ABUAD Journal of Engineering Research and Development (AJERD) ISSN: 2645-2685
Volume 1, Issue 2, 156-165

User Association and Enabling Technologies in Next Generation 5G Ultra-Dense Networks – A Review

Oladele Felix AFOLALU[1,2,*], Morufu Adewale AYOADE[2], Adegboyega Kifli ADEBAYO[2], Isaac Adekunle BAMIKEFA[2]

[1]*Department of Electrical Engineering, University of Cape Town, South Africa*
`aflola001@myuct.ac.za`
[2]*Department of Electrical/Electronics Engineering, Federal Polytechnic, Ede, Osun State*
`maruf672001@yahoo.com/adekifli@yahoo.com/adekunle626@gmail.com`

[*]*Corresponding Author:* `aflola001@myuct.ac.za`

Date of First Submission: `19/03/2018`
Date Accepted: `09/06/2018`

***Abstract**: Embedding small cells and relay nodes in a macro-cellular network is a promising method for achieving substantial gains in coverage and capacity compared to traditional macro only networks. These new types of base-stations can operate on the same wireless channel as the macro-cellular network, providing higher spatial reuse via cell splitting. However, these base-stations are deployed in an unplanned manner, can have very different transmit powers, and may not have traffic aggregation among many users. This could potentially result in much higher interference magnitude and variability. Hence, such deployments require the use of innovative cell association and inter-cell interference coordination techniques in order to realize the promised capacity and coverage gains. In this paper, we review various techniques for user association and interference mitigation which are required to meet increased data demand in next generation 5G ultra-dense network.*

***Keywords:** 5G, HetNets, massive MIMO, mmWave, user association.*

1. INTRODUCTION

Recently, the amount of data traffic on mobile communication systems has rapidly increased because of the widespread use of mobile Internet applications such as browsing and streaming. The data rates requested by end-users have also increased, and will be a factor in determining the future deployment of mobile networks [1]. Long Term Evolution-Advanced (LTE-A) wireless networks are being designed to improve spectral efficiency per unit area by shrinking cell size via deployment of diverse set of base-stations [2], [3]. Hence, future network is expected to be heterogeneous in nature. Heterogeneous Networks (HetNets) could mean a network comprising of different Radio Access Technologies (RATs) (WiFi, GSM, UMTS/HSPA, LTE/LTE-A). The coexistence between various multi-vendors is referred to as Multi-RATs. Furthermore, a HetNet also means a network consisting different access nodes, called small cells (SCs), such as macrocell, microcell, picocells, femtocells, RRHs (Remote Radio Heads), as well as relay stations. This set-up is aptly referred to as two (multiple) tier/layer networks [3]. Small cells are low-power wireless access points, increasing in size from femtocells (the smallest) to macrocells (the largest), that operate in licensed spectrum. They provide improved cellular coverage, capacity and applications for homes and enterprises as well as metropolitan and rural public spaces.

Figure 1 shows a cellular system of macro base-stations deployed in a planned regular manner with transmit power of up to 40 W, overlaid with pico, femto, and relay base-stations which transmit at substantially lower power (100 mW to 2 W), and are typically deployed in an unplanned manner. These overlaid base stations improve coverage and provide capacity gain via higher spatial reuse. Therefore, the difference between heterogeneous deployment paradigm and classical macro-only deployment can be summed up in the following ways [4]:

Interference: HetNet can experience much more severe interference since they are deployed in an unplanned manner. In particular, the closed femtocell group (i.e. where only users in its subscriber group are allowed to connect), strongly interferes with other users in the network. In a macro-dominated cellular network, a terminal always connects to the strongest BS which results to interfering signals being received at a low power than the desired signal [5]. On the contrary, in HetNets, a terminal may connect to a closer small cell for the purpose of cell splitting despite a higher received power from a macro BS. This invariably leads to a net gain in throughput if a careful resource management is performed, to ensure that the loss in capacity due to higher interference is matched with rate gain due to higher spatial reuse.

Traffic Load: Due to fewer number of users that can be served by small cells, the traffic and load can vary at a much faster time, which also leads to more interference as discussed earlier. Hence, a robust resource allocation algorithm in [6], [7] that is dynamic in nature is required to utilize the available resources in the network more efficiently. The aim of this algorithm will be in two-folds: it will incorporate association schemes by taking into account different transmit powers of small cells and macro BS, then jointly associate and allocate resources fairly across cells for maximum system efficiency.

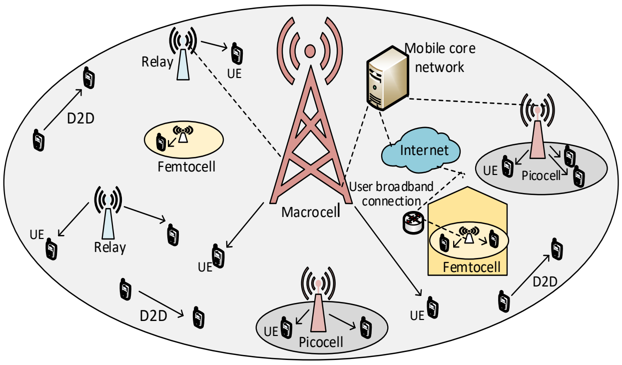

Figure 1: Ultra-dense deployment of BSs in 5G Network.

Improving the cell-edge performance and further increasing the system capacity will be very important because the link performance of LTE is already close to its theoretical limit. Several technical solutions have been discussed mainly by the 3rd Generation Partnership Project (3GPP) to improve the cell coverage, cell-edge throughput, and/or overall system capacity [8]. Coordinated multipoint transmission and/or reception (CoMP) is one of the key technologies employed to improve the cell-edge performance, in which a single user equipment (UE) simultaneously communicates with multiple cells. This can efficiently improve the performance of cell edge UE [1].

The deployment of different low-cost low-power base station nodes (LPN) such as picocells, femtocells and relays will provide the opportunities to increase the capacity within the macrocell area and avoid coverage holes by adapting to the varying nature of user traffic demand. According to a recent study by Ericsson, each macrocell will be overlaid with an average of six LPNs by 2019 to meet the demand for the coverage, mobility and thereby, improve user experience [9]. Also, it was predicted that by 2023, the number of 5G subscriptions for enhanced broadband would reach 1 billion.

Similarly, Cisco further corroborated this forecast in the latest VNI Mobile Forecast report [10]. The report predicted global mobile traffic per user to reach 8,423 megabytes per month by 2021, up from 1,456 megabytes per month in 2016, a compound annual growth rate (CAGR) of 42%.

However, massive deployment of small cells in the coverage of traditional macro BS comes with a lot of challenges. Prominent among these are the problem of Interference, self-organization, mobility management (e.g., handover), backhauling and access control [11]. In-depth explanation on these challenges is beyond the scope of this paper.

In conventional single-layer networks, base station selection was based on the highest reference signal received power (RSRP) measured at user equipment (UE). While this gives the optimum selection methodology for these networks, it does not always apply to the HetNets where base stations have different transmit powers. Macrocell and picocell base stations, namely MeNBs and pico-eNBs, respectively, differ by almost 16 dB in their downlink transmit power levels [3]. If the cell selection is based on RSRP only, UEs are more likely to connect to the MeNBs even when the path loss conditions between the pico-eNB and the UE are better. If the optimal cell selection were assigned to this case, UE could reduce its transmission power since it has higher uplink (UL) signal to interference plus noise ratio (SINR) at the closer pico-eNB. This action will consequently lead to a longer battery life for the UE and reduce the interference in the system. Also, this would lead to a more balanced loading within the macrocell footprint where the resources in MeNBs and LPNs are better utilized.

2. RELATED LITERATURE

Several works have been carried out to analyze both capacity gains and offloading benefits of user association techniques in heterogeneous cellular networks. In [12], combined usage of inter-cell interference coordination (ICIC) and cell range expansion (CRE) was proposed as a very effective way in improving the system and cell-edge throughput. With full buffer model and CRE offset value between 8 and 20dB, it was shown that almost the same user throughput performance is obtained by allocating the appropriate resources to protect UEs that connect to the picocells.

As a follow up to the work in [12], the authors in [13] investigated the transmission power control (TPC) method in the heterogeneous networks that employ CRE and evaluates the cell-edge user throughput and cell throughput performance. Simulation results show that almost the same cell-edge user throughput is obtained by setting the appropriate difference in the target received signal power between the macrocell and picocells according to the CRE offset value. Although, most studies assume fixed bias value for picocells, factors like geographical location of BS, network density and mobility pattern of users may render this assumption void in ultra-dense scenario. However, [14] substantiated this fact by taking the approach further to dynamically change the range of low power nodes in the presence of a time based ICIC method. This was achieved by designing a simple heuristic algorithm which adapts the size of the low power nodes to the load and interference situation. The outcome of this approach shows that dynamic range expansion provides better performance for the 5th and 50th percentile users but causes loss to the 95th percentile users of a system.

Other works in [15] - [18] adopted similar methods to balance load and improve spectral efficiency in macrocell-underlaid small cell networks. However, our previous work in [29] sought to find an optimal bias value in a network containing a single macro BS overlaid with four pico BSs. We formulated the problem as a mixed integer programming problem (MIP), in an attempt to determine users' association based on minimizing resource block allocation. The results of our proposed method showed a remarkable reduction in number of outages and an increase in the number of offloaded UE at the optimal offset value (12dB).

With the exponential growth of cellular devices, the ever-increasing data demand leads to more energy consumption of base stations in cellular networks. As reported in [16], BSs consume a significant portion, amounting to about 60%−80%, of energy in a wireless cellular network. Therefore, reducing the energy consumption of BSs in cellular networks has become a strategic target for the cellular mobile communications industry. Hence, the authors [16] considered a scenario where macro BSs are powered by on-grid energy while pico-cell BSs can be powered by either on-grid energy or green energy. A constrained energy cost saving optimization problem was formulated, with the objective of minimizing the total energy cost, while guaranteeing the users' Quality of Service (QoS) requirement. Thereafter, an adaptive range expansion (ARE) algorithm, which adaptively set the biasing factor for each pico BS according to its estimated energy drain ratio (EDR) to maximize the green energy utilization was proposed. With regards to the above scheme, simulation results validated the proposed algorithm to achieve significant improvement in terms of the total energy cost, compared with the max-RSRP method and the conventional RE method.

In recent years, Vehicular communications have attracted more attention from both industry and academia due to their strong potential to enhance road safety, improve traffic efficiency, and provide rich on-board information and entertainment services [19]. To this end, various communication standards have been developed across the globe to ensure interoperability in information exchange of vehicles, e.g., dedicated short-range communications (DSRC) standards in the US [20] and intelligent transportation system (ITS)-G5 standards developed by the European Telecommunications Standards Institute (ETSI) [7]. However, recent studies [21], [22] show that time variation within an Orthogonal Frequency Division Multiplexing (OFDM) symbol destroys orthogonality among subcarriers. This introduces ICI which, if not properly accounted for, would result in an error floor, which increases with vehicle mobility and carrier frequency.

Therefore, the main contribution of this paper is to identify and compare the different schemes of cell association and load balancing techniques recently studied in Ultra-dense/Heterogeneous cellular Networks. The aim is to obtain the potency and weakness of these technologies. Thus, future challenges in performing the techniques are proposed in order to meet year 2020 target in the deployment of 5G.

The remaining part of the paper is organized as follows. Section 3 presents an overview of LTE-A network architecture enabled for OFDM and multiple-input-multiple-output (MIMO) communications. Section 4 elaborates on the existing cell selection/user association algorithms for classical and Ultra Dense Network (UDN). The recent emerging issues and open challenges of user association in 5G networks are discussed in Section 5. Finally, the paper is concluded with conclusion and future works in Section 6.

3. LTE-A NETWORK ARCHITECTURE

The LTE/LTE-A system architecture consists of two layers: a radio access network, called Evolved-Universal Terrestrial Radio Access Network (E-UTRAN) and a core network known as Evolved Packet Core (EPC) [27], [28]. The architecture of the core network is also referred to as the Service Architecture Evolution (SAE) and the combination of E-UTRAN and EPC/SAE is also called the Evolved Packet System (EPS). The network architecture is illustrated in Figure 2, which has the

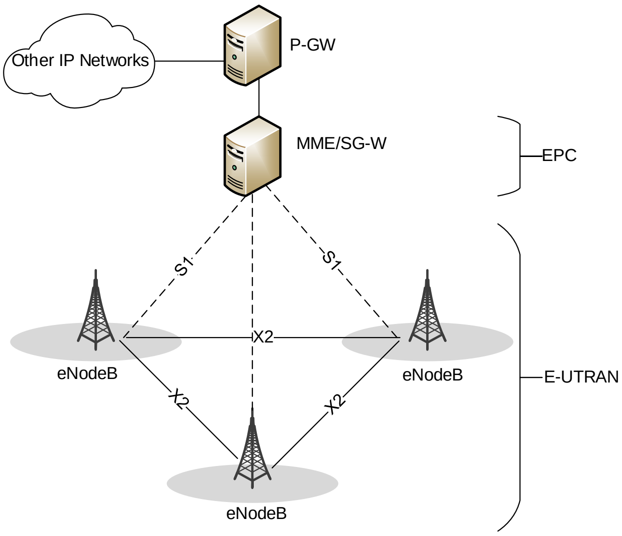

Figure 2: LTE architecture [27]

Mobility Management Entity (MME) and the Serving Gateway (S-GW) located at the core network and connected to the LTE/LTE-A base stations, called evolved-NodeBs (eNodeBs) via the S1 interface.

The MME entity handles several functions related to network access control, radio resource management, and mobility management, while the S-GW acts as a local mobility anchor point for inter-eNodeB handovers and for the handling of data packet transfer between the core network and the UEs [27]. The Packet data network Gateway (P-GW) provides connectivity between the core network and other Internet Protocol (IP) networks. It also serves as an anchor for mobility between 3GPP and non-3GPP technologies. The radio access network, which is the first point of entry for a UE to the LTE network, is comprised of eNodeBs and UEs. The eNodeBs are the single logical nodes in the radio access network normally inter-connected with each other through the X2 interface, that allows the exchange of signaling messages and information related to resource usage and power allocation [28]. In addition, E-UTRAN ensures link level reliability, segmentation, and reassembly of higher-layer Protocol Data Units (PDUs) and IP header compression. These protocols between the UE and the E-UTRAN are collectively referred to as Access Stratum (AS) protocols.

3.1 LTE Frame Structure

In LTE/LTE-A systems, Orthogonal Frequency Division Multiple Access (OFDMA) technique is selected as the multiple access technology on the downlink of the radio interface and a flat, all-IP (Internet Protocol) in the core network design [24], [27]. OFDMA enables multiple access among different UEs using OFDM transmission where different UEs belonging to the same eNodeB are allocated a different number of subcarriers according to their individual traffic and QoS requirements. The channel gain experienced by a UE in a subcarrier is also taken into account during allocation. The multicarrier structure of OFDMA introduces the possibility of dynamic resource allocation.

The available bandwidth is divided into several orthogonal subcarriers which eliminates intra-cell interference. For purposes of resource allocation, the LTE is divided in time and frequency into small scheduling unit called Resource Block (RB), and it consists of 12 subcarriers in the frequency domain, and six OFDM symbols in the time domain in the case of normal cyclic prefix, or seven OFDM symbols in the time domain in the case of short cyclic prefix. RB duration is 0.5 ms, and it occupies a spectrum of 180 kHz. The scheduling period is called Transmit Time Interval (TTI), and it equals 1 ms.

During one TTI, each RB is exclusively assigned to one UE in a given cell, and it could be simultaneously used in the neighboring cells for different UEs. Consequently, ICI problems occur due to the dense usage of the available frequency resources. However, LTE supports a variety of bandwidths up to 20MHz (100 RBs) with minimum bandwidth of 1.4MHz, corresponding to 6 RBs. The number of Physical Resource Blocks (PRBs) for different LTE bandwidths is listed in Table 1 [28].

Table 1: Number of RBs for different LTE Bandwidths [28]

Channel Bandwidth (MHz)	**Number of RBs**
1.4	6
3	15
5	25
10	50
15	75
20	100

4. CELL SELECTION SCHEME

In the traditional homogenous network, cell selection is performed based on the downlink received signal strength (RSS) which allows mobile users to connect to the base station from which the received power is the strongest [25]. However, dense HetNets are likely to become the dominant scheme during the wireless evolution towards 5G. Hence, the conventional max-RSS user association approach will become unsuitable for HetNets, since the transmit power disparity of marcocells and small cells will lead to the association of most of the users with the macro BS. This will potentially result in inefficient small cell deployment. To this end, some of the conventional approaches towards cell selection are discussed below.

4.1 RSRP-based Cell Selection

A mobile user camped on a particular cell will monitor the system information and paging of that cell, and at the same time continue to monitor the quality and strength of the other cells to determine if handover or cell reselection is required. The Reference Signal Received Power (RSRP)-based cell selection is carried out by choosing the base station that maximizes the received power of the reference signals such that [4]

$$CellID_{serving}(j) = \arg\max_{i}\{RSRP_{ij}\} \quad (1)$$

where $RSRP_{ij}$ represents the RSRP at user j coming from cell i and $CellID_{serving}(j)$ is the donor cell which is selected by user j.

According to (1), each mobile user selects its serving cell ID to correspond to the cell from which the largest RSRP is received. However, if this approach is applied in HetNets, it would result in most users being served by the Macro base station (MBS) due to the lower transmission power of small cells. Hence, there would be high competition for the available resources of the Macrocell, while the resources of the small cells would not be fully utilized.

4.2 RSRQ-based Cell Selection

The Reference Signal Received Quality (RSRQ) measurement provides additional information when RSRP is not sufficient to make a reliable handover or cell reselection decision. In the procedure of handover, the LTE specification provides the flexibility of using RSRP, RSRQ, or both [14]. It indicates quality of the received signal with a typical range of -19.5dB (bad) to -3dB (good).

$$RSRQ = N * (\mathrm{RSRP/\,RSSI}) \quad (2)$$

which can be re-written as

$$RSRQ = 10\log(\mathrm{N}) + \mathrm{RSRP(dBm)} - \mathrm{RSSI(dBm)} \quad (3)$$

where N is the number of RBs per channel bandwidth. The Reference Signal Strength Indicator (RSSI) represents the entire received power including the wanted power from the serving cell as well as all co-channel power and other sources of noise. In addition, UE usually measures RSRP or RSRQ based on the direction of radio resource control (RRC) message from the network and report the value, using the real RSRQ value. Thus, it sends a non-negative value ranging from 0 to 34 and each of these values are mapped to a specific range of real RSRQ value [19].

4.3 CRE-based Cell Selection

This scheme increases downlink coverage footprint of low power nodes such as picocells by adding a positive bias to their measured signal strengths during cell selection or association [20]. Although RSRP-based cell selection only reflects the received power from each cell and does not reflect the channel quality of the respective resources, simple extension by adding offset values to the picocells in cell range expansion (CRE) can be used to compensate for the difference in channel quality between the macrocells and picocells [13] - [15], [17], [27]. For RSRP-based cell selection using CRE, the UE selects the cell index based on the following criteria.

$$CellID_{serving}(j) = \arg\max_{i}\{RSRP_{ij} + \Delta_i\} \qquad (4)$$

$$\Delta_i = \begin{cases} \Delta & ; \quad if \quad i \quad is \quad a \quad Picocell \\ 0 & ; \quad if \quad i \quad is \quad a \quad Macrocell \end{cases} \qquad (5)$$

where Δ_i represents the bias which is applied to a user located around cell i.

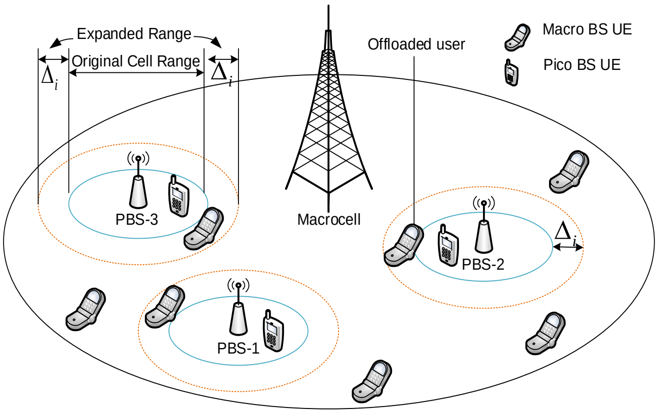

Figure 3: Cell Range Expansion technique in HetNets [14].

This cell selection scheme adds a bias value Δ_i in RSRP to allow more users select Pico BS as their serving cell (Figure 3). Typically, 0dB bias value is assigned to macrocell while a non-negative value is assigned to small cells. Such a strategy may offload significant amount of traffic from the macro cell to small cells, but it can also lead to more interference from the macro base-station, particularly the cell edge users who are no longer connected to BS with the strongest signal [25].
It has further been shown also that in a typical macro-pico deployment with a small CRE bias, downlink interference is not a problem. However, when a high CRE bias is used, the cell edge pico-UEs suffer serious interference from the macrocell as illustrated in Figure 4. Hence, if the bias value is not properly set, a large number of users will not access the small cells because of their small coverage.

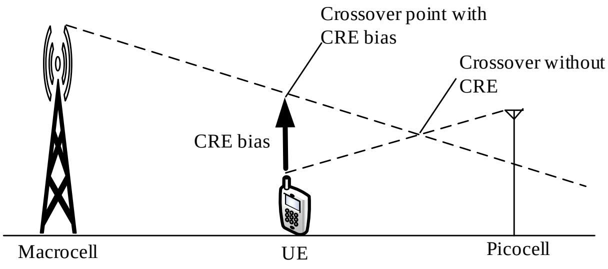

Figure 4: CRE bias for picocell expansion [9].

On the other hand, high bias value will admit more users to the small cells. These users will not be scheduled due to their relatively long distance from their serving cell, which may lead to scheduling outage [9], [18]. Also in Figure 4, without CRE bias the RSRP crossover point may occur quite close to the picocell due to its handicap in the transmission power with respect to the macrocell. This may not offer too much traffic offloading from the macro BS, defeating the purpose of picocell deployment [13]. Thus, the bias value for RE strategy should be carefully designed to get an optimal system performance.

4.4 SINR-based Cell Selection

In order to obtain accurate channel quality for cell selection, the average received SINR is the best criterion [16]. This approach is often used for optimal offloading effect and to manage interference for the user located in the expanded region using Almost Blank Subframe (ABS).

$$CellID_{serving}(j) = \arg\max_{i}\{SINR_{ij} + \Delta_i\} \qquad (6)$$

where $SINR_{ij}$ represents the SINR at user j coming from cell i and $CellID_{serving}(j)$ is the donor cell which is selected by user j. Hence, in other to protect the offloaded users connected to picocell, the macrocell only needs to stop transmission for a specific time resource in ABS, while only reference signals are transmitted.

In this manner, the users in range expanded region are scheduled within ABS to avoid strong cross-tier interference. Then, the remaining normal subframes are allocated to the users closely located to the picocells and macrocells. However, when a large bias value is used for more offloading effect, more subframes need to be kept silent to protect offloaded users. In other words, if the ratio of ABS to non-ABS is low, average data rate of the offloaded users might be lower due to heavily loaded users within ABSs [29]. Therefore, the ABS ratio is one of the important factors to be considered in cell selection process when range expansion is applied.

4.5 Classification of Cell Selection

The existing work [17]-[19] on cell association can be broadly classified into two groups:

4.5.1) Strategies based on Channel Borrowing: The channel borrowing technique has been used in conjunction with dynamic load balancing scheme to cope with the problem of tele-traffic overloads in hot spots [30]. The approach is to denote a particular cell within a hot spot as the center cell. Then the hot spot can be conceived as a stack of hexagonal ‘Rings’ around the center cell such that each ‘Ring’ consists of at least one hot cell. A hot spot with only hot cells is called complete, otherwise it is incomplete.

Hence, in load balancing approach, a hot cell in Ring i borrows channels from its adjacent cells in Ring $i+1$ to ease out its high channel demand. This structured lending mechanism decreases excessive co-channel interference and borrowing conflicts, which are prevented through channel locking in other schemes. Also, the number of channels to be borrowed by each cell will be predetermined by its class and its position in the hot spot.

Two basic methods proposed in the literature as criteria in the channel borrowing strategy are Channel Borrowing without Locking (CBWL) and Load Balancing with Selective Borrowing (LBSB) [18]. In the CBWL scheme, channel borrowing is utilized when the set of channels in a cell gets exhausted, but to use them under reduced transmission power. This is done to avoid interference with the other co-channel cells of the lending cell using the same frequency. Channels can be borrowed only from adjacent cells in an orderly fashion.

The set of channels in a particular cell is divided into seven groups. One group is exclusively for the users in that cell, while each of the six other groups caters for channel requests from one of the neighboring cells. If the number of channels in a channel-group is exhausted, a subscriber using one of the channels can be switched to an idle channel in another group, thereby freeing up one in the occupied group. Since borrowed channels are transmitted at low power, not all users (within range) are capable of receiving them. If such a user finds all the channels occupied, an ordinary user occupying a regular channel can handover its channel to the former while itself switching over to a borrowed channel if available [26].

However, in other to correct the high channel demand in CBWL, LBSB attempts to alleviate these problems by selectively borrowing channels by a cell before the available channel set is exhausted. A cell is classified as ‘hot’ if its degree of coldness (defined as the ratio of the number of available channels to the total number of channels allocated to that cell) is less than or equal to some chosen threshold value. The LBSB scheme proposes to migrate a fixed number of channels from certain cold cells to a hot one according to a channel borrowing algorithm which can be implemented either as a centralized scheme or a distributed scheme [30].

4.5.2) Strategies based on Traffic Transfer: In a cellular network, a fixed number of channels is normally assigned to each cell. However, under this approach, the channel usage may not be efficient because of the variability in the offered traffic. Different approaches such as dynamic channel allocation (DCA) [17] and mobile-assisted connection-admission (MACA) [19] algorithm, which is an extension of channel borrowing scheme, are used to accommodate variable traffic and to achieve load balancing in a cellular network. In this scheme, some special channels are used to connect mobile users from different cells; thus, a mobile user, which is unable to connect to its own base station because it is in a heavily-loaded cell, may be able to get connected to its neighboring lightly-loaded cell through a two-hop link. This technique can, therefore, greatly improve the performance of a cellular network.

Earlier works on traffic steering oriented load balancing were predominantly based on the concept of cell breathing in macro cell only networks [30], [36]. Cell breathing techniques were employed in legacy mobile networks such as GSM, WCDMA, and later on, other wireless networks such as WiFi and WiMaX [30]. The technique is basically a traffic congestion relieving mechanism that adjusts the coverage area of a cell according to the load. Different from CRE technique, an overloaded cell’s coverage area shrinks whilst a lightly loaded cell’s coverage area expands to compensate. The shrinking of the overloaded cell offloads users from overloaded cell, forcing them to associate with lightly loaded neighbouring cells to balance the overall network load.

Usually, cell breathing is implemented by controlling base station or access point transmit power to adjust the coverage area. There are different variants of cell breathing techniques. In CDMA systems for instance, the coverage area of a cell is adjusted by controlling the pilot signal power, whilst in GSM it could be done by adjusting power of other control channels. Even though the idea of adjusting transmit power to cover specific geographical areas has been in existence ever since the conception of cellular wireless networks, it was not initially used for traffic steering. Transmit power adjustment was mainly used for network planning purposes and interference mitigation amongst neighbouring cells.

5. EMERGING ISSUES AND CHALLENGES IN USER ASSOCIATION

The proliferation of multimedia applications and high-end devices (e.g., smartphones, tablets, wearable devices, laptops, machine-to-machine (M2M) communication devices) exacerbates the demand for high data rate services [23], [24] The fifth generation (5G) mobile networks are envisioned to support huge amount of data traffic with reduced energy consumption and improved (QoS) provision.

To this end, key enabling technologies, such as ultra-dense network (UDN), massive multiple-input multiple-output (MIMO), and millimeter wave (mmWave) techniques, have been identified to bring 5G to fruition. Regardless of the technology adopted, a user association mechanism is needed to determine whether a user is associated with a particular base station (BS) before data transmission commences. Hence, user association plays a pivotal role in enhancing the load balancing, the spectrum efficiency, and the energy efficiency of networks. In the following sub-sections, we present an extensive overview of different performance metrics and some state-of-the-art techniques in user association, envisioned for the fifth generation mobile networks. We therefore conclude by highlighting some key challenges like the uplink-downlink asymmetry, mobility support and the backhaul bottleneck which are inherent in HetNets user association design.

5.1 Performance Evaluation of User Association in Ultra-Dense Networks

Dense HetNets are likely to become the dominant theme during the wireless evolution towards 5G [31]. However, the traditional downlink reference signal received power (RSRP), which is the most basic cell association criterion will not lead to much offloading since the transmit power of a macrocell is much greater than that of the small cell. Therefore, adding a bias to the small cell RSRP is an example of an association method that increases offloading.

Although, this concept has been discussed in Section 4 of this paper, the performance metrics which provide a qualitative comparison of the user association scheme are treated here. The primary performance metric employed for user association analysis is outage/coverage probability [32]. When taking into consideration the effect of co-channel interference imposed on radio links during evaluation and planning, the probabilities that the SINR drops below and rises above a certain threshold are, respectively, defined as outage and coverage probability. This is an indicator of the average throughput of a randomly chosen user in the network, and serves as a fundamental metric for network performance analysis and optimization.

Other performance metrics that will be discussed briefly are energy efficiency, spectrum efficiency, quality of service (QoS) and fairness. As data traffic continues to grow due to the proliferation of smart phone devices, reducing greenhouse emission by improving the network energy efficiency is becoming increasingly important. Such improvements may be achieved by advanced system design at both the network and the mobile terminal sides, although the majority of savings may come from the infrastructure.

For instance, one promising technique for cellular network power reduction is through advanced radio resource management (RRM), to dynamically switch ON or OFF a base station based on daily traffic variation [33]. By switching OFF lightly loaded base station and offloading the traffic to neighboring base stations, significant energy reduction may be possible without compromising the network performance. Also, spectrum efficiency is an important metric for performance comparison, due to the scarcity of the frequency spectrum. With the evolution towards 5G networks, coupled with the high demand for data, a high spectrum efficiency is mandatory to meet this requirement. Therefore, it refers to the maximum information rate that can be transmitted over a given bandwidth in a specific communication system [24].

Fairness is a desirable metric in wireless networks. It assists in determining if there is fair sharing of network resources amongst users. Various methods of evaluating fairness are proposed in literature, which include min-max fairness, TCP fairness, and Jain's fairness index [30], [32]. The evaluation of a system for fairness generally depends on the resources shared amongst entities of a system. The resources can either be time or frequency-based resources or a combination of both resource dimensions, and are measured as time slots, sub-channels and resource blocks respectively [34]. To express this mathematically, let us assume that a system allocates a total of x resources to n contending users such that the $i\text{th}$ user receives allocation x_i, then Jain's fairness index is given by

$$f(X) = \frac{\left[\sum_{1=i}^{n} x_i\right]^2}{n\sum_{i=1}^{n} x_i^2} \tag{7}$$

where $0 \le f(X) \le 1$.

5.2 User Association in Massive MIMO and mmWave Networks

The massive MIMO and mmWave technologies provide vital means to resolve many technical challenges of the future 5G UDN, and they can be seamlessly integrated with the current networks and access technologies [35]. For next-generation wireless data networks, massive MIMO promises significant gains that offer the ability to accommodate more users at higher data rates with better reliability while consuming less power.

The deployment of a massive number of antennas at the transmitter and/or receiver can significantly enhance the spectral and energy efficiency of the wireless network. These performance gains can be achieved with simple beamforming strategies, such as maximum ratio transmission (MRT) or zero forcing (ZF) [36]. The mmWave frequency range (30–300 GHz) can offer a huge segment of spectrum that is still underutilized to cater for the capacity requirements of the next generation wireless network, with demand for the exploitation of frequency bands above 6 GHz.

Compared to current small scale MIMO networks, massive MIMO systems achieve high power and spectrum efficiencies, despite their low complexity transceiver designs [31]. The distinctive characteristics of both massive MIMO and mmWave technologies inevitably necessitate the redesign of user association algorithms. For example in [32], a linear transmit pre-coding (TPC) stochastic geometry based approach is utilized for analyzing the impact of massive MIMO on the max-RSS user association. The results show that with reduction in macro BS transmit power at per with the deployed pico BS, the large array gain brought by the massive MIMO macro BS resulted in a user to much more likely be associated with the micro BS than with the pico BS.

On the other hand, the current standards for mmWave communications, such as the IEEE 802.11ad and IEEE 802.15.3c, adopt RSS-based user association scheme which may lead to an inefficient use of resources [31]. However, a novel user association method using combinatorial optimization was introduced in [33], which considered the supported achievable rate and the number of users in each cell to allow small cell BSs to accommodate more data traffic and to maximize the system's data rate. To further extend the work of [33], user association was considered in a hybrid HetNet, where macro cells adopt massive MIMO and small cells adopt mmWave transmissions [34]. The work shows the capabilities of the proposed algorithm to efficiently coordinate massive MIMO and mmWave in the future wireless networks.

5.3 Uplink-Downlink Asymmetry

Most of the research on user association in HetNets investigated the problem from either a downlink or an uplink perspective. However, HetNets typically introduce an asymmetry between uplink and downlink in terms of the channel quality, the amount of traffic, coverage, and the hardware limitations [31]. Although, the uplink and downlink coverage asymmetry is the most severe out of these. In the downlink, due to the large power disparities between the different BS types in a HetNet, macrocells have much larger coverage areas than small cells. On the contrary, the users' devices may transmit at the same power level in the uplink, regardless of the BS type. Although some promising results related to decoupling of the uplink and downlink user association have been reported in [35], [36], this decoupling inevitably requires a tight synchronization as well as a high-speed and low-delay data connectivity between BSs. Hence, it is extremely important to use state-of-the-art joint uplink and downlink user association optimization in UDN.

5.4 Mobility Support

The increased cell densification, coupled with reduced transmit powers of small cells encountered in HetNets leads to reduced footprints and continuously poses challenges for mobility support [29]. Consequently, for a user having moderate or high mobility, a user association algorithm that does not consider the mobility issues may result in more frequent handovers among the cells in HetNets compared to conventional homogeneous cellular networks. Hence, it is imperative to account for user mobility, when making user association decisions in UDN in order to enhance the long-term system-level performance and to avoid excessive handovers.

5.5 Backhaul Bottleneck

Among the challenges faced by HetNets' paradigm shift is backhaul bottleneck in relation to 4G LTE network. Specifically, most of the research assumed a perfect backhaul between the BS and the network controller, and focused on the achievable performance gains of the wireless front-end without taking into account the specific details of the backhaul implementation and any possible backhaul bottleneck [21]. This assumption is usually valid for well-planned traditional macrocells. However, the potentially densely deployed small BSs may impose an overwhelming backhaul traffic in UDN. Then again, the current small cell backhaul solutions, such as xDSL and non-line-of-sight (NLOS) microwave, are far from an ideal backhaul solution owing to their limited data rate. Thus, it is important for UDN to incorporate backhaul-aware user association mechanisms, which fully take the backhaul capacity constraint into account.

In Summary, we have classified some of these technologies according to what they are trying to achieve in Table 2.

Table 2: Hierarchy of different Enabling Technologies

	Energy Efficiency	Capacity Enhancement	Coverage
Massive/3D MIMO	✓	✓	
Dense HetNets	✓	✓	✓
Multi-RAT		✓	
Millimetre Wave	✓	✓	✓

6. CONCLUSIONS AND FUTURE WORK

This work has surveyed various methods of user association techniques in the context of HetNets, massive MIMO and mmWave technologies. In each context, outage/coverage probability, energy efficiency, spectrum efficiency, QoS and fairness were adopted as metrics for determining the best BS to serve a particular user. At a point, much emphases were laid on the existing user association schemes, while recent advances and open challenges in these areas for the emerging 5G networks were subsequently highlighted. Since the traditional max- RSS user association approach is unsuitable for HetNets, more in-depth investigations need to be carried out in order to better accommodate the inherent features of future generation enabling technologies, so as to realize the full potential of 5G networks.

For the upcoming new 5G technology, advanced schemes and architectures for joint user association and interference mitigation should be developed so that the spectrum efficiency as well as the capacity of the network is enhanced. In addition, most of the research on user association in HetNets investigated the problem from either a downlink or an uplink perspective. However, HetNets typically introduce an asymmetry between uplink and downlink in terms of the channel quality, the amount of traffic, coverage, and the hardware limitations. Hence, to realize the goal of 5G networks, it is imperative to design a robust joint uplink and downlink user association optimization algorithm. In the future, we will focus on review of radio resource management (RRM) and various use cases of 5G ultra-dense networks in Vehicular communication.

ACKNOWLEDGMENT

The authors wish to acknowledge the assistance and support of the Federal Government of Nigeria through Tertiary Education Trust Fund (TETFund) intervention grant.

REFERENCES

[1] K. Kikuchi and H. Otsuka, "Proposal of adaptive control CRE in heterogeneous networks," in 2012 IEEE 23rd International Symposium on Personal, Indoor and Mobile Radio Communications - (PIMRC), 2012, pp. 910–914.

[2] R. Madan, J. Borran, A. Sampath, N. Bhushan, A. Khandekar, and T. Ji, "Cell Association and Interference Coordination in Heterogeneous LTE-A Cellular Networks," IEEE J. Sel. Areas Commun., vol. 28, no. 9, pp. 1479–1489, Dec. 2010.

[3] K. Davaslioglu and E. Ayanoglu, "Interference-based cell selection in heterogenous networks," in Information Theory and Applications Workshop (ITA), 2013, 2013, pp. 1–6.

[4] [4] M. Shirakabe, A. Morimoto, and N. Miki, "Performance evaluation of inter-cell interference coordination and cell range expansion in heterogeneous networks for LTE-Advanced downlink," 2011, pp. 844–848.

[5] A. Morimoto, N. Miki, H. Ishii, and D. Nishikawa, "Investigation on transmission power control in heterogeneous network employing cell range expansion for LTE-Advanced uplink," *in Wireless Conference (European Wireless)*, 2012 18th European, 2012, pp. 1–6.

[6] M. Al-Rawi, “A dynamic approach for cell range expansion in interference coordinated LTE-advanced heterogeneous networks,” *in 2012 IEEE International Conference on Communication Systems (ICCS)*, 2012, pp. 533–537.

[7] A. Daeinabi, K. Sandrasegaran, and P. Ghosal, “A dynamic cell range expansion scheme based on fuzzy logic system in LTE-advanced heterogeneous networks,” *in Telecommunication Networks and Applications Conference (ATNAC)*, 2014 Australasian, 2014, pp. 6–11.

[8] Q. Kong and B. Wang, “Adaptive Cell Range Expansion for Energy Cost Saving in Heterogeneous Cellular Networks with Hybrid Energy Supplies,” *in 2015 IEEE 29th International Conference on Advanced Information Networking and Applications*, 2015, pp. 829–836.

[9] R. Sasikumar, V. Ananthanarayanan, and A. Rajeswari, “An Intelligent Pico Cell Range Expansion Technique for Heterogeneous Wireless Networks,” *Indian J. Sci. Technol.*, vol. 9, no. 9, Mar. 2016.

[10] B. A. Yasir, G. Su, and N. Bachache, “Range expansion for pico cell in heterogeneous LTE #x2014; A cellular network,” *in 2012 2nd International Conference on Computer Science and Network Technology (ICCSNT)*, 2012, pp. 1235–1240.

[11] L. Liang, H. Peng, G. Y. Li, and X. Shen, “Vehicular Communications: A Physical Layer Perspective,” *IEEE Trans. Veh. Technol.*, vol. 66, no. 12, pp. 10647–10659, Dec. 2017.

[12] M. Yassin, “Inter-Cell Interference Coordination in Wireless Networks,” Université de Rennes 1, 2015.

[13] D. Tandel and T. Shah, “Cell Selection Techniques in Heterogeneous LTE-Advanced System,” *Int. J. Recent Innov. Trends Comput. Commun.*, vol. 2, no. 1, pp. 88–91, 2014.

[14] P. T. V. Bhuvaneswari, S. Indu, N. L. Shifana, D. Arjun, and A. S. Priyadharshini, “An analysis on Cell Range Expansion in 4G LTE networks,” *in 2015 3rd International Conference on Signal Processing, Communication and Networking (ICSCN)*, 2015, pp. 1–6.

[15] T. Qu, D. Xiao, D. Yang, W. Jin, and Y. He, “Cell selection analysis in outdoor heterogeneous networks,” *in 2010 3rd International Conference on Advanced Computer Theory and Engineering(ICACTE)*, 2010, vol. 5, pp. V5-554-V5-557.

[16] S. Moon et al., “Interference Management with Cell Selection Using Cell Range Expansion and ABS in the Heterogeneous Network Based on LTE-Advanced,” *Wirel. Pers. Commun.*, vol. 81, no. 1, pp. 151–160, Mar. 2015.

[17] J. Oh and Y. Han, “Cell selection for range expansion with almost blank subframe in heterogeneous networks,” *in Personal Indoor and Mobile Radio Communications (PIMRC), 2012 IEEE 23rd International Symposium on*, 2012, pp. 653–657.

[18] W. N. Nyembe, “Load Balancing using Cell Range Expansion in LTE Advanced Heterogeneous Networks,” University of Cape Town, 2015.

[19] D. Liu et al., “User Association in 5G Networks: A Survey and an Outlook,” *IEEE Commun. Surv. Tutor.*, vol. 18, no. 2, pp. 1018–1044, Secondquarter 2016.

[20] Q. Ye, B. Rong, Y. Chen, M. Al-Shalash, C. Caramanis, and J. G. Andrews, “User Association for Load Balancing in Heterogeneous Cellular Networks,” *IEEE Trans. Wirel. Commun.*, Jun. 2013, vol. 12, no. 6, pp. 2706–2716.

[21] T. E. Bogale and L. B. Le, “Massive MIMO and mmWave for 5G Wireless HetNet: Potential Benefits and Challenges,” *IEEE Veh. Technol. Mag.*, vol. 11, no. 1, Mar. 2016, pp. 64–75.

[22] A. He, L. Wang, Y. Chen, M. Elkashlan, and K. K. Wong, “Massive MIMO in K-Tier Heterogeneous Cellular Networks: Coverage and Rate,” *in 2015 IEEE Global Communications Conference (GLOBECOM)*, 2015, pp. 1–6.

[23] K. Sakaguchi et al., “Millimeter-wave evolution for 5G cellular networks,” *IEICE Trans. Commun.*, vol. 98, no. 3, pp. 388–402, 2015.

[24] B. Xu, Y. Chen, M. Elkashlan, T. Zhang, and K. K. Wong, “User association in massive MIMO and mmWave enabled HetNets powered by renewable energy,” *in 2016 IEEE Wireless Communications and Networking Conference*, 2016, pp. 1–6.

[25] K. Smiljkovikj et al., “Capacity Analysis of Decoupled Downlink and Uplink Access in 5G Heterogeneous Systems,” ArXiv14107270 Cs Math, Oct. 2014.

[26] K. Smiljkovikj, P. Popovski, and L. Gavrilovska, “Analysis of the Decoupled Access for Downlink and Uplink in Wireless Heterogeneous Networks,” *IEEE Wirel. Commun. Lett.*, vol. 4, no. 2, pp. 173–176, Apr. 2015.

[27] J. Acharya, L. Gao and S. Gaur. *Heterogeneous Networks in LTE-Advanced*, 1st ed., John Wiley & Sons, Ltd. West Sussex, UK. 2014.

[28] R. Qingyang Hu and Y. Qian. *Heterogeneous Cellular Networks*, 1st ed., John Wiley & Sons, Ltd. West Sussex, UK. 2013.

[29] O. Afolalu and N. Ventura, “Dynamic Cell Selection Technique for Interference Coordination in 5G Heterogeneous Networks,” *in Southern Africa Telecommunication Networks and Application Conference (SATNAC)*, 2016, pp. 4-8.

[30] Y. Chen, Z. Yang, and H. Zhang, “Opportunistic-based dynamic interference coordination in dense small cells deployment” *in 2017 IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, 2017, pp. 1-5.

[31] T.E. Bogale and L.B. Le, “Massive MIMO and mmWave for 5G Wireless HetNet: Potential Benefits and Challenges” *in IEEE Vehicular Technology Magazine*, 2016, pp. 64-75.

[32] C. Perfecto, D. S. Javier and B. Mehdi, “Millimeter Wave V2V Communications: Distributed Association and Beam Alignment”, *in IEEE Vehicular Technology Magazine*, 2016, pp. 40-55.

[33] S. Yujae, K. Peng-Yong and H, Youngnam, “Minimizing Energy Consumption Through Traffic Offloading in HetNets with Two-Class Traffic”, *IEEE*

Communications Letters, 2015, vol. 19, issue 8, pp. 1394-1397.

[34] D. Boya, S. Lingyang, L. Yonghui and L. Geoffrey Ye, “NOMA-Based Low-Latency and High-Reliable Broadcast Communications for 5G V2X Services”, *in IEEE Global Communications Conference (GLOBECOM),* 2017, pp. 1-6.

[35] D. Boya, S. Lingyang, L. Yonghui and L. Geoffrey Ye, “Non-Orthogonal Multiple Access for High-Reliable and Low-Latency V2X Communications in 5G Systems”, *IEEE Journal on Selected Areas in Communications*, 2017, vol. 35, issue 10, pp. 2383-2397.

[36] L. Liu, Y. Chang, S. Song and D. Yang, “On adaptive range expansion for UL-DL imbalance mitigation in HSPA HetNet”, *in IEEE Wireless Communications and Networking Conference (WCNC)”*, 2015, pp. 1440-1445.