An Overview of Mobility Management Mechanisms and the Related Challenges in 5G Networks and Beyond

Mustafa Mohammed Hasan Alkalsh[1,2]

[1] *Institute of Radiocommunications, Poznan University of Technology, Poznan, Poland,*
[2] *Department of Mobile Networks, Nokia Solutions and Networks, Wroclaw, Poland*

https://doi.org/10.26636/jtit.2023.171723

Abstract — Ensuring a seamless connection with various types of mobile user equipment (UE) items is one of the more significant challenges facing different generations of wireless systems. However, enabling the high-band spectrum – such as the millimeter wave (mmWave) band – is also one of the important factors of 5G networks, as it enables them to deal with increasing demand and ensures high coverage. Therefore, the deployment of new (small) cells with a short range and operating within the mmWave band is required in order to assist the macro cells which are responsible for operating long-range radio connections. The deployment of small cells results in a new network structure, known as heterogeneous networks (HetNets). As a result, the number of passthrough cells using the handover (HO) process will be dramatically increased. Mobility management (MM) in such a massive network will become crucial, especially when it comes to mobile users traveling at very high speeds. Current MM solutions will be ineffective, as they will not be able to provide the required reliability, flexibility, and scalability. Thus, smart algorithms and techniques are required in future networks. Also, machine learning (ML) techniques are perfectly capable of supporting the latest 5G technologies that are expected to deliver high data rates to upcoming use cases and services, such as massive machine type communications (mMTC), enhanced mobile broadband (eMBB), and ultra-reliable low latency communications (uRLLC). This paper aims to review the MM approaches used in 5G HetNets and describes the deployment of AI mechanisms and techniques in "connected mode" MM schemes. Furthermore, this paper addresses the related challenges and suggests potential solutions for 5G networks and beyond.

Keywords — B5G, beam switching, handover, HetNet, mobility management.

1. Introduction

The massive growth in the number of connected devices, services, and the advent of advanced applications has caused a significant increase in data volumes. Billions of new user equipment (UE) items, such as vehicles, drones, smart sensors, medical devices/applications, and home appliances, constitute what is nowadays known as the Internet of Things (IoT). All these devices and the related services require a continuous, seamless, and reliable connection. Consequently, in the coming years, demand for mobile data will increase many thousand times. The fifth generation (5G) new radio (NR) network standard is designed to support high data rates while offering extremely low latency (suitable for real-time applications), ensuring efficient mobility management (MM) of the connected devices and guaranteeing higher energy efficiency. It will provide a 10 Gbps peak data rate and per-user data rates that will be 10 Mbps higher than those available in current wireless networks [1].

Furthermore, it will increase system capacity and provide wide-scale connectivity by utilizing new frequency bands, such as the millimeter wave (mmWave) spectrum. In order to deal with the demands posed by increased network capacity and to provide high coverage to users, HetNets will be gaining in importance. The mmWave band is considered to be the most suitable and functional solution for 5G mobile networks and beyond [2], [3]. This leads to the deployment of small cells which are best suited to operate within the mmWave band due to their short-range characteristics. They will contribute mainly to higher throughput, increased energy efficiency, reduced coverage losses in indoor and outdoor environments, and to improved quality of service (QoS).

In wireless networks, the MM entity is primarily required to ensure an appropriate hand-over (HO) process to switch the connection between different sources without any interruptions [4], [5]. For example, network coverage should be maintained to ensure a seamless connection with UE items traveling in high-speed trains [6]. Wireless communication solutions, such as 5G networks, are capable of providing continuous access to specific services. Therefore, when a UE item is connected, e.g. during a call or data exchange session, the base station (BS), known as eNodeB (eNB) in LTE and gNodeB (gNB) in 5G, applies an active signaling mechanism through its control plane (CP) to monitor the connection, keep it stable, ensure its good quality and decide the time at which the HO is requested to switch to another gNB without interruption [6], [7].

Even though 5G will offer numerous features and solutions capable of satisfying the demand for increased system capacity, enhancing the user experience and providing high data rates, a number of MM-related issues have emerged as well. The anticipated increase in complexity of future wireless networks is a result of several factors that have contributed to challenges in the field of mobility management. These factors include mass-scale deployment of small cells – a phenomenon that constitutes a response to the need for more coverage in areas with limited-service availability. This deployment is expected to lead to an increase in handover probability. Additionally, the implementation of carrier aggregation (CA) and dual connectivity (DC) technologies will also create a wide range of handover scenarios that need to be addressed [8]. Consequently, ensuring a seamless connection with a number of mobile users will become one of the significant challenges faced while implementing HetNets and ultra-dense HetNets in 5G networks and beyond. These will significantly reduce HO performance if current MM mechanisms are retained. Thus, this paper aims to provide an overview of the concept of MM and discusses the enhancement of HO schemes in future wireless networks.

The rest of the paper is organized as follows. Section 2 describes related works, while Section 3 provides an overview of 5G networks, focusing on such areas as mmWaves, small cells, 5G HetNets, and ML. Section 4 discusses the MM mechanisms, presenting also the background of 5G HetNets and MM-related functional requirements for 5G networks and beyond. Section 5 presents the HO concepts, different types of procedures relied upon and identifies the specific performance parameters. Section 6 details MM-related challenges encountered in 5G HetNets. Section 7 identifies future directions for managing HO in 5G HetNets, while Section 8 concludes the paper.

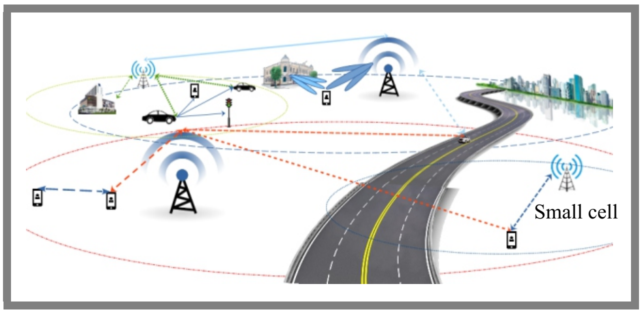

Fig. 1. 5G network architecture.

Figure 1 shows a general architecture of a 5G cellular network and presents different services that are supported in different scenarios and environments.

2. Related Works

Comprehensive studies have been performed in the literature to cope with management-related problems affecting HetNets, where the mobility of users is the major reason behind performing HO processes and switching the users from one cell to another. In [9], a prediction method was proposed for calculating user velocity based on the number of HO processes occurring in HetNets, by using the minimum variance-neutral technique. In [10], a mitigation algorithm for frequent HO has been investigated to reduce the number of HO processes and improve network efficiency in ultra-dense HetNets by classifying users with frequent HO processes as high mobility users and ping-pong (PP) users. In this algorithm, high mobility users are connected to macro cells. Simultaneously, this approach optimizes handover (HO) parameters for PP users to reduce the number of unnecessary handovers. If the adjustment of HO parameters does not prevent PP effects, these users are also connected to macro cells.

In [11], the authors attempted to optimize cross-layer HO processes in terms of delay encountered in HetNets. In paper [12], the authors proposed a framework for achieving seamless HO between neighboring macro cells in 5G network, with the proposed solution allowing for continuous switching by integrating the overlap area of an HO-assisted micro cell with a DC connection. The authors of [13] tried to identify the weight of HO metrics by proposing a method that uses the analytical hierarchy processing (AHP) technique and then sorts the cells for the purpose of selecting the best target cell for HO relying on the grey rational analysis (GRA) method. The results showed a reduction in the HO rate and a lower radio link failure (RLF) rate. In article [14], the impact that channel fading exerts on mobility management in HetNets has been examined. The results showed that by increasing the sampling period of the HO decision, the fading effect decreased, while the PP effect increased.

The work described in [15] showed the sensitive user mobility association rule. The rule attempts to overcome the problem of crowded areas by directing UEs to small cells and monitoring dynamic shifts in channel conditions caused by the continued mobility of users within the network's topology. Consequently, it prevents frequent HO and PP between small cells. However, specific aspects of mmWave communication are taken into consideration, e.g. directionality, sensitivity to blocking, non-line-of-sight propagation effects, and distributing the UEs within the network accordingly. Due to its small wave-length, the mmWave band may be blocked by various objects. To overcome this issue, a framework is provided in [16] that may predict the data rate degradation caused by obstacles before the degradation occurs, by expanding the status area based on sequential camera images, and by using deep reinforcement learning to decide HO timings and overcome issues encountered while addressing the problem of large dimensions.

3. Background

MmWave bands cover high frequencies between 30 and 300 GHz within the spectrum of 5G wireless networks. Thanks to their significant bandwidth available, these bands may be used to solve the primary concern of 5G networks, namely the demand for a higher network capacity [17]. However, mmWave bands are more vulnerable to weather conditions and propagation blockages caused by building walls and vegetation.

3.1. Small Cells

Nowadays, providing full coverage, especially in crowded cities with numerous high-rise buildings, is a big challenge. The vast increase in data volumes is another challenge faced

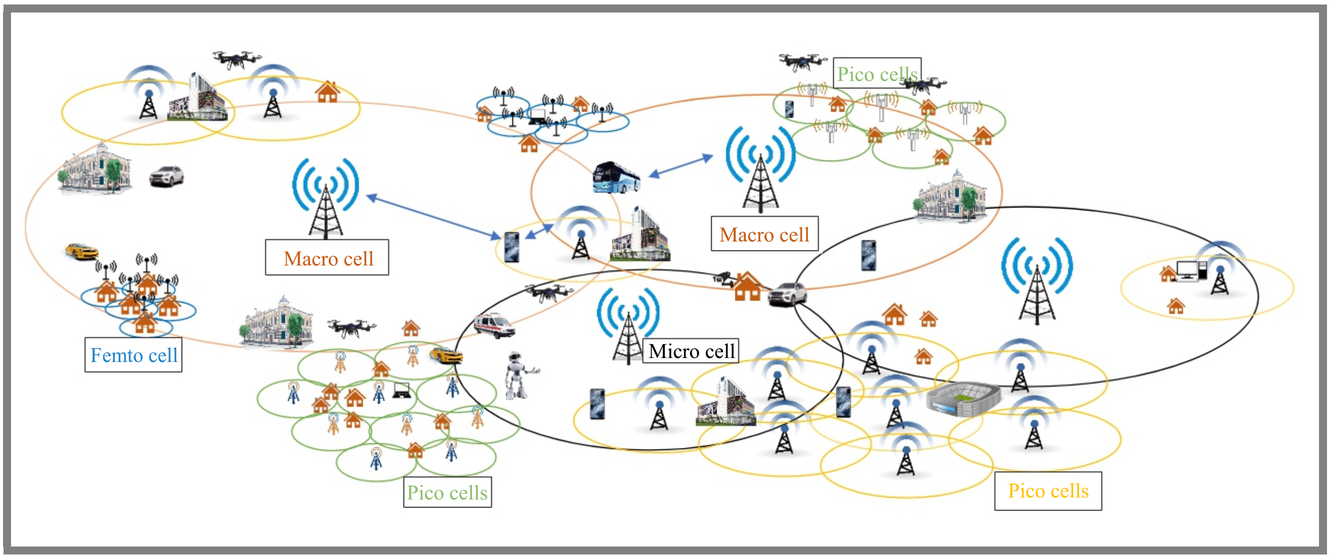

Fig. 2. Cell deployment scenarios in 5G networks.

by wireless networks. To cope with those obstacles, the size of macro cells must be reduced [18]. This approach is inefficient from the cost perspective. The introduction of mmWave for 5G networks and beyond has resulted in the fact that small cells have become a preferred solution due to the high path loss characteristics of these bands. These cells are of the low power variety and have the form of simple plug-in access nodes. They are designed to increase capacity, extend coverage, and enhance spectral efficiency of wireless networks [19].

Many small cell types may be deployed within a single macro cell. Small cells can be situated either inside or outside of buildings. They can cover a radius reaching from 10 to many hundreds of meters in order to ensure adequate coverage for both indoor and outdoor subscribers, as shown in Fig. 2. Small cells are also used as load-balancing radio access networks (RAN) to reduce the load in macro cells and improve QoS. The following small cell types may be distinguished:

- microcells cover an area of several kilometers [20]. Microcells are deployed at remote locations, such as small villages and suburban areas,
- picocells cover the area of approx. 20 to 200 m [21]. Picocells are deployed in congested places or at special occasions, such as stadiums during events, big shopping malls hosting seasonal parties, and other large gatherings,
- femtocells cover areas of approx. 10 to 30 m and are used in offices, homes or subways [22]. They are useful in scenarios where a specific user requires a high data rate and the available network offers insufficient throughput.

3.2. 5G Hetnets

Improving the coverage of wireless networks is one of the important requirements. Hence, network expansion and densification is recognized as one of the most effective mechanisms. To provide coverage to users in offices, crowded areas or at homes, small cells can be deployed within a given macro cell. The deployment of small cells within a macro cell is known as a HetNet [23]. HetNets are capable of improving the overall performance of a wireless network. HetNet deployments may be classified into four different categories:

- overlapping/non-overlapping,
- intra/inter frequency,
- sparse/dense,
- indoor/outdoor.

In the first category, the macro cell and the low power node offer overlapping coverages. The low power node operates in the macro cell's dead area to enhance coverage. If the coverage of both components overlaps, low power nodes may be used for load balancing and for improving capacity. In the second category, low power nodes and macro cells may operate using the same carrier frequency (intra-cell frequency) or different frequencies (inter-cell frequency).

In the third category, low power nodes can be used to provide service in small areas with high user density levels. Such scenarios are known as sparse deployment cases. In dense deployment configurations, however, low-power nodes can be deployed to enhance the overall capacity over a wider area. Finally, in the fourth category, low-power nodes may be applied, in a scalable manner, to secure indoor and outdoor data transmissions. For instance, femtocells are used for indoor applications.

3.3. Machine Learning (ML)

Machine learning is an innovative idea where an algorithm is employed to enable the system to self-learn. This approach may be implemented in 5G HetNets to enhance their overall performance [24]. Nowadays, many novel machine learning models are available, with three key approaches identified below: unsupervised learning, supervised learning, and reinforcement learning. Unsupervised learning algorithms classify data in the unlabeled form. According to that, the machine model is in charge of clustering data points in terms of simi-

Tab. 1. MM-related functional requirements for 5G networks and beyond.

5G and beyond scenario	Resulting MM functional requirement
eMBB, mMTC and URLLC applications that require different QoS, i.e. minimum data rate, latency, reliability, etc. [27]	Support based on context, i.e., UE mobility, network conditions, application requirements etc.
UEs density $\geqslant 10^6$ devices per km^2 [28]	Provide adequate support effectively as long as the density of users continues to increase
Multiple RATs connections	Support of multi-RAT MM in addition to the efficient RAT selection methods
Drones as APs and relay stations [29]	Support mobility of both UEs and APs in a 3D form
High speed mobility – up to and 500 kph and beyond [28]	Ensure the flexibility required to deal with mobility profile demands and to avoid the one-size-fits-all approach
Utilizing sub-6 GHz, millimeter wave (mmWave), terahertz communication [30], [31]	Increased robustness for seamless mobility, bearing in mind that mmWave will be significantly influenced by what is around, i.e. buildings, forests, etc.
Network softwarization [32]	Evolve to utilize the benefits provided by software enablers, such as SDN, network function virtualization (NFV), etc.
Granularity of tracking and localization on the beam level [33]	Utilize the advanced level of granularity to ensure better MM and tracking performance in high-speed scenarios or in dense urban networks

larities between each other. Supervised learning, on the other hand, is a method that requires supervision to determine specific relations between the input and the output. By definition, supervised learning algorithms classify data in the labelled form and feed the model which includes both input and output data. Generally, supervised machine learning techniques are utilized for classification and regression, i.e processes relied upon for predicting continuous and separated values, respectively.

In the third ML model – reinforcement learning (RL) – an agent learns to make decisions by interacting with its environment and by receiving feedback in the form of rewards or punishments for the decisions made, thus adjusting its behavior to maximize long-term rewards. The key components of RL include a policy (used for mapping states to actions), a value function (an estimate of the expected reward for being in a specific state and following a specific policy), and a learning algorithm which updates the policy and/or value function based on feedback from the environment.

4. Mobility Management in 5G Hetnets

The MM feature is considered to be one of the fundamental aspects of wireless communication networks. MM enables wireless networks to perform various functions, including locating a UE in the idle state and determining which cell should be used for delivering data packets, i.e. tracking areas and managing locations. Additionally, MM is used for maintaining the connection with UE during its movement within the network. MM switches the UE connection from the serving cell (source gNB) to a new cell (target gNB), provided that coverage is available during its movement. For instance, when a UE is in a connected state (a call or a data session are in progress), the data flowing over the user plane (UP) can be transferred from the source gNB to the target gNB.

MM enhances the UE's experience and renders the network services available for many purposes at any time. Future wireless networks will provide services for several industry fields. They will also be dense, heterogeneous and extensively programmable, while relying on the same network infrastructure [25]. These changes represent a paradigm shift in the network architecture design [26]. Consequently, MM mechanisms need to be re-evaluated and/or re-designed. In light of the above, this section presents, in Tab. 1, the functional MM-related requirements for future wireless networks, based on specific network scenarios.

One may observe from Tab. 1 that MM-related mechanisms for 5G networks and beyond should be adapted and improved in order to be able to manage future wireless networks efficiently and to be flexible, scalable, and reliable to ensure the required QoS and seamless mobility. Apart from these requirements, there are specific criteria that can impact the development of future MM solutions. These include the following:

- User context. This category includes several parameters (i.e., user mobility profiles, flow types, network/user policies, signal quality) that allow the MM mechanisms of future networks to serve users with different mobility profiles and accessing different services. For instance, in [34], network load-aware MM methods show a throughput enhancement of 75% at the cell edge, compared to the context-agnostic methods. This means that the availability of contextual information boosts the performance of the network in terms of the previously mentioned criteria.

- D2D service availability. D2D services may determine when the mobility management mechanism is executed. D2D may be utilized to provide seamless mobility by forwarding CP information. This may be relevant for V2X scenarios. For instance, vehicles outside the coverage area or experiencing a deep fade with the infrastructure network could transfer data, over the PC5 interface, through other vehicles which are close by and are simultaneously within the coverage area, or are experiencing better channel conditions when communicating with the network's infrastructure.
- Physical layer considerations. The use of massive MIMO and mmWaves would certainly influence the MM methods. In urban environments, the mmWave band will suffer from significant interception in addition to its limited range caused by peculiar propagation characteristics. Hence, densification will be required which, in return, calls for frequent HO. Moreover, the use of beamforming with massive MIMO antennas may track the mobility of UEs, allowing to provide them with high QoS through a higher throughput.
- Control plane signaling. Reduction of CP signaling that occurs during the HO phase is an important target of future MM mechanisms. Such a process will positively affect QoS experienced during the HO process [35].

5. Handover Concepts

Handover is one of the most significant functionalities of MM. It refers to the ability to maintain the connection of a UE that is in a connected state with the network during its mobility, ensuring a seamless transition between different cells while minimizing service disruption. The fact that UE remains mobile changes the strength of the received signal, which may cause the connection to be unstable. To maintain connectivity, the UE is required to switch between different cells providing a higher signal strength as it moves within the network. This process is called a handover (HO) and allows the UE to stay connected to the network, even as it crosses cell boundaries. This process is also used for other purposes, such as steering UE traffic to better frequencies with better throughput, load balancing within the cell, etc.

Two parameters are used commonly to indicate the power of the received signals: reference signal received power (RSRP), and reference signal received quality (RSRQ). RSRP identifies the average power of the received signal without considering interference and noise, while RSRQ is a more complicated metric that identifies the average power of the received signal including interference and noise components.

The decision to trigger the HO process is taken according to the related changes in the strength values of the received signal reported by the UE for the current and neighboring cells. The process is performed in the following manner.

5.1. Handover Procedure

As soon as the strength of the received signal falls, at any location within the network, below the threshold level, the HO procedure is triggered. A measurement report includes the RSRP or RSRQ values of the serving and neighbor cells, submitted by the UE to the serving cell. There are several conditions and events in trigger the transmission of MRs such as A3 event to trigger HO process. The process is initiated when the power of the signal received from the serving cell becomes lower than the threshold value, and the received power of a neighbor cell becomes higher than the threshold value [36]. The HO process is executed once the serving cell receives the MR that identifies the neighboring cells which meet the applicable conditions. The serving cell selects the neighboring cell that meets the criteria, identifies it as a target cell to and sends the HO request message in order to transfer the UE's connection to the target cell.

The HO process may be divided into four phases, although they may vary across different technologies: measurement, preparation, execution, and completion stages may be distinguished [37], [38]. During the first phase of the HO process, when the UE reaches threshold values triggering a specific event, it reports the source and neighboring cells to the source cell through the CP. It is important to note that this step may vary depending on the technology used.

If the HO condition is met, the second phase is initiated, where the source cell sends an HO request message to the target cell. In response, the target cell sends an acknowledgment message to the source cell, including the RRC message. Once the source cell receives the acknowledgment message, the third phase begins. In this phase, the source cell forwards the RRC message to the UE through the HO command message. When the UE is attached to the target cell, the synchronization, authentication, and network configuration process is finalized, marking the end of the fourth phase.

Figure 3 shows the A3 condition triggering the HO process to the target cell at UE as:

$$M_{Source} - M_{Target} > O \;, \tag{1}$$

where M_{Source} and M_{Target} are the RSRP levels of the serving cell and the target cell, measured at UE, respectively. O is a cell-specific parameter, known as the HO margin. The UE sends an MR message to the serving cell when Eq. (1) remains satisfied for the period of time known as time-to-trigger (TTT). When the serving cell receives the MR from the UE, it starts the HO process with the target cell.

Several classifications of the HO process, based on different perspectives, are available. These are presented below [24].

5.2. Handover Classifications

The HO procedure may be classified based on the purpose for which it is triggered, e.g., based on load-balancing, QoS and coverage. In the load-balancing scenario, the HO process is triggered by the network for UEs located at the edge of the gNB, in order to prevent gNB from being overloaded when

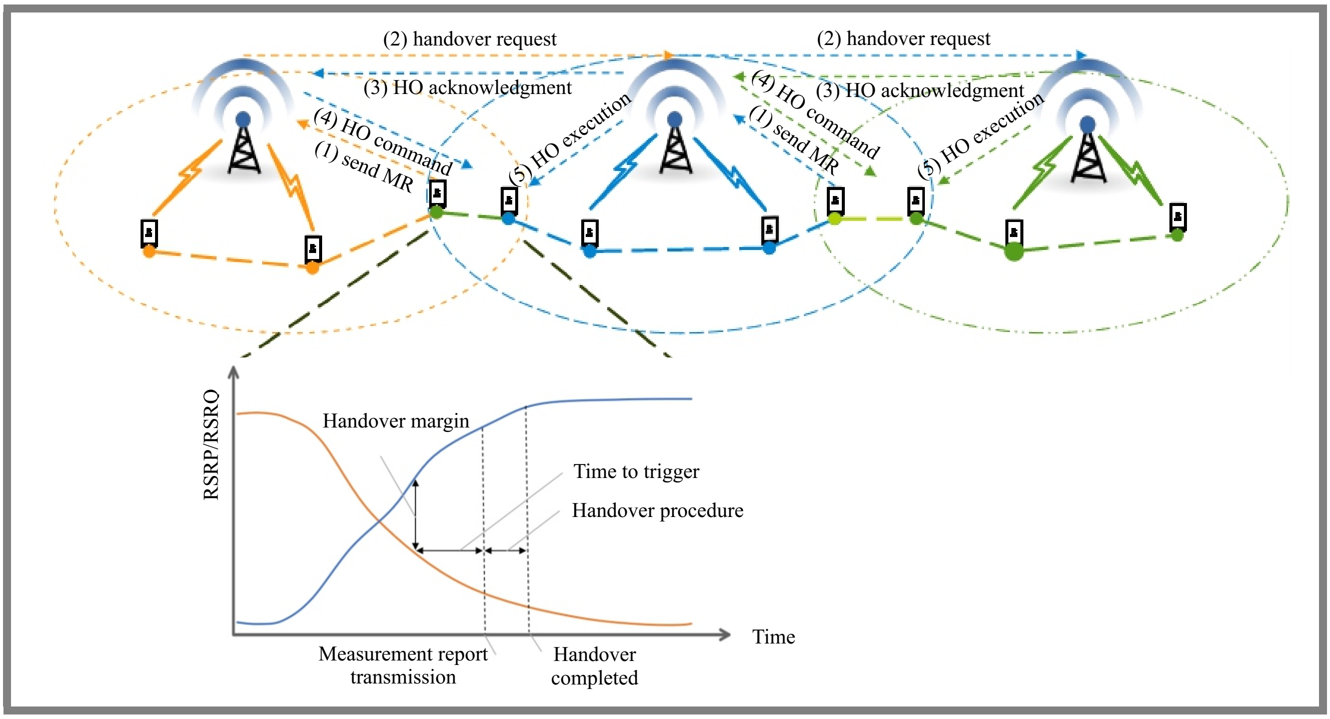

Fig. 3. Handover procedure diagram.

the covered area is crowded. For the QoS scenario, the HO process is triggered if any neighboring gNBs provide better signal quality than the current one. In the coverage-based case, HO is triggered to reduce the connection loss rate for UEs located in areas in which the serving gNB cannot provide sufficient coverage.

Network densification is one of the most popular solutions relied upon to manage the vast increase in data traffic in future wireless networks, as it is capable of providing excellent coverage. The idea behind this specific approach is to create HetNets by deploying small cells, such as microcells, picocells, and femtocells. The HO process would be much more complicated in a HetNet than in a macro cell.

Two classifications of HO processes are known: they may be categorized as being of the inter- and intra-cell layer variety. An intra-cell layer HO occurs between macro-macro/or micro-micro cells etc. An inter-cell layer HO occurs, in turn, when the quality of a pico cell signal becomes better than the threshold value. In such a case, the network may can detach the UE from the macro cell and attach it to the picocell.

The remaining HO process classifications divide them into those relying on inter and intra radio access technology (RAT) HO. Intra-RAT HO occurs in the same RAT, while Inter-RAT HO occurs between different RATs, for instance, when the UE access network changes from 4G to 3G or to edge when the user is traveling on a fast train or is passing through a forest.

The last classification of HO processes distinguishes soft and hard HO variations. The soft HO procedure allows UE to connect to source and target gNBs simultaneously, in order to prevent connection interruptions. In the hard HO procedure, in turn, UE may connect to a single gNB during the HO process only. Therefore, connection interruptions take place during the HO process.

HO processes may be evaluated based on their performance parameters which are described below.

5.3. Handover Performance Parameters

There are specific performance indicators that exert a significant impact on network parameters and are employed to measure HO performance of wireless networks. They are listed below:

- HO rate, outlining the number of HO processes occurring per second,
- HOF rate, calculated by dividing the number of HOF by the number of HO processes,
- number of PP events – PP is the second HO to the source gNB after the first HO from the source gNB to the target gNB,
- HO interruption time, being the period of time during which UE could not receive any data packets from source and target gNBs during the HO process,
- HO latency, being the duration of the HO execution phase.

HO processes involve switching a user's connection from one cell to another, and a failure that occurs during this process is called a handover failure (HOF). One of the most common causes of HOF is a radio link failure (RLF) which can be attributed to three key causes described below [19]:

- **Too-early HO**. It means that the HO process starts before the signal strength of the target gNB is good enough to attach the UE, thus an RLF affecting the connection with the target gNB takes place. A radio link reestablishment (RLR) with the source gNB takes place.

- **Too-late HO**. It means the HO process starts and the RLF toward the source gNB occurs due to a reduction in channel quality of the source gNB, a RLR with the target gNB taken place.
- **HO to a wrong cell**. This means that the HO process starts toward a specific gNB and, at the same time, the direction in which UE is traveling changes and it enters the area covered by another gNB. The UE may also be moving along the edge of two neighboring gNBs, with their signals overlapping. An RLR toward the new gNB occurs and an RLF message is sent from the connected gNB to inform the target gNB.

6. MM Challenges in 5G Hetnets

Good understanding of the challenges/issues related to MM mechanisms is of key significance for this review. For instance, such HO-related issues as frequent HOs, HOF, RLF, and PP effects, will remain a key challenge in future network environments. Meanwhile, current MM mechanisms are unable to ensure a seamless and effective HO between cellular and non-3GPP networks, which exacerbates the existing challenges. HO signaling is yet another challenge, as a reduction in HO signaling ensures good scalability and reliability of the network. Additionally, an introduction of network slices to ensure that different services are rendered in accordance with their own resource-related demands will create even more challenges concerning MM design strategies providing joint solutions for multiple network slices or individual solutions for each network slice.

Furthermore, context-based MM factors, such as network load, user preference, network policy, mobility profiles, also need to be taken into consideration to ensure the best possible provision of the requested QoS, challenging future MM mechanisms. The criticality of this challenge is consolidated by the fact that low computational complexity of these solutions will be of the essence in order to meet strict latency-related requirements. On the other hand, several factors and technologies that have already been implemented create severe challenges for mobility management mechanisms used in 5G networks.

6.1. MmWave and Small Cell Deployment

The deployment of HetNets is a solution that is capable of satisfying the demand of 5G networks for higher coverage and capacity. HetNets, made up of low power nodes, such as micro cells, femtocells, picocells, or even relays, improve the functionality of macro cell networks. This results from the introduction of mmWave bands that provide short coverage due to their path loss characterization, where the path loss increases as the frequency band used becomes higher [39]. For instance, 28 GHz is one of the best frequency bands capable of providing coverage of up to 200 m [40]. However, to provide 5G coverage over a distance of several kilometers, several small 5G cells must be deployed. Thus, the number of small cells will continue to increase with the use of higher mmWave bands, due to the decrease in cell coverage. On the other hand, a single 4G cell may cover an area within the radius of up to 1.5 km [41]. Replacement of one 2.1 GHz 4G macro cell will require approximately fourteen 28 GHz 5G small cells to achieve the same coverage as ensured by one 4G cell.

Although HetNets are capable of satisfying capacity- and coverage-related demands, they also create a number of other issues and challenges related to MM mechanisms, with a particular emphasis placed on the HO process. The management of HO parameters is getting much more challenging with 5G cells characterized by small coverage areas and more complex radio transmission environments. Therefore, large-scale deployments of small cells increase the probability of triggering a HO process when UE is moving. This, in turn, leads to an increase in the probability of PP events, RLF, interruption time, and throughput regression. The increase in the number of HOs required, caused by the smaller cell size, boosts the number of HOFs expected in the network as well.

This poses a significant challenge in high-speed mobility scenarios in which UEs travel across a single cell over a period of several seconds, limiting the time needed to send the MR and complete the HO procedure. Therefore, the introduction of mmWaves and HetNets/ultra-dense HetNets will be one of the more significant challenges for MM mechanisms.

6.2. Dual Connectivity (DC)

DC is one of the important technologies that allows UE to connect to two different cells at the same time [42]. It was introduced in Release 12 of the 3rd Generation Partnership Project (3GPP) standard. The first connection is established with a macro cell, while the other with a small cell [43]. UE may communicate over 4G and 5G networks simultaneously, e.g., when the NSA mode is used, the 4G macro cell is considered the primary node, while the 5G small cell is considered the secondary node. Such an approach allows to enhance the data rate and supports UE mobility. The number of HOs will increase if UE is mobile within the network, since UE may be simultaneously connected to 4G and 5G networks over varying frequency bands. The increase in the number of HO scenarios will affect HO probability compared to a single connection.

The new HO scenarios occur in two cases. The first case is when UE switches the connection between a macro cell and another macro cell at the primary node level. The second case occurs when UE switches the connection between a small cell and another small cell at the secondary node level. Several factors need to be considered in addition to the normal aspects taken into account in MM mechanisms in order to maintain the connection with a mobile UE, as HO probability will be increased and additional mobility-related challenges will be experienced, including those concerned with interruption time, UE power consumption, and RRC signaling to monitor both connections by the network.

6.3. Carrier Aggregation (CA)

The CA technique was introduced in LTE-A systems from Release 10 of the 3GPP standard. The basic idea behind CA consists in aggregating multiple carrier components (CC) to serve a given piece of UE. It allows UE to support, simultaneously, multiple connections with the serving cell, using different frequency bands, in order to achieve a higher data rate over a wider bandwidth. The CA approach enhances wireless connectivity by providing better coverage. One of the assigned carriers is configured as the primary carrier component (PCC), used to transfer both control and user data flows between the UE and the serving network. The other assigned carrier is configured as the secondary carrier component (SCC) and is used to extend the UE's bandwidth, and deliver user data flow only.

The aim of enhancing the CA technique by relying on different carrier aggregation deployment scenarios (CADSs) is to increase the total performance of the network by increasing network coverage, offering extended bandwidth to UEs, and improving the overall UE experience. However, the addition of various scenarios to this technique has expanded the range of mobility-related challenges.

A configuration in which multiple CCs serve a single piece of UE introduces new HO scenarios. For example, an HO scenario between CCs could be defined as a CC-to-CC HO. Such a situation exists when the system needs to change the PCC, selected as the best among multiple configured CCs. This new HO scenario focuses on switching the PCC, which is assigned based on its channel conditions and quality of signal to the UE. The other HO scenario, in turn, materializes when an inter-cell HO occurs. As UE allows the use of multiple CCs, the HO is needed for switching connections of both PCC and SCC to the new cells that support the CA technique. These HO scenarios increase the HO probability, leading to an additional increase in the number of MM-related challenges that need to be taken into consideration.

6.4. Emergence of UAV Communications

Connected drones are expected to be used in 5G wireless networks and beyond. Currently, the target scenarios involve using connected UAVs serving as sky base stations or acting as mobile UEs when deployed in order to perform other roles. Such an approach is expected to enhance data rates of UE, thus providing better wireless services in remote areas [17], [44].

However, connected drones require more stable communications. The fact that drones or other UAVs are mobile in three dimensions is a big challenge that leads to fast variations in the received signal strength [17], [45]. Moreover, drones move at faster speeds and their paths vary from those taken by vehicles or regular UE, which results in a considerable deterioration in received signal strength. This significantly contributes to increasing HO probability. The time required for the HO process to switch connections to a target cell may cause some interruptions in calls before the drones switch their connections. Therefore, the interruption in such a scenario will be increased beyond the value that is typical of UE.

7. Future Outlook

The expected complexity of future wireless networks results from the need to adapt to many different use cases by introducing new communication technologies that are capable of meeting growing data traffic demands in order to accommodate the rapid increase in the number of devices and services that require an Internet connection. Furthermore, the diverse range of connection-related requirements (reliability, latency, minimum data rate, high mobility speed, etc.) is playing an important role as well and creates numerous challenges for future MM mechanisms.

The need for intelligent MM solutions is increasing due to various user scenarios, such as high-speed mobility at different altitudes and various densities. These solutions should address the challenges associated with the available technologies, such as DC or CA, in order to improve MM mechanisms. Proper solutions must be identified to exploit the advantages of these technologies. For instance, utilizing context awareness techniques will ensure the consideration of user, network, and application context while selecting the appropriate MM solution. On the other hand, network slicing, which is an on-demand strategy, could be utilized to assist networks in dedicating specific slices to dealing with specific scenarios in order to create MM solutions that supplement the usual scenarios. Classifications of potential scenarios may be based on their expected materialization times, estimated by considering a history file stored in a module located somewhere within the network (e.g. many high-speed trains with high user density, passing specific areas at the same time, when the network is facing high traffic loads at the beginning of a work day).

Furthermore, frequent learning of network conditions, mobility profiles, and their corresponding impact on HO processes is a complicated task. Many machine learning techniques can help predict/estimate valuable system parameters (such as RSRP, RSRQ, user location, user speed, used direction and load condition) to avoid frequent HOs as well as PP, HOF, and RLF conditions by selecting an appropriate HO method and time to trigger the HO process. The ability to understand HO complexities will render ML methodologies substantial in determining the best association in an increasingly dynamic and multi-dimensional network.

Edge computing platforms can assist in faster and more effective handover decisions, given their ability to provision computation power closer to the access network. D2D networks may also assist with extended coverage and, hence, smoother handovers. Finally, the higher the number of potential scenarios in future networks, the higher the number of queries concerning MM mechanisms that can be actively verified by research and industrial communities.

8. Conclusion

Mobility management remains a critical aspect of wireless communication networks, especially in the 5G and beyond context. However, MM in such networks presents a unique set of challenges, such as those related to the mmWave band, deployment of small cells, dual connectivity, carrier aggregation, and the emerging field of UAV-based communications. Therefore, further studies and research are required to address these challenges and to ensure that MM continues to be an effective and efficient process in the dynamic environment of wireless communication networks. By doing so, we can ensure that users can experience seamless connectivity and high-quality services, while network operators can continue to provide reliable and efficient solutions for their customers, enabling new use cases.

Acknowledgment

The author would like to acknowledge the financial support provided by the MEiN Program "Doktorat Wdrożeniowy" (Agreement no. DWD/6/0057/2022, date: 23.12.2022) administered by the Ministry of Higher Education, Republic of Poland. The author would also like to acknowledge the affiliation with Nokia Solutions and Networks in relation to the implementation of the cooperation agreement for the "Doktorat Wdrożeniowy" program.

References

[1] I. Chih-Lin *et al*., "New paradigm of 5G wireless Internet", *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 3, pp. 474–482, 2016 (`https://doi.org/10.1109/JSAC.2016.2525739`).

[2] W. Saad, M. Bennis, and M. Chen, "A vision of 6G wireless systems: Applications, trends, technologies, and open research problems", 2019, arXiv:1902.10265 [Online]. Available: `http://arxiv.org/abs/1902.10265`

[3] F. Tariq *et al*. "A speculative study on 6G", 2019, arXiv:1902.06700 [Online]. Available: http://arxiv.org/abs/1902.06700

[4] A.C. Morales, A. Aijaz, and T. Mahmoodi, "Taming mobility management functions in 5G: Handover functionality as a service (FaaS)", in *Proceedings of the 2015 IEEE Globecom Workshops*, San Diego, USA, pp. 1–4, 2015 (`https://doi.org/10.1109/GLOCOMW.2015.7414151`).

[5] V. Yajnanarayana, H. Ryden, and L. Hevizi, "5G handover using reinforcement learning", in *Proceedings of the 2020 IEEE 3rd 5G World Forum* (*5GWF*), India, pp. 349–354, 2020 (`https://doi.org/10.48550/arXiv.1904.02572`).

[6] Y. Li et al., "Beyond 5G: reliable extreme mobility management", in *Proceedings of the Annual conference of the ACM Special Interest Group on Data Communication on the Applications, Technologies, Architectures, and Protocols for Computer Communication*, pp. 344–358, 2020 (`https://dl.acm.org/doi/proceedings/10.1145/3387514`).

[7] E. Gures *et al*., "A comprehensive survey on mobility management in 5G heterogeneous networks: Architectures, challenges and solutions", *IEEE Access*, vol. 8, pp. 195883–195913, 2020 (`https://doi.org/10.1109/ACCESS.2020.3030762`).

[8] I. Shayea *et al*. "Key challenges, drivers and solutions for mobility management in 5G networks: A survey", *IEEE Access*, vol. 8, pp. 172534–172552, 2020 (`https://doi.org/10.1109/ACCESS.2020.3023802`).

[9] R. Tiwari and S. Deshmukh, "MVU estimate of user velocity via gamma distributed handover count in HetNets", *IEEE Communication Letters*, vol. 23, no. 3, pp. 482–485, 2019 (`https://doi.org/10.1109/LCOMM.2019.2892962`).

[10] M.M. Hasan, S. Kwon, and S. Oh, "Frequent-handover mitigation in ultra-dense heterogeneous networks", *IEEE Transactions on Vehicular Technology*, vol. 68, no. 1, pp. 1035–1040, 2019 (`https://doi.org/10.1109/TVT.2018.2874692`).

[11] X. Xu, X. Tang, Z. Sun, X. Tao, and P. Zhang, "Delay-oriented cross-tier handover optimization in ultra-dense heterogeneous networks", *IEEE Access*, vol. 7, pp. 21769–21776, 2019 (`https://doi.org/10.1109/ACCESS.2019.2898430`).

[12] Z. Zhang, Z. Junhui, S. Ni, and Y. Gong, "A seamless handover scheme with assisted eNB for 5G C/U plane split heterogeneous network", *IEEE Access*, vol. 7, pp. 164256–164264, 2019 (`https://doi.org/10.1109/ACCESS.2019.2952737`).

[13] M. Alhabo, L. Zhang, and N. Nawaz, "GRA-based handover for dense small cells heterogeneous networks", *IET Communications*, vol. 13, no. 13, pp. 1928–1935, 2019 (`https://doi.org/10.1049/iet-com.2018.5938`).

[14] K. Vasudeva, M. Simsek, D. Lopez-Perez, and I. Guvenc, "Impact of channel fading on mobility management in heterogeneous networks", in *Proc. IEEE International Conference on Communication Workshop* (*ICCW*), London, UK, pp. 2206–2211, 2015 (`https://doi.org/10.1109/ICCW.2015.7247509`).

[15] A.S. Cacciapuoti, "Mobility-aware user association for 5G mmWave networks", *IEEE Access*, vol. 5, pp. 21497–21507, 2017 (`https://doi.org/10.1109/ACCESS.2017.2751422`).

[16] Y. Koda *et al*., "Handover management for mmWave networks with proactive performance prediction using camera images and deep reinforcement learning", *IEEE Transactions on Cognitive Communications and Networking*, vol. 6, no. 2, pp. 802–816, 2020 (`https://doi.org/10.1109/TCCN.2019.2961655`).

[17] M.A. Habibi, M. Nasimi, B. Han, and H.D. Schotten, "A comprehensive survey of RAN architectures toward 5G mobile communication system", *IEEE Access*, vol. 7, pp. 70371–70421, 2019 (`https://doi.org/10.1109/ACCESS.2019.2919657`).

[18] L.F. Ibrahim *et al*., "A survey on heterogeneous mobile networks planning in indoor dense areas", *Personal and Ubiquitous Computing*, vol. 24, no. 4, pp. 487–498, 2019 (`https://doi.org/10.1007/s00779-019-01243-y`).

[19] I. Shayea *et al*., "Real measurement study for rain rate and rain attenuation conducted over 26 GHz microwave 5G link system in Malaysia", *IEEE Access*, vol. 6, pp. 19044–19064, 2018 (`https://doi.org/10.1109/ACCESS.2018.2810855`).

[20] A.S. Bhosle, "Emerging trends in small-cell technology", in *2017 IEEE International Conference on Electrical, Instrumentation and Communication Engineering* (*ICEICE*), pp. 1–4, 2017 (`https://doi.org/10.1109/ICEICE.2017.8191847`).

[21] A. Mukherjee, P. Deb, and D. De, "Small cell zooming based green congestion control in mobile network", *CSI transactions on ICT*, vol. 5, no. 1, pp. 35–43, 2016 (`https://doi.org/10.1007/s40012-016-0141-9`).

[22] S.A. Khan, M. Asshad, K. Küçük, and A. Kavak, "A power control algorithm (PCA) and software tool for femtocells in LTE-A networks", *Sakarya University Journal of Science*, vol. 22, no. 4, pp. 1124–1129, 2018 (`https://doi.org/10.16984/saufenbilder.373293`).

[23] S.A. Khan, A. Kavak, A. Colak, and K. Küçük, "A novel fractional frequency reuse scheme for interference management in LTE-A HetNets", *IEEE Access*, vol. 7, pp. 109662–109672, 2019 (`https://doi.org/10.1109/ACCESS.2019.2933689`).

[24] M.H. Alsharif, A.H. Kelechi, K. Yahya, and S.A. Chaudhry, "Machine learning algorithms for smart data analysis in Internet of Things environment: taxonomies and research trends", *Symmetry*, vol. 12, no. 1, 2020 (`https://doi.org/10.3390/sym12010088`).

[25] G. Liu *et al*., "Coverage enhancement and fundamental performance of 5G: Analysis and field trial", *IEEE Communications Magazine*,

vol. 57, no. 6, pp. 126–131, 2019 (`https://doi.org/10.1109/MCOM.2019.1800543`).

[26] P.-J. Hsieh, W.-S. Lin, K.-H. Lin, and H.-Y. Wei, "Dual-connectivity prevenient handover scheme in control/user-plane split networks", *IEEE Transactions on Vehicular Technology*, vol. 67, no. 4, pp. 3545–3560, 2018 (`https://doi.org/10.1109/TVT.2017.2778065`).

[27] A. Mohamed, M. A. Imran, P. Xiao, and R. Tafazolli, "Memory-full context-aware predictive mobility management in dual connectivity 5G networks", *IEEE Access*, vol. 6, pp. 9655–9666, 2018 (`https://doi.org/10.1109/ACCESS.2018.2796579`).

[28] M. Hassanalian and A. Abdelke, "Classifications, applications, and design challenges of drones: A review", *Progress in Aerospace Sciences*, vol. 91, pp. 99–131, 2017 (`https://doi.org/10.1016/j.paerosci.2017.04.003`).

[29] R. Amorim *et al*., "Pathloss measurements and modeling for UAVs connected to cellular networks", in *Proc. IEEE 85th Vehicular Technology Conference* (*VTC Spring*), Sidney, Australia, 2017 (`https://doi.org/10.1109/VTCSpring.2017.8108204`).

[30] X. Wang, S. Poikonen, and B. Golden, "The vehicle routing problem with drones: Several worst-case results", *Optimization Letters*, vol. 11, no. 4, pp. 679–697, 2017 (`https://doi.org/10.1007/s11590-016-1035-3`).

[31] S. Khunteta and A.K.R. Chavva, "Deep learning-based link failure mitigation", in *2017 16th IEEE International Conference on Machine Learning and Applications* (*ICMLA*), Cancun, Mexico, pp. 806–811, 2017 (`https://doi.org/10.1109/ICMLA.2017.00-58`).

[32] L. Yan *et al*., "Machine learning-based handovers for sub-6 GHz and mmWave integrated vehicular networks", *IEEE Transactions on Wireless Communications*, vol. 18, no. 10, pp. 4873–4885, 2019 (`https://doi.org/10.1109/TWC.2019.2930193`).

[33] M.D. Renzo *et al*., "Smart radio environments empowered by reconfigurable AI meta-surfaces: An idea whose time has come", *EURASIP Journal on Wireless Communications and Networking*, art. no. 129, 2019 (`https://doi.org/10.1186/s13638-019-1438-9`).

[34] E. Basar *et al*., "Wireless communications through reconfigurable intelligent surfaces", *IEEE Access*, vol. 7, pp. 116753–116773, 2019 (`https://doi.org/10.1109/ACCESS.2019.2935192`).

[35] S.E. Elayoubi *et al*., "5G service requirements and operational use cases: Analysis and METIS II vision", in *2016 European Conference on Networks and Communications* (*EuCNC*), Athens, Greece, pp. 158–162, 2016 (`https://doi.org/10.1109/EuCNC.2016.7561024`).

[36] R. Ahmad, E.A. Sundararajan, N.E. Othman, and M. Ismail, "Handover in LTE-advanced wireless networks: state of art and survey of decision algorithm", *Telecommunication Systems*, vol. 66, no. 4, pp. 533–558, 2017 (`https://doi.org/10.1007/s11235-017-0303-6`).

[37] M. Tayyab, X. Gelabert, and R. Jäntti, "A survey on handover management: From LTE to NR", *IEEE Access*, vol. 7, pp. 118907–118930, 2019 (`https://doi.org/10.1109/ACCESS.2019.2937405`).

[38] 3GPP standardization, "Evolved Universal Terrestrial Radio Access (EUTRA) Radio Resource Control (RRC) Protocol specification", TS 36.331 v9.1.0, January 2010 [Online]. Available `http://www.3gpp.org/`

[39] M.A. Esmail *et al*., "5G-28 GHz signal transmission over hybrid all-optical FSO/RF link in dusty weather conditions", *IEEE Access*, vol. 7, pp. 24404–24410, 2019 (`https://doi.org/10.1109/ACCESS.2019.2900000`).

[40] ITU, IMT Vision – Framework and overall objectives of the future development of IMT for 2020 and beyond, M Series, Recommendation ITU-R M.2083-0 (09/2015), 2015 [Online]. Available: `https://www.itu.int/rec/R-REC-M.2083-0-201509-I`

[41] W. Khawaja *et al*., "A survey of air-to-ground propagation channel modelling for unmanned aerial vehicles", *IEEE Communications Surveys and Tutorials*, vol. 21, no. 3, pp. 2361–2391, 2019 (`https://doi.org/10.1109/COMST.2019.2915069`).

[42] A.A.A. Boulogeorgos *et al*., "Terahertz technologies to deliver optical network quality of experience in wireless systems beyond 5G", *IEEE Communications Magazine*, vol. 56, no. 6, pp. 144–151, 2018 (`https://doi.org/10.1109/MCOM.2018.1700890`).

[43] M.Z. Chowdhury *et al*., "Optical wireless hybrid networks for 5G and beyond communications", in *9th International Conference on Information and Communication Technology Convergence* (*ICTC*), pp. 709–712, 2018 (`https://doi.org/10.1109/ICTC.2018.8539460`).

[44] H. Wymeersch, G. Seco-Granados, G. Destino, and D. Dardari, "5G mmwave positioning for vehicular networks", *IEEE Wireless Communications*, vol. 24, no. 6, pp. 80–86, 2017 (`https://doi.org/10.1109/MWC.2017.1600374`).

[45] A. Jain, E. Lopez-Aguilera, and I. Demirkol, "Enhanced handover signaling through integrated MME-SDN controller solution", in *2018 IEEE 87th Vehicular Technology Conference*, Porto, Portugal, pp. 1–7, 2018 (`https://doi.org/10.1109/VTCSpring.2018.8417719`).

Mustafa Mohammed Hasan Alkalsh, M.Sc.
E-mail: mustafa.alkalsh@doctorate.put.poznan.pl
Institute of Radiocommunications, Poznan University of Technology, Poznan, Poland
Department of Mobile Networks, Nokia Solutions and Networks, Wroclaw, Poland