phase-only weights are applied for the analog precoder and combiner in a mm-Wave MU-MIMO transmission. The proposed design is applied in fully-connected structures and is then extended to their sub-connected counterparts. In [19], the authors propose an energy efficient optimization model for multi-user massive MIMO, based on the joint optimization problem involving computation and communication power. An upper bound of energy efficiency is provided for multi-user massive MIMO systems with partially-connected structures of RF transmission systems. The proposed design aims to find an EE suboptimal solution capable of jointly improving computation and communication power in massive MIMO systems. Paper [20] addresses the implications of highly directional communication for the design of an efficient medium access control (MAC) layer. It discusses the key MAC layers options available and design physical control channel aspects that may improve performance, spectral efficiency and energy efficiency. An energy-efficient hybrid beamforming design for a partially-connected structure is investigated in [21]. The authors design a two-layer optimization method to solve the non-convex problem by exploiting interference alignment and fractional programming. First, an analog beamformer is obtained using the alternating-direction optimization method, and then the minimum mean square error (MMSE)-based digital precoder and combiner are formed.

5. Energy Efficiency Using Machine Learning Techniques

In recent research, machine learning techniques are widely explored in various sections of the 5G networks to enhance their performance in terms of spectral and energy efficiency. In [22], the authors propose a centralized resource allocation scheme that exploits online learning techniques, guaranteeing mitigation of interference and maximization of energy efficiency, while avoiding dropping QoS requirements for all users. The priority of users in resource block (RB) allocation and compact state representation-based learning methodology was considered to enhance the learning process in the design with a model-free reinforcement learning. The outcome of the simulation clearly shows that the proposed resource allocation scheme is capable of mitigating interference, considerably increasing energy and spectral efficiency, as well as sustaining the users' QoS requirements. Centralized joint RBs and the power allocation scheme rely on a single controller integrated into the centralized baseband unit. This controller gathers the network state information through its interface with the macro BSs and uses this knowledge to select the most appropriate actions that ensure energy efficiency enhancements, while simultaneously maintaining QoS requirements from different tier users. The proposed online learning model ensures compact state representation in order to reduce the size of the state space, augments the algorithm's convergence and manages the curse of dimensionality. The conventional base station is separated into two parts within a heterogeneous cloud radio access network structure. This allows to centralize the signal processing part as the baseband unit, and the signal transceiver part as the radio remote head: the centralized baseband unit and the radio remote head unit are associated through fronthaul links. Paper [23] proposes a joint transmission mode selection and power control algorithm with reinforcement learning to ensure energy optimization via fifth-generation communication for the heterogeneous cloud radio access network architecture in a vehicular social network. It considers cellular and D2D communication coexisting in a single cell environment. Two Q-learning algorithms have been designed that enable to take optimal communication mode selection and transmission power control decisions by altering the target SINR. The two Q-learning models are centralized and distributed Q-learning. The agent in the baseband unit uses centralized Q-learning to maximize the system's energy efficiency while guaranteeing compliance with QoS constraints. The vehicles use distributed Q-learning to maximize their achievable data rate.

In [24], a sub-connected switch-based hybrid precoding architecture with lens array for beamspace MIMO systems is proposed to overcome the issue of high energy consumption. The hybrid precoding scheme aiming to maximize the achievable sum rate is based on the cross-entropy (CE) optimization approach developed in machine learning [24]. But this paper does not focus on inter-user interference in the beam domain.

In [25], a machine learning approach was applied to a joint hybrid beamforming and radio resource management for a rank constrained mm-Wave MU-MIMO downlink. Formulation of the optimization problem aimed to maximize the total throughput constraint while taking into account the total transmit power and the number of RF chains. The energy efficiency aspect has not been investigated in this paper. Three different machine learning methods, namely random forest, deep learning, and ridge regression, were applied in [26] to forecast the resulting uplink transmission power based on the passive network quality indicators and application-level information available. The random forest model was reported as offering the best performance. However, this framework cannot be utilized to optimize energy efficiency without the channel feedback overhead. Machine type communications also need a low latency 5G technology. Fast establishment of uplink transmission for machine type communications (MTCs) is one of the main challenges affecting future wireless systems. The optimal selection of machine type devices (MTDs) to be used in the establishment of a fast uplink can be achieved by applying machine learning-based techniques, namely multi-armed bandit theory and deep reinforcement learning [27]. The deep reinforcement learning and multi-armed bandit theory-based uplink resource allocation methodology presented only works for dominant uplink traffic networks, such as sensor networks or MTDs. It cannot be applied to mobile users who require mostly downlink data transmissions. The authors of [28] proposed a genetic algorithm (GA) approach to resource block allocation in a multi-cell network. This approach allows to minimize inter-cell interference by maximizing channel capacity in the resource allocation process. The proposed genetic algorithm-based resource allocation methods for cell-edge users uses the multipath Rayleigh channel. However, 5G communication is supposed to rely on the mm-Wave channel, with dominant line-of-sight (LOS) transmission paths. The resource allocation problem of small LTE-U base stations (SBSs) was modeled based on the deep learning approach in [29]. Dynamic channel selection, carrier aggregation, and fractional spectrum access operations are performed proactively, while guaranteeing fairness with existing Wi-Fi networks and other LTE-U operators. Again, this work does not consider mm-Wave channels and the proposed solution is intended for the multipath channel model.

The author of [30] proposed an actor-critic reinforcement learning (RL) approach based on the stochastic policy gradient to improve the system's throughput and to boost D2D throughput. Compared to the value-based scheme, such as Q-learning, the policy-based method was reported to be better, since it maximized the expected throughput by searching in the policy space. Due to the near-optimal and direction beam transmission in the mm-Wave channel, D2D communication will exert lower impact on the existing cellular network. Reinforcement learning has also been proposed for the resource allocation stage in mobile edge computing systems [31]. It offers a near-optimal solution for the joint task of offloading and resource allocation, simultaneously minimizing energy consumption and delays. This work does not consider transmission power consumption. In [32], the resource allocation problem in virtual reality wireless networks is considered for both uplink and downlink scenarios. Formulation of the has led to the establishment of a non-cooperative game and a distributed algorithm based on the machine learning framework of echo state networks (ESNs) being proposed to solve this game. Such a model is used for determining VR metrics, such as tracking accuracy, processing delay, and transmission delay. Deep reinforcement learning was applied to develop a decentralized resource allocation mechanism with minimum transmission overhead for vehicle-to-vehicle (V2V) communications [33]. The decisions concerned with finding the optimal sub-band and power level for the transmission are autonomously taken to support each V2V link. This approach does not use massive MIMO or hybrid beamforming. In [34], the authors proposed a simple reinforcement learning algorithm based on an epsilon-greedy multi-hop to achieve significant energy savings in low-power wide area networks relying on multi-hop topologies. In single-hop topologies, uplink (UL) communications from distant nodes are established with high power levels, whereas transmissions to closer hops are enabled by multi-hop routing in UL, in order to reduce energy consumption. This LoWAN network uses remote with IEEE 802.15.4 physical and MAC layers. The proposed multi-hop dynamic routing approach is not always energy efficient. There is a tradeoff between circuit power consumption and transmission power. There is a need for cross-layer machine learning techniques covering PHY, MAC, and network layers to ensure optimal EE and SE levels.

6. Performance Study

In this section, we summarize our preliminary studies concerning machine learning techniques deployed for user-beam pair selection in the hybrid beamforming design used in 5G communications. Hybrid beamforming is an imperative part of 5G massive MIMO millimeter wave communications [35]. Hybrid beamforming divides the processing of signal for beamforming into two stages: analog beamforming and digital beamforming. Analog beamforming is realized by analog phase shifters, while digital beamforming usually relies on zero-forcing precoders/combiners. This arrangement reduces the number of radio frequency chains required in the digital beamforming and reduces energy consumption. Energy consumption may be further reduced by user-beam allocation in the analog beam domain.

6.1. Hybrid Beamforming and User-Beam Selection Architecture

In hybrid beamforming, beamforming is performed in two stages: an analog beamforming $\mathbf{F}_{AB} \in \mathbb{C}^{N \times N_{RF}}$, and digital beamforming $\mathbf{F}_{DB} \in \mathbb{C}^{N_{RF} \times N_s}$ as shown in Fig. 2.

We consider a single base station scenario. The base station is equipped with $N = 128$ transmit antennas and is serving $K = 8$ users. There are $N_{RF} = K$ RF chains available within the base station. We use an extended Saleh-Valenzuela model which accurately captures the mathematical structure of in mm-Wave channels [36]. For simplicity, we assume that each scattering cluster around the transmitter and the receiver contributes a single propagation path [37]. The near optical line-of-sight (LOS) wave propagation at mm-Wave frequencies results in a limited number of scattering paths, say L. The $N \times K$ MIMO channel matrix at the BS can be written as:

$$\mathbf{H} = \sqrt{\frac{KN}{L}} \sum_{l=1}^{L} \alpha_l \mathbf{a}(\phi_l), \tag{2}$$

where α_l represents the complex gain of the l-th path with i.i.d. $\mathcal{CN}(0,1)$. Moreover, $\mathbf{a}$ is the transmit steering vector. The variable $\phi_l \in [0, 2\pi)$ is the l-th path's azimuth angle.

The analog beamforming matrix, e.g. discrete Fourier transform (DFT) $N \times N$ matrix $\mathbf{U}$ may be used to convert spatial channel $\mathbf{H}$ to a beam domain channel as:

$$\mathbf{H}_b = \mathbf{U}\mathbf{H}\,. \tag{3}$$

Now, the row of beam domain channel $\mathbf{H}_b$ corresponds to N beams. The selection of K beams out of N to maximize the energy efficiency poses a complex user-beam matching

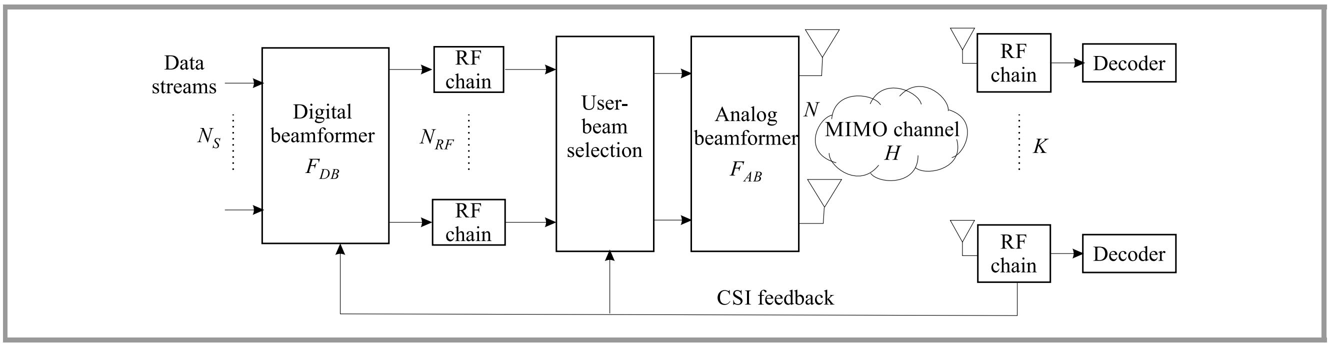

Fig. 2. Hybrid beamforming and user-beam selection architecture.

problem. User-beam allocation is an NP-hard mixed integer programming optimization problem. A global optimal solution to this problem with $K \leq N_{RF}$ (K is the number of users and N_{RF} is the number of RF chains) may only be obtained by using an exhaustive search method. This solution has very high computational complexity of:

$$\sum_{i=1}^{N_{RF}} {}^{K}C_i . {}^{N}C_i . {}^{i}P_i \,,$$

where ${}^{a}C_b = \frac{a!}{b!(a-b)!}$ and ${}^{a}P_b = \frac{a!}{(a-b)!}, a \geq b$.

For the data symbol vector $\mathbf{s} \in \mathbb{C}^{K\times 1}$, the beam domain representation of the system is given by:

$$\mathbf{y} = \mathbf{H}_b^H \mathbf{F}_{DB,b}\mathbf{s} + \mathbf{z} \,, \tag{4}$$

where $\mathbf{F}_{DB,b} = \mathbf{U}^H \mathbf{F}_{DB} \in \mathbb{C}^{M\times K}$ is the beam domain digital precoder and $\mathbf{z} \in \mathbb{C}^{N\times 1}$ is the independent and identically distributed (i.i.d.) complex Gaussian noise with zero-mean and variance of one, i.e. $\mathscr{CN}(\mathbf{0}, \mathbf{I}_N)$. The signal received at user k is given as:

$$y_k = \mathbf{h}_{k,b}^H \mathbf{f}_{DB,k,b}\mathbf{s}[n] + z_k \,. \tag{5}$$

The received SNR γ_k of the user k is:

$$\gamma_k = \frac{|\tilde{\mathbf{h}}_{k,b}^H \mathbf{f}_{DB,k,b}|^2}{\sum\limits_{j\neq k} |\tilde{\mathbf{h}}_{j,b}^H \mathbf{f}_{DB,k,b}|^2 + \frac{1}{\rho}||\mathbf{f}_{DB,k,b}||_2^2} \,, \tag{6}$$

where ρ is the average received SNR per antenna per user and $\tilde{\mathbf{h}}_{k,b}$ is the beam domain channel after the user-beam selection. The sum-rate is given as:

$$R = \sum_{k=1}^{K} \log_2(1+\gamma_k) \,. \tag{7}$$

Finally, the energy efficiency is calculated by Eq. (1).

In our machine learning model, training data is obtained from the Hungarian method [38] and is then applied to the user-beam matching problem. We use a shallow neural network with the input layer of 128 neurons, the hidden layer of 20 neurons and the output layer of 8 neurons. The

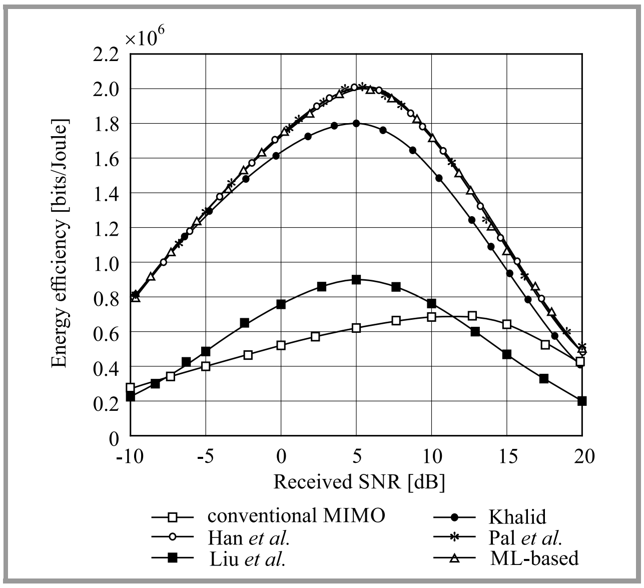

Fig. 3. Energy efficiency versus received SNR with 8 users and 128 transmit antennas.

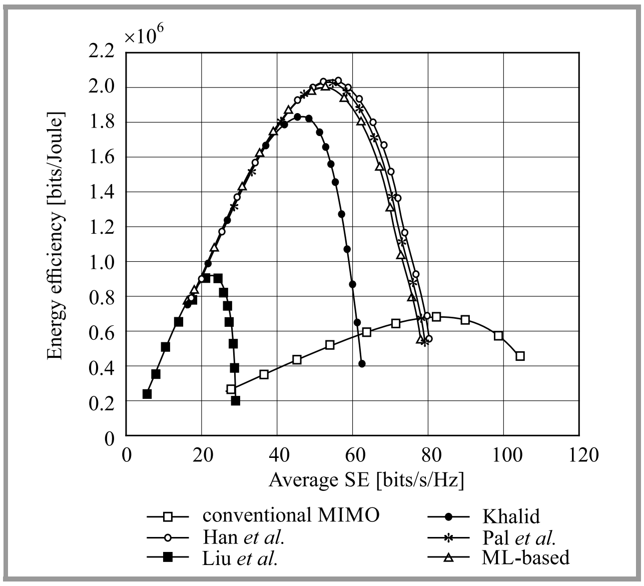

Fig. 4. Energy efficiency and spectral efficiency trade-off with 8 users and 128 transmit antennas.

Table 1
Energy efficiency comparison of some recent research works devoted to signal processing

Authors	Reference paper	Energy efficiency
Han *et al.*	[39]	High
Liu *et al.*	[41]	Low
Khalid	[42]	Medium
Pal *et al.*	[40]	high
ML-based	This paper	High
Conventional MIMO	Benchmark	Low

hidden and output layers use sigmoid activation functions. Figure 3 shows the average system level EE for the received SNR. EE increases with the received SNR, up to the level of 5 dB, and then starts decreasing due to interference and high transmit power. ML-based hybrid beamforming offers performance that is comparable to that of DFT-based [39] and DFT+Greedy-based [40] solutions, but required fewer computational resources (Table 1). It may be observed that conventional MIMO precoding offers the lowest EE performance because of $N_{RF} = N$ RF chains, while in other hybrid beamforming designs $N_{RF} = K$. The trade-off between EE and spectral efficiency (SE) is shown in Fig. 4. EE increases with SE up to a certain level, and then EE starts decreasing. ML-based hybrid beamforming may be used to reduce the computational complexity and achieve comparable, good performance.

7. Conclusions

This article reviews energy efficiency of 5G wireless networks relying on conventional and machine learning techniques. Energy efficiency may be improved by deploying the following techniques: base station on/off, energy harvesting with simultaneous transfer of wireless information and power, small cells, coexistence of LTE and 5G and signal processing algorithms. The latest machine learning techniques may be applied as well. Having analyzed the use of various energy efficiency techniques in 5G networks, we present a comparison of different hybrid beamforming designs in mm-Wave massive MIMO systems. It has been shown that the two-layered shallow feedforward neural network-based ML scheme provides a negligible performance gap. It is expected that machine learning will replace adaptive and computationally intensive portions of future communication systems.

Acknowledgments

This work was supported under HCT Research Grants, grant number 2307 from Higher Colleges of Technology (HCT), Abu Dhabi, United Arab Emirates.

References

[1] "What Will 5G Mean for the Environment?", The Henry M. Jackson School of International Studies", University of Washington [Online]. Available: https://jsis.washington.edu/news/what-will-5g-mean-for-the-environment/

[2] "Power a Greener 5G Era with Huawei 5G Power", Mobile World Live" [Online]. Available: https://www.mobileworldlive.com/huawei-updates/power-a-greener-5g-era-with-huawei-5g-power/

[3] M. Usama and M. Erol-Kantarci, "A survey on recent trends and open issues in energy efficiency of 5G", *Sensors* (Basel), vol. 19, no. 14, 2019 (DOI: 10.3390/s19143126).

[4] 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 (DOI: 10.1109/ACCESS.2019.2907284).

[5] Q. Zhao and M. Gerla, "Energy efficiency enhancement in 5G mobile wireless networks", in *Proc. 2019 IEEE 20th IEEE Int. Symp. on a World of Wirel., Mob. and Multim. Netw. WoWMoM 2019*, Washington, DC, USA, 2019 (DOI: 10.1109/WoWMoM.2019.8792998).

[6] C. Bouras and G. Diles, "Energy efficiency in sleep mode for 5G femtocells", in *Proc. 2017 Wireless Days WD 2017*, Porto, Portugal, 2017, no. 3, pp. 143–145 (DOI: 10.1109/WD.2017.7918130).

[7] F. Mukhlif, K. A. B. Nooridin, Y. A. Al-Gumaei, and A. S. Al-Rassas, "Energy harvesting for efficient 5G networks", in *Proc. 2018 Int. Conf. on Smart Comput. and Electron. Enterp. ICSCEE 2018*, Shah Alam, Malaysia, 2018 (DOI: 10.1109/ICSCEE.2018.8538376).

[8] A. Li and C. Masouros, "Energy efficient MIMO SWIPT by hybrid analog-digital beamforming", in *Proc. 2017 IEEE 18th Worksh. on Sig. Process. Adv. in Wirel. Commun. SPAWC 2017*, Sapporo, Japan, 2017 (DOI: 10.1109/SPAWC.2017.8227793).

[9] I. Alqerm and B. Shihada, "Energy efficient traffic offloading in multi-tier heterogeneous 5G networks using intuitive online reinforcement learning", *IEEE Trans. on Green Commun. and Network.*, vol. 3, no. 3, pp. 691–702, 2019 (DOI: 10.1109/TGCN.2019.2916900).

[10] S. Rizvi *et al.*, "An investigation of energy efficiency in 5G wireless networks", in *Proc. 2017 Int. Conf. on Circ., Syst. and Simul. ICCSS 2017*, London, UK, 2017, pp. 142–145 (DOI: 10.1109/CIRSYSSIM.2017.8023199).

[11] V. Poirot, M. Ericson, M. Nordberg, and K. Andersson, "Energy efficient multi-connectivity algorithms for ultra-dense 5G networks", *Wirel. Netw.*, vol. 26, no. 3, pp. 2207–2222, 2020 (DOI: 10.1007/s11276-019-02056-w).

[12] H. Kour and R. K. Jha, "Power optimization using spectrum sharing for next generation cellular networks", in *Proc. 2019 11th Int. Conf. on Commun. Syst. and Netw. COMSNETS 2019*, Bengaluru, India, 2019, vol. 2061, pp. 562–564 (DOI: 10.1109/COMSNETS.2019.8711197).

[13] S. Noh *et al.*, "Training sequence design for feedback assisted hybrid beamforming in massive MIMO systems", *IEEE Trans. on Commun.*, vol. 64, no. 1, pp. 187–200, 2016 (DOI: 10.1109/TCOMM.2015.2498184).

[14] L. Yang, Y. Zeng, and R. Zhang, "Wireless power transfer with hybrid beamforming: How many RF chains do we need?", *IEEE Trans. on Wirel. Commun.*, vol. 17, no. 10, pp. 6972–6984, 2018 (DOI: 10.1109/TWC.2018.2865313).

[15] K. Roth, H. Pirzadeh, A. Lee Swindlehurst, and J. A. Nossek, "A comparison of hybrid beamforming and digital beamforming with low-resolution ADCs for multiple users and imperfect CSI", *IEEE J. on Selec. Topics in Sig. Process.*, vol. 12, no. 3, pp. 484–498, 2018 (DOI: 10.1109/JSTSP.2018.2813973).

[16] K. Roth and J. A. Nossek, "Achievable rate and energy efficiency of hybrid and digital beamforming receivers with low resolution ADC", *IEEE J. on Selec. Areas in Commun.*, vol. 35, no. 9, pp. 2056–2068, 2017 (DOI: 10.1109/JSAC.2017.2720398).

[17] S. Han, C. L. Io, C. Rowell, Z. Xu, S. Wang, and Z. Pan, "Large scale antenna system with hybrid digital and analog beamforming structure", in *Proc. IEEE Int. Conf. on Commun. Worksh. ICC 2014*, Sydney, NSW, Australia, 2014, pp. 842–847 (DOI: 10.1109/ICCW.2014.6881305).

[18] P. Zhu, H. Mao, J. Li, and X. You, "Energy efficient joint energy cooperation and power allocation in multiuser distributed antenna systems with hybrid energy supply", *IET Commun.*, vol. 13, no. 2, pp. 153–161, 2019 (DOI: 10.1049/iet-com.2018.5181).

[19] X. Ge, Y. Sun, H. Gharavi, and J. Thompson, "Joint optimization of computation and communication power in multi-user massive MIMO systems", *IEEE Trans. on Wirel. Commun.*, vol. 17, no. 6, pp. 4051–4063, 2018 (DOI: 10.1109/TWC.2018.2819653).

[20] H. Shokri-Ghadikolaei, C. Fischione, G. Fodor, P. Popovski, and M. Zorzi, "Millimeter wave cellular networks: A MAC layer perspective", *IEEE Trans. on Communi.*, vol. 63, no. 10, pp. 3437–3458, 2015 (DOI: 10.1109/TCOMM.2015.2456093).

[21] S. He, C. Qi, Y. Wu, and Y. Huang, "Energy-efficient transceiver design for hybrid sub-array architecture MIMO systems", *IEEE Access*, vol. 4, pp. 9895–9905, 2016 (DOI: 10.1109/ACCESS.2017.2649539).

[22] I. Al Qerm and B. Shihada, "Enhanced machine learning scheme for energy efficient resource allocation in 5G heterogeneous cloud radio access networks", in *Proc. IEEE Int. Symp. on Pers., Indoor and Mob. Radio Commun. PIMRC 2018*, Montreal, QC, Canada, 2018, vol. 2017-October (DOI: 10.1109/PIMRC.2017.8292227).

[23] H. Park and Y. Lim, "Reinforcement learning for energy optimization with 5g communications in vehicular social networks", *Sensors* (Basel), vol. 20, no. 8, 2020 (DOI: 10.3390/s20082361).

[24] T. Ding, Y. Zhao, L. Li, D. Hu, and L. Zhang, "Hybrid precoding for beamspace MIMO systems with sub-connected switches: a machine learning approach", *IEEE Access*, vol. 7, pp. 143273–143281, 2019 (DOI: 10.1109/ACCESS.2019.2944061).

[25] I. Ahmed and H. Khammari, "Joint machine learning based resource allocation and hybrid beamforming design for massive MIMO systems", in *Proc. 2018 IEEE Globecom Worksh. GC Wkshps 2018*, Abu Dhabi, United Arab Emirates, 2019 (DOI: 10.1109/GLOCOMW.2018.8644454).

[26] R. Falkenberg, B. Sliwa, N. Piatkowski, and C. Wietfeld, "Machine learning based uplink transmission power prediction for LTE and upcoming 5G networks using passive downlink indicators", in *Proc. 2018 IEEE 88th Veh. Technol. Conf. VTC-Fall 2019*, Chicago, IL, USA, 2019, vol. 2018-August (DOI: 10.1109/VTCFall.2018.8690629).

[27] S. Ali, N. Rajatheva, and W. Saad, "Fast uplink grant for machine type communications: challenges and opportunities", *IEEE Commun. Mag.*, vol. 57, no. 3, pp. 97–103, 2019 (DOI: 10.1109/MCOM.2019.1800475).

[28] S. Essassi, S. Cherif, and M. Siala, "RB allocation based on genetic algorithm and coordination over the $\times 2$ interface in the LTE uplink", in *Proc. 2013 IEEE 24th Int. Symp. on Pers., Indoor and Mob. Radio Commun. PIMRC 2013*, London, UK, 2013, pp. 2424–2428 (DOI: 10.1109/PIMRC.2013.6666552).

[29] U. Challita, L. Dong, and W. Saad, "Proactive resource management for LTE in unlicensed spectrum: A deep learning perspective", *IEEE Trans. on Wirel. Commun.*, vol. 17, no. 7, pp. 4674–4689, 2018 (DOI: 10.1109/TWC.2018.2829773).

[30] P. Khuntia and R. Hazra, "An actor-critic reinforcement learning for device-to-device communication underlaying cellular network", in *Proc. 2018 IEEE Region 10 Ann. Int. Conf. TENCON 2019*, Jeju, South Korea, 2019, vol. 2018-October, pp. 50–55 (DOI: 10.1109/TENCON.2018.8650160).

[31] L. Huang, X. Feng, C. Zhang, L. Qian, and Y. Wu, "Deep reinforcement learning-based joint task offloading and bandwidth allocation for multi-user mobile edge computing", *Digit. Commun. and Netw.*, vol. 5, no. 1, pp. 10–17, 2019 (DOI: 10.1016/j.dcan.2018.10.003).

[32] M. Chen, W. Saad, and C. Yin, "Virtual reality over wireless networks: Quality-of-service model and learning-based resource management", *IEEE Trans. on Commun.*, vol. 66, no. 11, pp. 5621–5635, 2018 (DOI: 10.1109/TCO).

[33] H. Ye and G. Y. Li, "Deep reinforcement learning for resource allocation in V2V communications", in *Proc. IEEE Int. Conf. on Commun. ICC 2018*, Kansas City, MO, USA, 2018, vol. 2018-May (DOI: 10.1109/ICC.2018.8422586).

[34] S. Barrachina-Munoz, T. Adame, A. Bel, and B. Bellalta, "Towards energy efficient LPWANs through learning-based multi-hop routing", in *Proc. IEEE 5th World Forum on Internet of Things WF-IoT 2019*, Limerick, Ireland, 2019, pp. 644–649 (DOI: 10.1109/WF-IoT.2019.8767193).

[35] I. Ahmed *et al.*, "A survey on hybrid beamforming techniques in 5G: Architecture and system model perspectives", *IEEE Commun. Surv. and Tutor.*, vol. 20, no. 4, pp. 3060–3097, 2018 (DOI: 10.1109/COMST.2018.2843719).

[36] F. Sohrabi and W. Yu, "Hybrid digital and analog beamforming design for large-scale MIMO systems", in *Proc. IEEE Int. Conf. on Acoust., Speech and Sig. Process. CASSP 2015*, Brisbane, QLD, Australia, 2015, pp. 2929–2933 (DOI: 10.1109/ICASSP.2015.7178507).

[37] O. E. Ayach, R. W. Heath, S. Abu-Surra, S. Rajagopal, and Z. Pi, "Low complexity precoding for large millimeter wave MIMO systems", in *Proc. IEEE Int. Conf. on Commun. ICC 2012*, Ottawa, ON, Canada, 2012, pp. 3724–3729 (DOI: 10.1109/ICC.2012.6363634).

[38] H. W. Kuhn, "The Hungarian method for the assignment problem", *Naval Research Logistics Quarterly*, vol. 2, no. 1-2, pp. 83–97, 1955. [Online]. Available: https://onlinelibrary.wiley.com/doi/full/10.1002/nav.3800020109 (DOI: 10.1002/nav.3800020109).

[39] Y. Han, S. Jin, J. Zhang, J. Zhang, and K. K. Wong, "DFT-based hybrid beamforming multiuser systems: Rate analysis and beam selection", *IEEE J. of Selec. Topics in Sig. Process.*, vol. 12, no. 3, pp. 514–528, 2018 (DOI: 10.1109/JSTSP.2018.2821104).

[40] R. Pal, A. K. Chaitanya, and K. V. Srinivas, "Low-complexity beam selection algorithms for millimeter wave beamspace MIMO systems", *IEEE Commun. Lett.*, vol. 23, no. 4, pp. 768–771, 2019 (DOI: 10.1109/LCOMM.2019.2902147).

[41] Y. Liu and J. Wang, "Low-complexity OFDM-based hybrid precoding for multiuser massive MIMO systems", *IEEE Wirel. Commun. Lett.*, vol. 9, no. 8, pp. 263–266, 2020 (DOI: 10.1109/LWC.2019.2929518).

[42] F. Khalid, "Hybrid beamforming for millimeter wave massive multiuser MIMO systems using regularized channel diagonalization", *IEEE Wirel. Commun. Lett.*, vol. 8, no. 3, pp. 705–708, 2019 (DOI: 10.1109/LWC.2018.2886882).

Muhammad Khalil Shahid received his B.Sc. in Electrical Engineering from UET Lahore, Pakistan in 1998 and M.Sc. in Electrical Engineering in 2005. He got his Ph.D. in Engineering from BUPT, China in 2008. He has approximately 15 years of hands-on experience, working as an engineer and consultant in the telecom industry. He also has more than 10 years of experience as a trainer and faculty member. His areas of research include wireless communication and networks, optical communication and networks, ICT, as well as telecom policies and regulations. Currently, he is working as Program Team Leader/Assistant Professor at Higher Colleges of Technology, Al-Dhafra Region, UAE.

E-mail: kshahid@hct.ac.ae
Faculty of Electrical Engineering
Higher Colleges of Technology
RUWC/MZWC
Abu Dhabi, United Arab Emirates

Filmon Debretsion is currently an undergraduate student at the Electrical Engineering Department of Higher Colleges of Technology, UAE.

Higher Colleges of Technology
Abu Dhabi, United Arab Emirates

Aman Eyob Abraham is currently an undergraduate student at the Electrical Engineering Department of Higher Colleges of Technology, UAE.

Higher Colleges of Technology
Abu Dhabi, United Arab Emirates

Irfan Ahmed received his B.E. in Electrical Engineering degree and the M.Sc. in Computer Engineering degree from the University of Engineering and Technology, Taxila, Pakistan, in 1999 and 2003, respectively, and the Ph.D. degree in Telecommunication Engineering from the Beijing University of Posts and Telecommunications, Beijing, China, in 2008. Currently, he is working as Associate Professor at Higher Colleges of Technology, UAE. Between 2011 and 2017, he worked as Assistant/Associate Professor in Taif University, KSA. He was post-doctoral fellow with Qatar University from 2010 to 2011, where he worked on two research projects: wireless mesh networks in cooperation with Purdue University, USA, and radio resource allocation for LTE, with Qtel. He was also involved in a National ICT Pakistan-funded research project titled "Design and development of MIMO and Cooperative MIMO test-bed" at Iqra University, Islamabad, Pakistan, from 2008 to 2010. His research interests include wireless LAN (WLAN), medium access control (MAC) protocol design and analysis, cooperative communications, MIMO communications, performance analysis of wireless channels, energy constrained wireless networks, cognitive radio networks, and radio resource allocation. He is the author of more than 25 international publications.
E-mail: irfanahmed44@gmail.com
Higher Colleges of Technology
Abu Dhabi, United Arab Emirates

Tarig Faisal received his Ph.D. degree in Signal Processing from University of Malaya, Malaysia, in 2011. His research focuses on signal processing, intelligent systems, robotics and control. He has almost 20 years of academic and industry experience. Currently, he is the Dean of Academic Operations at Higher Colleges of Technology.
E-mail: tfaisal@hct.ac.ae
Higher Colleges of Technology
Abu Dhabi, United Arab Emirates