Optimal Slicing of mmWave Micro Base Stations for 5G and Beyond

Abdulhalim Fayad[1], Tibor Cinkler[1, 2, 3],
[1]Department of Telecommunications and Media Informatics, Budapest University of Technology and Economics
[2]ELKH-BME Cloud Applications Research Group, Műegyetem rkp. 3., H-1111 Budapest, Hungary
[3]Department of Computer Communications, Gdansk University of Technology, Gdańsk, Poland

5G and beyond 5G mobile networks are expected to cater to diverse needs by efficiently allocating network resources based on demand. Network slicing is a fundamental approach that involves segregating and allocating network resources distinctly to a group of users based on their individual needs, and it is widely recognized as an essential concept that caters to various requirements. Allocating such slices will encounter conflicting requests, and effectively implementing network slicing presents multiple challenges. Effective network slicing necessitates efficient management of priority levels among diverse slices. Network slicing necessitates efficient management of priority levels across various slices, specifically focusing on three distinctive categories: Ultra-Reliable Low Latency Communications (URLLC), enhanced Mobile Broadband (eMBB), and massive Machine Type Communications (mMTC). This paper proposes an optimization framework utilizing a Mixed Integer Linear Program (MILP) to allocate network resources for multiple slices efficiently. Our framework aims to maximize user satisfaction while ensuring that the specific requirements of each slice are met. We categorize the slices into three priority levels: the URLLC slice holds the highest priority, followed by the eMBB slice, and finally, the mMTC slice receives the least priority. By leveraging our proposed MILP-based approach, we dynamically assign network resources to different slices, considering their priority levels. This allocation strategy enables us to optimize resource utilization and effectively meet the diverse demands of users across various slices. Our framework provides a balance between meeting the stringent requirements of the URLLC slice, delivering high-quality services to the eMBB slice, and accommodating the massive connectivity needs of the mMTC slice.

***Index Terms*—5G and beyond, millimeter wave, slicing, URLLC, eMBB, mMTC, Blocking ratio, optimization, MILP**

I. Introduction

Driven by the exponential growth in mobile data traffic and the increasing demand for bandwidth-intensive and time-critical applications, the fifth generation (5G) of mobile networks has emerged as a revolutionary technology. With its high-speed connectivity, ultra-low latency, and extensive device support, 5G offers a promising platform for a wide range of applications. Compared to its predecessors, 5G represents a significant leap forward, capable of meeting diverse communication requirements across three primary categories [1], *Ultra-Reliable and Low-Latency communications* (URLLC), empowers applications such as self-driving vehicles, unmanned aerial vehicles, industrial automation, remote medical care, and other critical missions. By providing unparalleled reliability and minimal latency, URLLC ensures seamless and instantaneous data transmission. *Enhanced Mobile BroadBand* (eMBB), offers the necessary support for bandwidth-demanding applications like high-definition video, three-dimensional video, cloud-based tasks, and augmented or virtual reality (AR/VR). With its exceptional speed and capacity, eMBB enables immersive multimedia experiences and seamless connectivity. *Massive Machine Type Communications* (mMTC), facilitates extensive device connectivity and serves applications such as intelligent homes, buildings, cities, and the Internet of Things (IoT). mMTC enables a vast network of interconnected devices, fostering smart environments and efficient communication between machines.

Looking ahead, the future of mobile networks, specifically the upcoming 6G technology, will introduce new categories to advance communication capabilities further based on [2], [3]. *Ultra-Reliable low-latency Broadband Communications* (ULBC) will continue to enhance reliability and low-latency communications, enabling even more critical applications and services. It will provide an ultra-fast, highly dependable broadband connection supporting diverse industries. *Massive Ultra-reliable Low-latency Communications* (mULC), Built upon URLLC, mULC will deliver even higher reliability and lower latency, empowering mission-critical applications that demand utmost precision and responsiveness. It will be vital for autonomous transportation, remote surgery, and advanced robotics industries. *Ubiquitous Mobile BroadBand* (uMBB), will further enhance mobile broadband capabilities, ensuring seamless connectivity and high-speed data transmission in any location. It will enable ubiquitous access to the internet, empowering users to stay connected and enjoy rich multimedia experiences wherever they go.

Implementing millimeter wave (mmWave) frequency bands is an indispensable catalyst for revolutionizing the performance of 5G and beyond. By harnessing the power of mmWave, 5G networks can achieve remarkable strides in bandwidth expansion and data rates, surpassing the limitations of conventional sub-6 GHz bands. This technology harnesses the potential of high-frequency signals, unlocking new opportunities for enhanced connectivity and immersive experiences [4], [5].

However, mmWave communications have to face several challenges. These challenges include increased propagation loss, penetration loss, and sensitivity to blockage. Network operators have taken proactive steps to address these difficulties by gradually adopting the deployment of micro base stations (μBS). Integrating these μBS amplifies network coverage and

Manuscript received July 6, 2023; revised October 2, 2023. Corresponding author: Abdulhalim Fayad (email: fayad@tmit.bme.hu).

enhances overall capacity, thereby ensuring a more robust and reliable communications network [6].

Additionally, network slicing has emerged as a promising solution for effectively managing a wide range of services. Through the utilization of network function virtualization (NFV), network slicing empowers the establishment of multiple virtual networks on a single physical infrastructure. This innovative approach enables allocating resources to each slice based on their specific requirements, thereby addressing the ever-increasing demand for data while significantly enhancing network performance and cost efficiency.

Network slicing can assist network operators in allocating resources efficiently and preparing the network to meet various use cases and service requirements, primarily focusing on URLLC, eMBB, and mMTC.

To ensure continuous support and successful implementation of the evolving network slicing paradigm, ongoing network design and development efforts are indispensable. These endeavors are essential for effectively managing the ever-increasing traffic demands and guaranteeing optimal performance in the face of evolving technological advancements [7], [8].

In recent years, a significant amount of research has been dedicated to developing network slicing for 5G and beyond.

For example, the authors of [9] examined the benefits of non-orthogonal allocation of radio access network (RAN) resources in uplink communications for eMBB, mMTC, and URLLC. The authors of [10] analyzed the performance of a 5G network with real-world deployment and testing scenarios. They focused on network slicing mainly four slices (URLLC, eMBB, mMTC, and voice) and incorporated a Blockchain-based model for transparency and security. The study highlighted the benefits of end-to-end slice allocation and demonstrated improved resource handling flexibly and efficiently.

The authors of [11] presented a novel heuristic-based admission control mechanism for network slicing in 5G mobile networks. The mechanism dynamically allocates network resources to maximize user satisfaction while meeting slice requirements. Through simulations, the study showed improved user experience, increased resource utilization, and enhanced scalability as the number of users in each slice increased. The authors of [12] proposed Upper-tier First with Latency-bounded Over-provisioning Prevention (UFLOP) algorithm for optimizing resource allocation in 5G networks. The objective is to minimize over-provisioning while meeting latency constraints and service-level agreement (SLA) for different services. Experimental results showed optimal resource allocation ratios for eMBB, URLLC, and mMTC services. The authors of [13] investigated the problem of URLLC and eMBB resource allocation in future 6G networks. The authors of [14] introduced the concept of network slicing as a critical challenge in mobile network infrastructure, specifically for 5G. They highlighted the advantages of resource isolation and programming flexibility. They focused on expanding network slicing in 5G to accommodate applications (URLLC, eMBB, mMTC) with different requirements and ensuring optimal performance. The authors of [15] studied the problem of network slicing in the context of a wireless system having a time-varying number of users that require two types of slices: reliable low latency (RLL) and self-managed (capacity limited) slices. They proposed a control framework for stochastic optimization that enables the system to maintain slice isolation and provide reliable and low-latency end-to-end communication for RLL slices. The authors of [16] introduced the coupling of Software Defined Networking (SDN) and NFV for flexible resource provisioning in future networks. The authors proposed a mathematical formulation and a low-cost heuristic algorithm to optimize the deployment of network slices for different 5G use cases.

Moreover, In [17], the authors address challenges in NFV related to Service Function Chain (SFC) deployment. They identify two main issues: complex resource scheduling and SFC vulnerability. Their simulations confirm that this approach surpasses traditional resource-based deployment techniques. Finally, The authors of [18] proposed a new method for resource allocation in 5G networks in order to address the side-channel attacks issue.

To the best of the authors' knowledge, the literature has not covered the problem of resource allocation in 5G and beyond BS utilizing mmWaves, considering different user slices while meeting the Quality of Service (OoS) constraints.

A. *Motivations and paper contributions:*

The 5G and beyond mobile networks can be characterized by their dense nature. This density necessitates the use of smaller BSs (μBS) that utilize high-frequency signals (mmWaves), differentiating them from prior generations [19].

Additionally, there is an inhomogeneity in the users demands depending on their category, which poses a need to cater to diverse user requirements by efficiently allocating network resources. The primary mechanism to achieve this is through 'network slicing', which segregates and allocates network resources distinctly to user groups based on their unique needs.

Driven by that, this paper addresses the challenge of efficiently allocating network resources in 5G and beyond mmWave mobile networks based on demand and diverse requirements. The problem becomes complex due to the following reasons: *Conflicting Requests:* As more slices are introduced and different slices have varying needs, conflicting requests arise. Allocating resources without a systematic approach can lead to inefficiencies and unmet demands. *Priority Management:* Efficiently managing the priority levels of different slices is paramount. Each category has its significance and demand, with URLLC requiring the highest priority due to its ultra-reliability and low latency characteristics, eMBB focusing on broadband capabilities, and mMTC aiming at connecting a vast number of devices. *Maximization of User Satisfaction:* Any allocation framework must ensure that the diverse requirements of users across the different slices are met, leading to enhanced user satisfaction.

To address this multifaceted problem, the main contributions of this paper, the main contributions of this paper are as follows:

- We propose a novel optimization framework based on MILP tailored for efficient network resource allocation considering multiple slices in 5G and beyond 5G mmWave mobile networks.
- The proposed MILP categorizes and allocates network resources to slices based on a defined priority hierarchy, covering URLLC, eMBB, and mMTC slices. This ensures optimal resource utilization while addressing the distinct requirements of each slice category.
- We run the simulation to determine the effectiveness of the proposed MILP for different numbers of users, BSs, and various BS capacities.
- Our results pave the way for future research work, emphasizing the potential integration of artificial intelligence and machine learning to enhance resource allocation efficiency and adaptability in network slicing.

The rest of the paper is organized as follows. Section II overviews the Micro base station scope, network slicing concept, and the detailed network model. Section III describes in detail the problem formulation and the proposed MILP formulation. Analyzing the obtained results and discussion are shown in Section IV. Finally, Section V concludes the paper.

II. Network model

A. Micro base station scope

Micro base station are small and lightweight base stations that enhance the capacity and coverage of wireless networks. They are typically used in dense urban areas, where high user density and high traffic demand can cause network congestion and limited coverage. Due to their small size and low power consumption, μBSs can be easily deployed on street lamps, traffic lights, or building facades where traditional base stations cannot be installed. They provide localized wireless coverage, typically within 10 meters up to a few hundred meters. This enables network operators to deploy 5G networks more quickly and efficiently while providing better coverage and capacity than traditional macro base stations.

One of the key advantages of μBSs is their ability to support: (i) a high-capacity energy-efficient channel transceiver, (ii) robust and (omni)directional antennas, (iii) multiple forms of renewable energy harvesting (i.e., light, wind, hydro), (iv) energy-efficient modulation formats, (v) on-board edge processing functionalities (caching, virtualization, and software-defined), and (vi) in-mode supercapacitor storage supplying PA's peak power reducing demands on power supply unit requirements [20].

Based on that the use cases of μBSs in 5G Networks can be summarized as follows:

- *Increased Network Capacity and improved Coverage:* As 5G networks support higher data rates and accommodate a massive number of connected devices, μBSs act as crucial building blocks in meeting the increased demand for data traffic. μBSs reduce congestion and improve overall network capacity by offloading traffic from macro cells and distributing it among μBSs [21].
- *Improved Signal Quality:* 5G networks require extensive coverage, even in densely populated urban or indoor environments. Due to signal attenuation and obstacles, macro base stations alone cannot provide uniform coverage in such scenarios. μBSs fill these coverage gaps by extending the network's reach, especially in areas with weak or no signal reception. By bringing the network closer to the users, μBSs ensure improved signal quality, reduced latency, and higher data rates [21], [22].
- *Dense Urban Deployments:* Urban environments pose challenges for 5G networks due to high user density, increased demand for connectivity, and limited available spectrum. μBSs are crucial in addressing these challenges by enabling dense urban deployments. By leveraging many μBSs, 5G networks can provide localized coverage in heavily populated areas, ensuring consistent performance even during peak usage [23].
- *Offloading Macro Base Stations:* μBSs help offload traffic from macro cells, which provides coverage over larger areas. By distributing the load to μBSs, μBSs relieve the burden on macro cells and optimize their performance. This offloading mechanism is particularly effective in scenarios with concentrated user activity, such as shopping malls, stadiums, or transportation hubs, where the demand for connectivity is significantly higher [24].
- *Millimeter Wave Deployment:* As mmWave signals have a limited range and are susceptible to obstructions. μBSs are essential for mmWave deployment, as they facilitate the densification of network infrastructure required to support these high-frequency signals. By deploying μBSs nearby, mmWave coverage can be extended effectively [25].
- *Enhanced Energy Efficiency:* μBSs are designed to operate at lower power levels compared to macro base stations. This lower power consumption not only contributes to reducing energy costs but also leads to improved energy efficiency in 5G networks. By deploying μBSs strategically, operators can optimize network energy consumption and lower the overall operational expenses associated with powering and cooling network infrastructure [26].

B. Network Slicing concept

The concept of 5G network slicing enables the creation of multiple virtual network instances on a single physical infrastructure. Each virtual network instance, known as a slice, is tailored to meet the specific requirements of diverse applications and services [7]. We consider three different types slices (users categories) starting from the highest to lowest priority URLLC, eMBB, and mMTC as shown in Fig. 1.

Network slicing allows service providers to tailor the network resources, QoS parameters, and performance characteristics based on the specific needs of different applications and industries. This customization enables optimal utilization of network resources, improved efficiency, and enhanced user experiences. Moreover, network slicing enables service providers to differentiate their offerings by providing specialized slices for various applications. This differentiation helps cater to diverse customer needs, improve customer satisfaction, and support the development of new revenue streams.

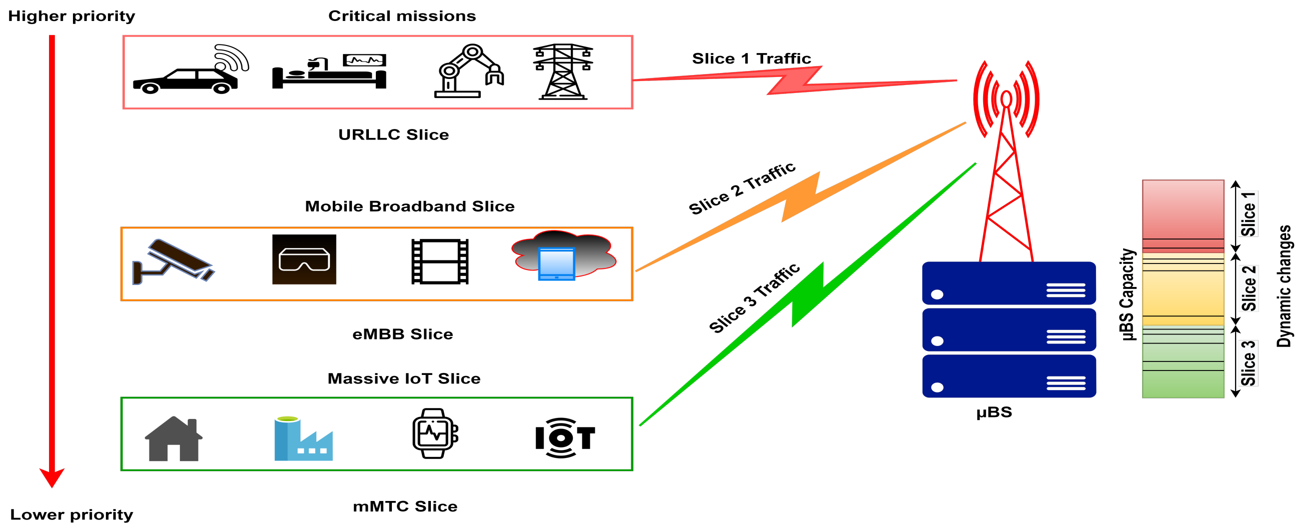


Fig. 1: 5G µBS slicing considering three slice types (URLLC, eMBB, mMTC)

C. Network architecture

Consider a mmWave 5G network comprising a set S of µBSs, denoted as $s \in \{1, \dots, S\}$. Each $\forall s \in \mathcal{S}$ is uniformly distributed in a two-dimensional area with a size of A. We have a set of users $u \in \mathcal{U} \triangleq \{1, \dots, U\}$, where each u represents a user of type URLLC. Similarly, we have a set of users $e \in \mathcal{E} \triangleq \{1, \dots, E\}$ representing users of type eMBB, and a set of users $m \in \mathcal{M} \triangleq \{1, \dots, M\}$ representing users of type mMTC. The complete set of all users is denoted as T, where $T = U \cup E \cup M$. All users, regardless of type, are uniformly distributed across the area, as illustrated in Figure 2. Both the µBSs and user equipment (UEs) are equipped with a single omnidirectional antenna. Additionally, we assume that all UEs are moving based on a random walk scenario [27].

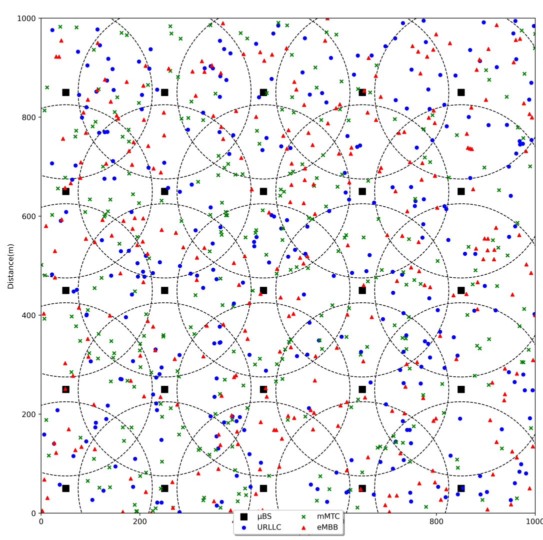


Fig. 2: Example of the user deployments for the case of 900 users (URLLC, eMBB, mMTC) (300 each)

D. Wireless network model

Our research specifically examines the utilization of mmWave frequencies; thus, we adopt a path loss model known as the Floating-Intercept Path Loss Model, which is applicable to mmWave bands [28]. This path loss model is expressed as follows:

$$PL(d) = PL(d_0) + \eta 10 \log_{10}(d) + \xi, \quad \xi \sim N(0, \sigma^2), \tag{1}$$

where

$$PL(d_0) = 20 \log_{10} \left(\frac{4\pi d_0}{\lambda} \right), \tag{2}$$

where $PL(d_0)$ is the path loss in decibels (dB) at a reference distance of d_0=1 meter between the transmitter and the receiver, η is the least square fits of the floating intercept and slope over the measured distances up to 200 meters, σ^2 is the variance of the lognormal shadowing, ξ represents a normally distributed random variable with mean 0 and variance σ^2, and λ is the wavelength.

The received signal power P_{rx}^{sk} for a given user equipment k ($k \in T$) at distance d from a µBS s can be calculated using the expression

$$P_{rx}^{sk} = P_{tx}^{sk} + G_{tx} - PL(d) + G_{rx}, \tag{3}$$

where P_{tx}^{sk} is the transmitted power from µBS s to UE k, and G_{tx} and G_{rx} are the gains of the transmitter and receiver antennas, respectively. The antenna gain can be calculated as follows:

$$G = 20 \log_{10} \left(\frac{\pi \cdot l}{\lambda} \right), \tag{4}$$

Where l is the antenna length, either for the µBS or the UE. While λ is the wavelength based on the used frequency. The transmission power P_{tx}^{sk} from µBS s to UE k must satisfy the power constraint.

$$\sum_{k \in T} P_{tx}^{sk} \leq P_{max} \quad \forall s \in S, \tag{5}$$

where P_{max} is the maximum RF output power of µBS s at its maximum traffic load. The signal-to-interference-plus-noise ratio (SINR) for user k ($k \in T$) from µBS s ($s \in S$) is given by

$$SINR_{sk} = orracP_{sk}^{rx}L(d_{sk})I_s + \text{Pn}, \quad (6)$$

where I_s is the inter-cell interference and Pn is the additive white Gaussian noise power. The interference I caused by other µBSs can be calculated as

$$I_s = \sum_{n \in S, n \neq s} P_{tx}^{nk} L(d_{nk}), \quad (7)$$

while Pn can be calculated using the expression

$$\text{Pn} = -174 + 10 \log(B), \quad (8)$$

where B is the bandwidth in Hertz (Hz) offered by the µBS. According to Shannon's capacity formula we can determine the maximum achievable throughput (data rate) R_{sk} for user k ($k \in T$) of any slice from µBS s as follows:

$$R_{sk} = B \cdot log_2(1 + SINR_{sk}), \quad (9)$$

III. SYSTEM MODEL AND PROBLEM FORMULATION

Given a set of possible locations for the µBSs, a set of locations for URLLC users, a set of locations for eMBB users, and a set of locations for mMTC users, the objective is to assign users of different types to the µBSs while prioritizing their allocations. This prioritization starts with the highest priority slice (URLLC) and ends with the lowest priority slice (mMTC), thus establishing three levels of priority.

In other words, the allocation process considers the different types of users. It assigns them to the µBSs, prioritizing URLLC users over eMBB users and eMBB users over mMTC users. This prioritization ensures that the blocking ratio (BR) is highest for the mMTC slice, indicating that these users have the lowest priority regarding resource allocation.

The capacity of the µBSs is dynamically sliced based on the connection request slices. This means that the available capacity of the µBSs is allocated and adjusted dynamically to accommodate the connection requests from different types of users. Each user is assumed to be assigned to one µBS, ensuring efficient resource utilization and optimal allocation based on the user's requirements and priorities. The problem can be formulated as a Mixed Integer Linear Program (MILP) as follows:

Optimization parameters

The optimization parameters are described in Table. I.

Decision variables:

The considered decision variables are illustrated in Table. II

Objective function:

$$\max \sum_{i \in S} \sum_{u \in U} \alpha x_{iu} + \sum_{i \in S} \sum_{e \in E} \beta y_{ie} + \sum_{i \in S} \sum_{m \in M} \gamma z_{im}, \quad (10)$$

Subject to:

TABLE I: Optimization parameters

Parameter	Description
α	URLLC user priority value.
β	eMBB user priority value.
γ	mMTC user priority value.
D	µBS maximum coverage distance.
d_{iu}	the distance between URLLC user u and µBS i.
d_{ie}	the distance between eMBB user e and µBS i.
d_{im}	the distance between mMTC user m and µBS i.
C	the capacity of µBS in Gbps.
C_u	the required capacity for each URLLC user.
C_e	the required capacity for each mMTC user.
C_m	the required capacity for each mMTC user.
R_{min}	minimum throughput for each user.

TABLE II: Decision variables

Decision variable	Description
Ω	continues variable presents the capacity used for allocating URLLC users to a µBS, where $0 \leq \Omega \leq 1$.
Φ	continues variable presents the capacity used for allocating eMBB users to a µBS, where $0 \leq \Phi \leq 1$.
Ψ	continues variable presents the capacity used for allocating mMTC users to a µBS, where $0 \leq \Psi \leq 1$.
x_{iu}	binary variable equals 1 if URLLC user u is allocated to µBS i, 0 otherwise.
y_{ie}	binary variable equals 1 if eMBB user e is allocated to µBS i, 0 otherwise.
z_{im}	binary variable equals 1 if mMTC user m is allocated to µBS i, 0 otherwise.
R_{iu}	the throughput of user u from the µBS i.
R_{ie}	the throughput of user e from the µBS i.
R_{im}	the throughput of user m from the µBS i.

1) C1: each user from any slice can only assigned to one µBS:

$$\sum_{i \in S} x_{iu} \leq 1, \quad \forall u \in U, \quad (11)$$

$$\sum_{i \in S} y_{ie} \leq 1, \quad \forall e \in E, \quad (12)$$

$$\sum_{i \in S} z_{im} \leq 1, \quad \forall m \in M, \quad (13)$$

2) C2: The distance between each user of any slice and its µBS must be less or equal to to the µBS radius:

$$x_{iu} \cdot d_{iu} \leq D, \quad \forall u \in U, i \in S, \quad (14)$$

$$y_{ie} \cdot d_{ie} \leq D, \quad \forall e \in E, i \in S, \quad (15)$$

$$z_{im} \cdot d_{im} \leq D, \quad \forall m \in M, i \in S, \quad (16)$$

3) C3: capacity constraints:

$$\sum_{u \in U} C_u \cdot x_{iu} \leq \Omega \cdot C \quad \forall i \in S, \quad (17)$$

$$\sum_{e \in E} C_e \cdot y_{ie} \leq \Phi \cdot C \quad \forall i \in S, \quad (18)$$

$$\sum_{m \in M} C_m \cdot z_{im} \leq \Psi \cdot C \quad \forall i \in S, \quad (19)$$

4) C4: continues variables constraint:

$$\Omega + \Phi + \Psi \leq 1, \tag{20}$$

5) C5: Qality of service constraint, the data rate allocated to each user from each slice must be at least equal to the minimum threshold R_{min}:

$$R_{iu} \cdot x_{iu} \geq R_{min}, \quad \forall u \in U, i \in S, \tag{21}$$

$$R_{ie} \cdot y_{ie} \geq R_{min}, \quad \forall e \in E, i \in S, \tag{22}$$

$$R_{im} \cdot z_{im} \geq R_{min}, \quad \forall m \in M, i \in S, \tag{23}$$

IV. Simulation results and discussion

A. Simulation setup

In this section, we analyze the performance of our proposed MILP using simulations in a 1000m x 1000m two-dimensional area. We obtain the optimal solutions by utilizing the commercially available IBM ILOG CPLEX solver with Python programming. Our setup includes an Intel(R) Core(TM) i5-1035G1 CPU running at 1.00 GHz and 8 GB RAM, ensuring efficient computations.

The parameters used in the simulation are summarized in Table. III.

TABLE III: Simulation parameters

Parameter	Value
Area size	1000m*1000 m
µBS radius (D)	175, 200, 250 m
Number of users $\|T\|$	300, 600, 900, 1200, 1500
Number of URLLC users $\|U\|$	100, 200, 300, 400, 500
Number of eMBB users $\|E\|$	100, 200, 300, 400, 500
Number of mMTC users $\|M\|$	100, 200, 300, 400, 500
Number of BSs ($\|S\|$)	25, 36
α, β, γ	100, 10, 1
Capacity of the µBS C	5, 10, 15, 20 Gbps
C_u	120 Mbps [29]
C_e	720 Mbps [29]
C_m	80 Mbps [29]
Frequency	60 GHz
Bandwidth (B)	1200 MHz
d_0	1 m
$L(d_0)$	71 dB [30]
η	1.8 [30]
P_{max}	30 dBm
ξ	2.9 [30]
v_{max}	1.3 m/s
λ	5 mm
$l(tx)$	0.1 m
$l(rx)$	0.01 m
R_{min}	20 Mbps

B. Results discussion

To demonstrate the efficacy of our proposed mmWave µBS slicing MILP model, this study focuses on a comprehensive evaluation of the BR for three distinct slice categories: URLLC, eMBB, and mMTC. We investigate various scenarios involving different numbers of users, diverse µBS capacities, and varying numbers of µBSs.

Furthermore, our evaluation extends to analyzing the BR values across a range of µBS radius values and considering different numbers of µBSs. These aspects provide a clear understanding of the performance and robustness of our proposed MILP approach.

Figure 3 illustrates the optimal BR as a function of the number of users for different µBS capacities, namely 5, 10, 15, and 20 Gbps, with a fixed number of 25 µBSs. The BR for any slice type is determined using the following equation:

$$\text{BR} = \frac{\text{Number of allocated users of the slice type X}}{\text{Total number of users of the slice type X}}, \tag{24}$$

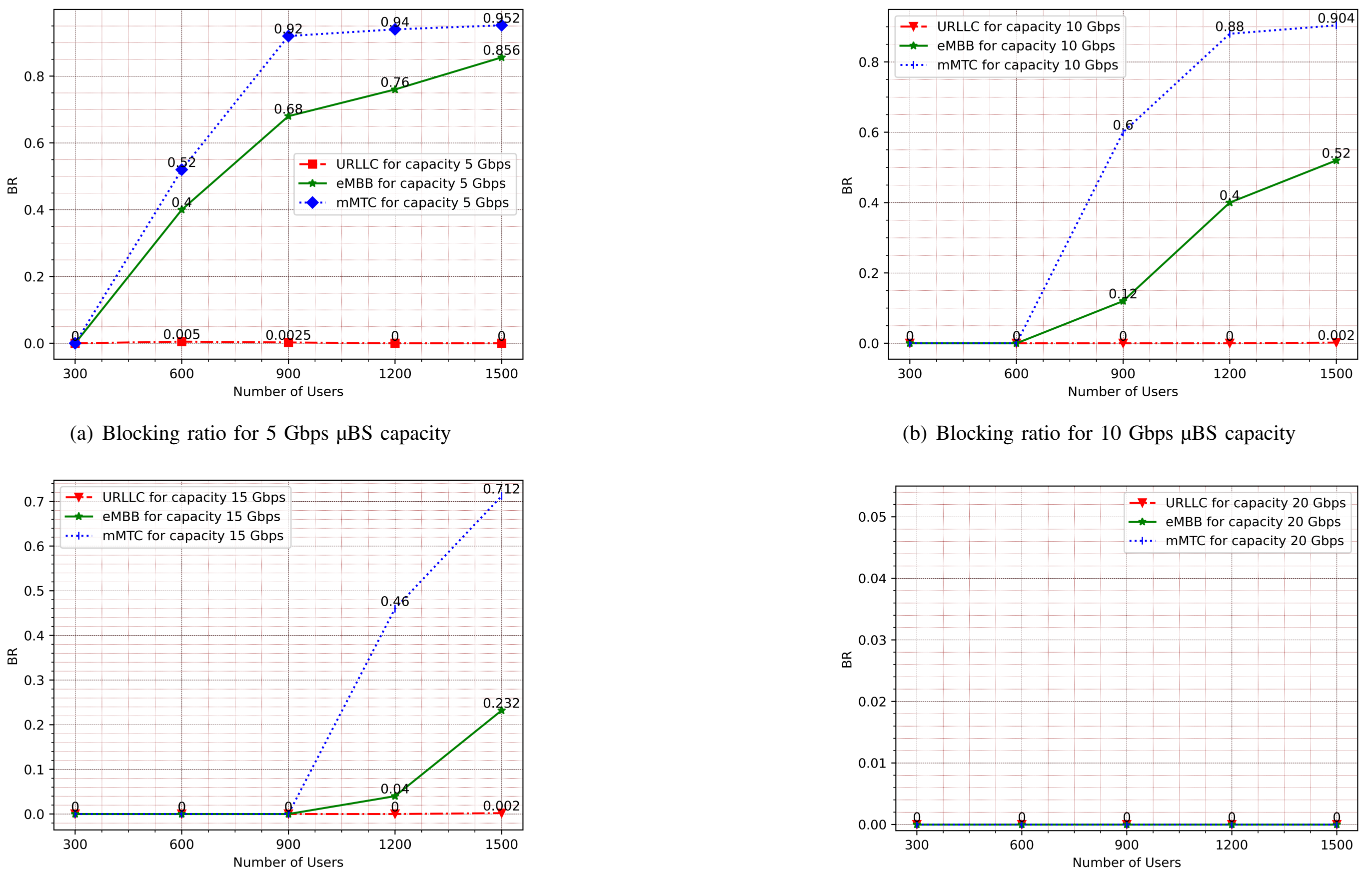
As expected, the results demonstrate that a higher µBS capacity leads to a lower BR. Furthermore, an increase in the number of users within a slice results in a higher BR, particularly for the eMBB and mMTC slices. Figure 3(a) specifically shows the case where the µBS capacity is 5 Gbps. Notably, the BR for the URLLC slice approaches zero for various numbers of users, owing to the high allocation priority assigned to URLLC slice users by the optimization framework. Conversely, the mMTC slice users experience higher BR values, indicative of their relatively lower priority.

Additionally, Figures 3(b), 3(c) and 3(d) illustrate the BR for scenarios where the µBS capacity is 10, 15, and 20 Gbps, respectively. It becomes evident that as the µBS capacity increases, the BR decreases. This is due to the fact that a higher-capacity µBS can accommodate more users. We can observe from Fig 3(d) that the BR value equals 0 for all slices, that is because the is sufficient capacity for allocating all users of all slices.

Figure 4 illustrates the free network capacity (non-allocated resources) as a function of the µBS capacity (5, 10, 15, and 20 Gbps) for different numbers of users. The total capacity can be calculated by multiplying the number of µBSs by the capacity of the µBS. It is evident from the figure that as the µBS capacity increases, so does the amount of available free resources across the network. However, an interesting observation can be made for the case when the µBS capacity is set to 5 Gbps. In this scenario, there is a noticeable increase in free capacity over the network as the number of users changes from 600 to 900 and then to 1200. This can be attributed to the high bandwidth requirements of eMBB users, which necessitate a significant amount of capacity. By blocking these eMBB users, additional free resources can be obtained, resulting in a higher overall network capacity.

Furthermore, Figure 5 presents the total BR as a function of the number of users for two different numbers of µBSs (25 and 36) across the URLLC, eMBB, and mMTC slices. The respective capacities of the µBSs are maintained at 5, 10, 15 and 20 Gbps. The graph clearly demonstrates a significant reduction in BR values when the number of µBSs is increased from 25 to 36, as there will be more capacity that can be allocated to the blocked users.

Figure 6 provides a clear illustration of the relationship between the BR and the µBS radius (175m, 200m, and 250m) for various numbers of users across different slices, given that the capacity of the µBS is equal to 10 Gbps. The graph reveals that as the number of users grows, the BR also experiences an increase. In contrast, when the µBS radius becomes larger,



(a) Blocking ratio for 5 Gbps µBS capacity

(b) Blocking ratio for 10 Gbps µBS capacity

(c) Blocking ratio for 15 Gbps µBS capacity

(d) Blocking ratio for 20 Gbps µBS capacity

Fig. 3: Blocking ratio vs µBS capacity for different number of users

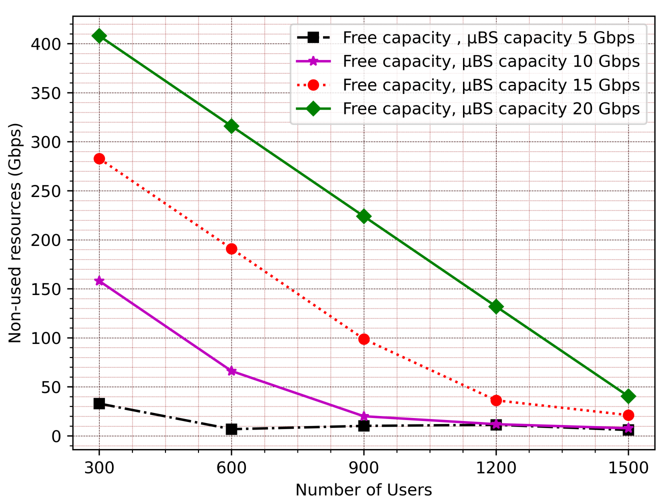


Fig. 4: Free resources vs µBS capacity for different numbers of users after the allocation

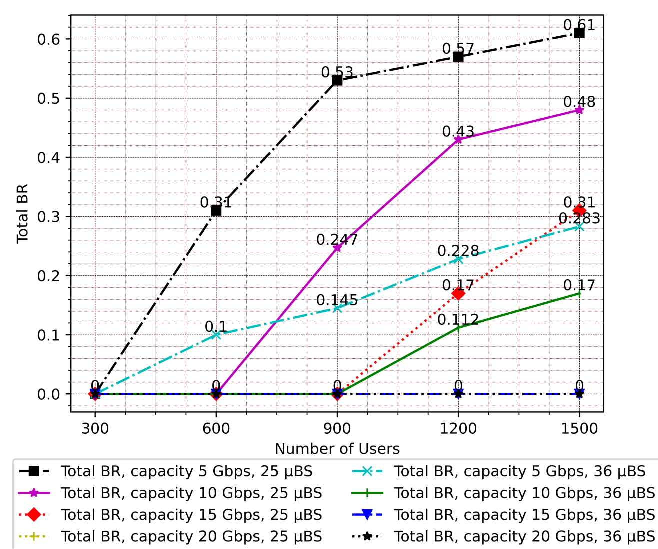
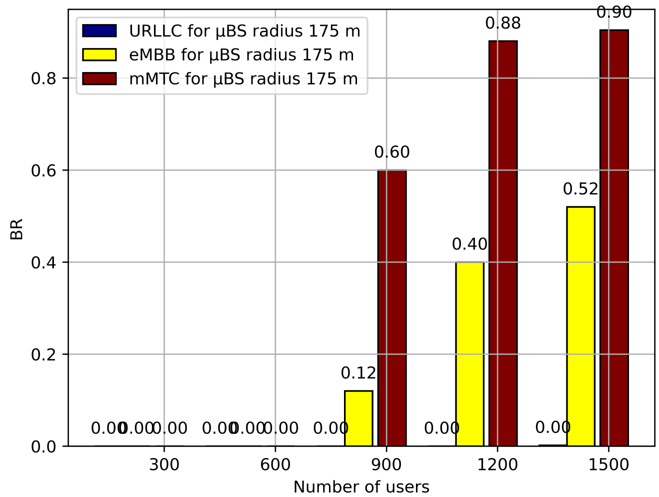


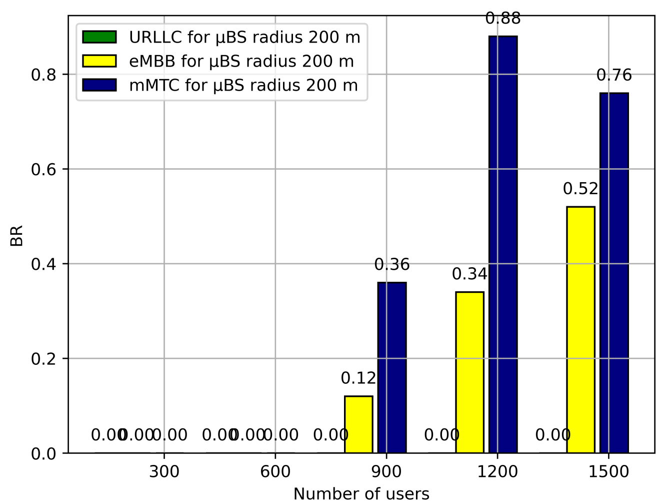
Fig. 5: Total blocking ratio vs µBS capacity for different numbers of users

the BR experiences a decrease. This behavior can be attributed to the improved load balancing between µBSs as the distance between them increases.

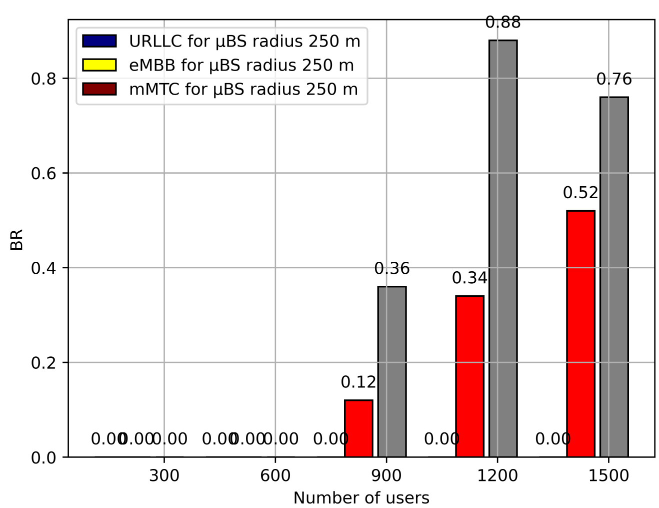
Figure 6(a) displays the BR for a range of user counts when the µBS radius is set at 175m. Similarly, Figures 6(b) and 6(c) show the scenarios with µBS radii of 200m and 250m, respectively. We can observe that the BR value decrease with



(a) Blocking ratio for μBS radius of 175 m



(b) Blocking ratio for μBS radius of 200 m



(c) Blocking ratio for μBS radius of 250 m

Fig. 6: Blocking ratio vs μBS radius for different number of users with μBS capacity of 10 Gbps

the increase of the μBS. However, in the case of μBS radius either equals to 200 m or 250 m and the number of users changes from 1200 to 1500 users, we can see that the BR value for eMBB users increased while we can see that the BR value mMTC decreases. The reason is that the number of eMBB users allocated to each μBS equals 10 users, where the capacity of the μBS equals 10 Gbps only. When the number of eMBB increased from 400 to 500, the BR value for the eMBB users increased as there was insufficient capacity for allocating new eMBB users. In contrast, the BR value for the mMTC users decreased as the free capacity of each μBS was used for allocating more mMTC users.

Figure 7 illustrates the non allocated network resources (free resources) corresponding to different μBS radius values (175 m, 200 m, and 250 m) considering a capacity of 10 Gbps for each μBS. The findings from the figure suggest that when the radius of the μBS is set to 175 m, the non-allocated network capacity surpasses that of the μBS with a radius of 200 m or 250 m because more users can be connected to the μBS due to the larger distance. Additionally, it is noticeable that the same amount of free resources is achieved for both μBS radius values of 200 m and 250 m. This is attributed to the fact that the utilized capacity remains constant due to the identical value of BR for both radius values.

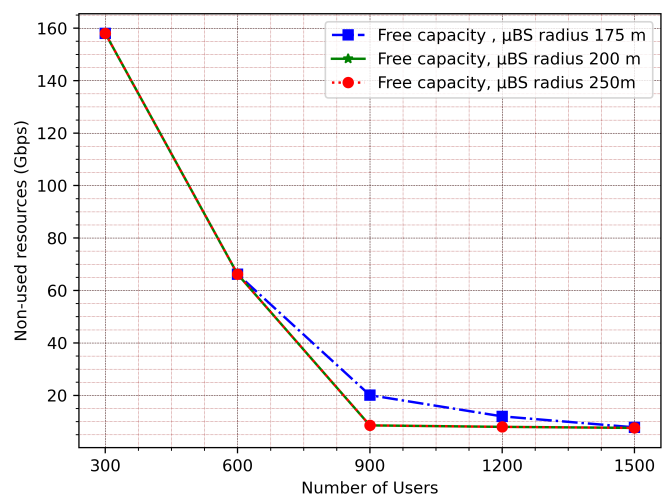


Fig. 7: Free resources vs μBS radius for different numbers of users after the allocation

Moreover, Figure 8 shows the total BR in relation to different μBS radii (175m, 200m, and 250m) and two different numbers of μBSs (25 and 36). In general, a higher number of users results in an elevated BR, and a larger cell radius leads to a reduced BR. Notably, when the μBS radius changes from 200m to 250m, there is no change in the BR value. This occurs because users are already covered when the μBS radius is at 200m, and no additional capacity is available to allocate to the blocked users. However, when the number of μBSs increases from 25 to 36, the BR is significantly reduced.

V. Conclusion

In this paper, we have presented MILP-based optimization framework for mmWave μBSs for 5G and beyond mobile

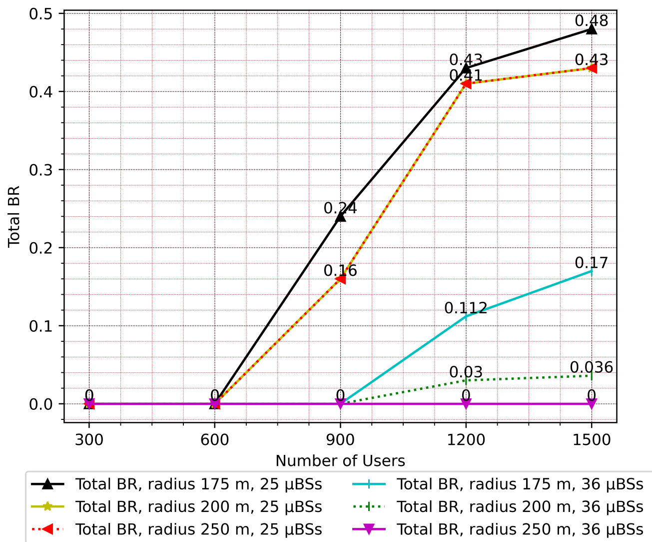


Fig. 8: Total blocking ratio vs µBS radius for different numbers of users

networks optimal slicing. The aim was to effectively allocate network resources for multiple slices with different priorities. Our proposed framework addresses the diverse requirements of the three primary categories: URLLC, eMBB, and mMTC. By dynamically assigning network resources to different slices based on their priority levels, our approach optimizes resource utilization while ensuring that the specific needs of each slice are met. We have applied the proposed MILP on a two-dimensional area with different numbers of users of different slices, considering various numbers of µBSs with different capacities and sizes. We have shown that the proposed model maximizes the efficiency of resource allocation for different slices and QoS requirements.

The results of this work can serve as a basis for further research in network slicing and resource allocation for 5G and beyond 5G mobile networks. Future work could explore integrating machine learning and artificial intelligence techniques to enhance the adaptability and efficiency of resource allocation algorithms. Additionally, investigating the impact of real-world constraints, such as energy consumption and hardware limitations, on the performance of the proposed framework would be valuable in developing more practical and robust solutions for network slicing.

Acknowledgment

This work was supported by the CHIST-ERA SAMBAS grant (CHIST-ERA-20-SICT-003) funded by FWO, ANR, NKFIH, and UKRI.

References

[1] M. Agiwal, A. Roy, and N. Saxena, "Next generation 5g wireless networks: A comprehensive survey," *IEEE communications surveys & tutorials*, vol. 18, no. 3, pp. 1617–1655, 2016.

[2] W. Jiang, B. Han, M. A. Habibi, and H. D. Schotten, "The road towards 6g: A comprehensive survey," *IEEE Open Journal of the Communications Society*, vol. 2, pp. 334–366, 2021.

[3] A. Fayad, T. Cinkler, and J. Rak, "5g/6g optical fronthaul modeling: cost and energy consumption assessment," *Journal of Optical Communications and Networking*, vol. 15, no. 9, pp. D33–D46, 2023.

[4] Y. Niu, Y. Li, D. Jin, L. Su, and A. V. Vasilakos, "A survey of millimeter wave communications (mmwave) for 5g: opportunities and challenges," *Wireless networks*, vol. 21, pp. 2657–2676, 2015.

[5] A. Fayad, T. Cinkler, and J. Rak, "5g millimeter wave network optimization: Dual connectivity and power allocation strategy," *IEEE Access*, 2023.

[6] W. Guo, J. Koo, I. F. Siddiqui, N. M. F. Qureshi, and D. R. Shin, "Qos-aware energy-efficient microbase station deployment for 5g-enabled hetnets," *Journal of King Saud University-Computer and Information Sciences*, vol. 34, no. 10, pp. 10 487–10 495, 2022.

[7] X. Foukas, G. Patounas, A. Elmokashfi, and M. K. Marina, "Network slicing in 5g: Survey and challenges," *IEEE communications magazine*, vol. 55, no. 5, pp. 94–100, 2017.

[8] S. Zhang, "An overview of network slicing for 5g," *IEEE Wireless Communications*, vol. 26, no. 3, pp. 111–117, 2019.

[9] P. Popovski, K. F. Trillingsgaard, O. Simeone, and G. Durisi, "5g wireless network slicing for embb, urllc, and mmtc: A communication-theoretic view," *Ieee Access*, vol. 6, pp. 55 765–55 779, 2018.

[10] P. Gorla, V. Chamola, V. Hassija, and D. Niyato, "Network slicing for 5g with ue state based allocation and blockchain approach," *IEEE Network*, vol. 35, no. 3, pp. 184–190, 2020.

[11] M. Jiang, M. Condoluci, and T. Mahmoodi, "Network slicing management & prioritization in 5g mobile systems," in *European wireless 2016; 22th european wireless conference*. VDE, 2016, pp. 1–6.

[12] H.-T. Chien, Y.-D. Lin, C.-L. Lai, and C.-T. Wang, "End-to-end slicing with optimized communication and computing resource allocation in multi-tenant 5g systems," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 2, pp. 2079–2091, 2019.

[13] M. Al-Ali and E. Yaacoub, "Resource allocation scheme for embb and urllc coexistence in 6g networks," *Wireless Networks*, pp. 1–20, 2023.

[14] A. Nakao, P. Du, Y. Kiriha, F. Granelli, A. A. Gebremariam, T. Taleb, and M. Bagaa, "End-to-end network slicing for 5g mobile networks," *Journal of Information Processing*, vol. 25, pp. 153–163, 2017.

[15] A. T. Z. Kasgari and W. Saad, "Stochastic optimization and control framework for 5g network slicing with effective isolation," in *2018 52nd Annual Conference on Information Sciences and Systems (CISS)*. IEEE, 2018, pp. 1–6.

[16] A. Kammoun, N. Tabbane, G. Diaz, A. Dandoush, and N. Achir, "End-to-end efficient heuristic algorithm for 5g network slicing," in *2018 IEEE 32nd International Conference on Advanced Information Networking and Applications (AINA)*. IEEE, 2018, pp. 386–392.

[17] P. Zou, S. Yang, T. Zhang, and S. Wei, "Risk-aware sfc placement method in edge cloud environment," in *International Conference on Emerging Networking Architecture and Technologies*. Springer, 2022, pp. 105–116.

[18] T. Zhang, C. Xu, B. Zhang, X. Li, X. Kuang, and L. A. Grieco, "Towards attack-resistant service function chain migration: A model-based adaptive proximal policy optimization approach," *IEEE Transactions on Dependable and Secure Computing*, 2023.

[19] B. T. Tinh, L. D. Nguyen, H. H. Kha, and T. Q. Duong, "Practical optimization and game theory for 6g ultra-dense networks: Overview and research challenges," *IEEE Access*, vol. 10, pp. 13 311–13 328, 2022.

[20] "sambas-project." [Online]. Available: http://sambas-project.com/index.php/objectives/

[21] M. A. Adedoyin and O. E. Falowo, "Combination of ultra-dense networks and other 5g enabling technologies: A survey," *IEEE Access*, vol. 8, pp. 22 893–22 932, 2020.

[22] Á. Ladányi and T. Cinkler, "Resilience–throughput–power trade-off in future 5g photonic networks," *Photonic Network Communications*, vol. 37, pp. 296–310, 2019.

[23] F. Richter and G. Fettweis, "Cellular mobile network densification utilizing micro base stations," in *2010 IEEE International Conference on Communications*. IEEE, 2010, pp. 1–6.

[24] L. N. Huynh, Q.-V. Pham, X.-Q. Pham, T. D. Nguyen, M. D. Hossain, and E.-N. Huh, "Efficient computation offloading in multi-tier multi-access edge computing systems: A particle swarm optimization approach," *Applied Sciences*, vol. 10, no. 1, p. 203, 2019.

[25] G. E. Athanasiadou, P. Fytampanis, D. A. Zarbouti, G. V. Tsoulos, P. K. Gkonis, and D. I. Kaklamani, "Radio network planning towards 5g mmwave standalone small-cell architectures," *Electronics*, vol. 9, no. 2, p. 339, 2020.

[26] M. De Ree, G. Mantas, A. Radwan, S. Mumtaz, J. Rodriguez, and I. E. Otung, "Key management for beyond 5g mobile small cells: A survey," *IEEE Access*, vol. 7, pp. 59 200–59 236, 2019.

[27] T. Camp, J. Boleng, and V. Davies, "A survey of mobility models for ad hoc network research," *Wireless communications and mobile computing*, vol. 2, no. 5, pp. 483–502, 2002.
[28] T. S. Rappaport, G. R. MacCartney, M. K. Samimi, and S. Sun, "Wideband millimeter-wave propagation measurements and channel models for future wireless communication system design," *IEEE transactions on Communications*, vol. 63, no. 9, pp. 3029–3056, 2015.
[29] C. Song, M. Zhang, Y. Zhan, D. Wang, L. Guan, W. Liu, L. Zhang, and S. Xu, "Hierarchical edge cloud enabling network slicing for 5g optical fronthaul," *Journal of Optical Communications and Networking*, vol. 11, no. 4, pp. B60–B70, 2019.
[30] B. De Beelde, Z. Verboven, E. Tanghe, D. Plets, and W. Joseph, "Outdoor mmwave channel modeling for fixed wireless access at 60 ghz," *Radio Science*, vol. 57, no. 12, p. e2022RS007519, 2022.