This is the post-print of the following article: N. Chukhno, O. Chukhno, S. Pizzi, A. Molinaro, A. Iera and G. Araniti, "Efficient Management of Multicast Traffic in Directional mmWave Networks", in IEEE Transactions on Broadcastig, doi: 10.1109/TBC.2021.3061979.
Article has been published in final form at: https://ieeexplore.ieee.org/document/9376272.

Efficient Management of Multicast Traffic in Directional mmWave Networks

Nadezhda Chukhno*,†, Olga Chukhno*,‡, Sara Pizzi*, Antonella Molinaro*,§, Antonio Iera¶, Giuseppe Araniti*
*University Mediterranea of Reggio Calabria, Italy and CNIT, Italy
†Universitat Jaume I, Castelló de la Plana, Spain
‡Tampere University, Finland
§Université Paris-Saclay, Gif-sur-Yvette, France
¶University of Calabria, Italy and CNIT, Italy
e-mail: {*nadezda.chukhno, olga.chukhno, sara.pizzi, antonella.molinaro, araniti*}*@unirc.it*, *antonio.iera@dimes.unical.it*

***Abstract*—Multicasting is becoming more and more important in the Internet of Things (IoT) and wearable applications (e.g., high definition video streaming, virtual reality gaming, public safety, among others) that require high bandwidth efficiency and low energy consumption. In this regard, millimeter wave (mmWave) communications can play a crucial role to efficiently disseminate large volumes of data as well as to enhance the throughput gain in fifth-generation (5G) and beyond networks. There are, however, challenges to face in view of providing multicast services with high data rates under the conditions of short propagation range caused by high path loss at mmWave frequencies. Indeed, the strong directionality required at extremely high frequency bands excludes the possibility of serving all multicast users via a single transmission. Therefore, multicasting in directional systems consists of a sequence of beamformed transmissions to serve all multicast group members, subgroup by subgroup. This paper focuses on multicast data transmission optimization in terms of throughput and, hence, of the energy efficiency of resource-constrained devices such as wearables, running their resource-hungry applications. In particular, we provide a means to perform the beam switching and propose a radio resource management (RRM) policy that can determine the number and width of the beams required to deliver the multicast content to all interested users. Achieved simulation results show that the proposed RRM policy significantly improves network throughput with respect to benchmark approaches. It also achieves a high gain in energy efficiency over unicast and multicast with fixed predefined beams.**

***Index Terms*—Multicast, Millimeter Wave Communication, Radio Resource Management, Wearable Devices.**

I. Introduction

Recently, the popularity of millimeter wave (mmWave) wireless networks has increased due to their capability to cope with the escalation of mobile data demands caused by the unprecedented proliferation of smart devices in the fifth-generation communication system (5G). Extremely high frequency (EHF) or mmWave band is a fundamental pillar in the provision of the expected gigabit data rates. Hence, according to both academic and industrial communities, mmWave technology, e.g., 5G New Radio (NR) and WiGig (60 GHz), is considered as one of the main components of 5G and beyond 5G (B5G) networks [1], [2]. Particularly, the 3GPP provides for the use of licensed mmWave sub-bands (e.g., 24.25-27.5, 27.5-29.5, 37-40, 64-71 GHz) for the 5G mmWave cellular networks [3], whereas IEEE actively explores the unlicensed band at 60 GHz for the next-generation wireless local area networks (WLANs) [4]. Bandwidths of cellular systems range between 500 MHz and 2 GHz, which results in a cell capacity of several gigabits per second [5], [6], whereas the IEEE 802.11ay enables Wi-Fi devices to achieve up to 100 Gbps [7]. In this regard, mmWave has been envisaged as a new technology layout for *real-time heavy-traffic applications*, such as ultra-high definition (UHD) video streaming [8], extended reality (XR) broadcasting [9] that includes augmented, virtual, and mixed reality (AR/VR/MR), and proximate gaming [10].

Meanwhile, multicast transmission can provide effective system bandwidth usage and energy efficiency improvement, thus playing a crucial role in emerging applications [11], [12]. In multicast transmissions, a device, which acts as a personal basic service set (PBSS) central point (PCP) or access point (AP), may transmit the same packet to a group of receivers simultaneously by utilizing the same frequency and the same modulation and coding scheme (MCS) [13]. Further, multicasting can improve the total network throughput [14], which is a critical feature for ultra-high-speed data transmissions. Hence, a perfect candidate for multicast scenarios requiring the distribution of the large volume of data with low latency is undoubtedly the multi-gigabit rate communication enabled at EHF bands such as mmWave.

The benefits of coupling directional mmWave communications with multicast traffic delivery are manifold. On the one hand, the mmWave frequencies offer the availability of huge bandwidths, high data rates, and, simultaneously, a decrease in the antenna size. On the other hand, multicast transmissions have been proven to gain bandwidth efficiency in various scenarios. While much effort has been invested in optimizing the performance of individual radio systems, e.g., in [15], [16], limited research attention has been dedicated to the joint use of mmWave and multicast networks.

Transmissions at mmWave use highly directional antennas to guarantee the gigabit capabilities and overcome the short propagation range, thereby *(i)* suffering from the limited coverage caused by the oxygen absorption and severe path loss and *(ii)* making the multicast fashion more complex. The former drawback makes it unfeasible to serve users spread over

large regions at a time with one beam due to the decrease in antenna gain. The latter is an effect of the directionality of mmWave systems, which complicates multicast deployment by posing additional challenges [17] (e.g., beam steering and proper selection of beamwidth). Hence, the proper beamwidth and data rate setting is one of the most challenging issues in multicast with directional antennas.

This work analyzes the performance of multicast, unicast, and sequential multicast transmissions in the directional system and provides solutions for radio resource management (RRM). The NR AP performs RRM decisions based on the number of users in a certain area (i.e., the user density). Our method aims to guarantee QoS requirements of bandwidth-hungry applications (e.g., VR/AR multicasting), while reducing energy consumption of devices. The main contributions of this paper are three-fold and summarized as follows:

- We propose the policy and potential thresholds for switching from unicast to multicast transmissions and vice versa in directional systems by relying upon stochastic geometry methods. We approximate the coverage area of an antenna as a drop, which shows close results to real ones.
- We design a flexible resource management algorithm for the multicast transmission in mmWave networks with the goal of the system performance optimization in terms of *(i)* energy efficiency and *(ii)* network throughput. The proposed algorithm exploits a drop-based approach while traversing a beam tree to reduce algorithm complexity and, hence, computational time.
- We investigate the influence of different parameters of heuristic algorithms to optimize the multicast transmission of large-volume data. We then derive a practical conclusion that network throughput optimization of each transmission is required to maximize energy efficiency of the system from both user and network sides.

The remainder of this paper is organized as follows. Section II reviews the related work. Section III gives a brief overview of the challenges in directional multicast systems and describes the considered scenario and network topology features. Our system model is then specified in Section IV. The proposed beam switching policy and RRM scheme are presented in Section V. In Section VI, we discuss achieved numerical results. Conclusive remarks are given in the last section.

II. Related Work

Several works in the literature propose strategies for group-oriented communications in directional systems. In [18], a group-aware multicast scheme (GAMS) compatible with the IEEE 802.11ad to manage steerable beamforming for multicast devices is proposed. Specifically, multicast beamforming is performed during an association beamforming training (A-BFT) interval of 802.11ad beacon interval (BI). However, the most suitable data rate for the multicast group is obtained using the cosine law, which is not applicable to real-life scenarios as it cannot capture antenna radiation features. In [19], a rate adaptation algorithm, which seeks to find out appropriate transmission rates that preserve fairness among sectors and accommodate as many multicast devices as possible, is proposed. The authors focus on delay-sensitive applications, whereas energy efficiency aspects are not considered. Moreover, the approach proposed in [19] fails to serve all multicast users with the high data rate required by emerging applications. An alternative method is presented in [20], wherein the beamwidth is adaptively generated depending on the locations of multicast devices, their number, and data rates. This approach assumes an exhaustive search for a multicast beam, which affects the computational time and, consequently, influences the multicast beamforming duration.

Since unnecessary sector switching in multicast transmissions with directional antennas leads to a long delay, and, hence, to a low *throughput*, in work [21], asymmetric sectorization for the irregular deployment pattern of multicast group members is optimized by sweeping different sizes of beams to cover all multicast group members with the minimum number of directional transmissions. However, this approach relies on sector antenna models, which provide trivial cut-off solutions [22], thereby leading to non-optimal results. Another approach is investigated in [23], wherein spatial reuse scheduling problem with multicast transmissions is under investigation to enhance network capacity in a time division multiple access (TDMA) based mmWave system. More precisely, the authors state that leveraging the spatial sharing, wherein the simultaneous transmissions of single-hop unicast and multicast sessions are enabled, may increase the network efficiency. In [24], based on the training information and starting with only the finest beams, a scalable beam grouping algorithm to achieve the minimum multicast group data transmission time is designed. In particular, the algorithm traverses a codebook tree[1] in descending order to maximize the *throughput* delivered to multicast groups. Differently from other works, a method for selecting the MCS (an appropriate bitrate for the multicast stream) based on the analysis of the distribution of users in the service area is proposed in [25].

The possibility to exploit device-to-device (D2D) communications in directional multicast systems has been investigated in several works. In [26], the efficient multicast scheduling (EMS) is developed, where D2D multi-hop and concurrent transmissions (spatial reuse) are jointly exploited to achieve lower energy consumption in comparison with mmWave multicast performed through serial unicast transmissions. Recently in [27], the optimal multicast scheduling problem with D2D communications, concurrent transmission, multicast group partition, and beam selection in a multilevel codebook is formulated. A similar approach is proposed in [28], where multicast scheduling jointly exploits relaying and spatial sharing properties of mmWave networks that work at 73 GHz. The proposed multicast scheme aims to minimize the overall data delivery time for all multicast group members. To this end, transmitting nodes and their target destinations are properly selected at each time slot.

In [29] and more recently in [17], a multicast transmission

[1]Traversing means visiting every node in the tree composed of precomputed transmit and receive beams of different beamwidths (that is, codebook).

strategy for mmWave in NR that aims to find an optimal trade-off between the base station (BS) resource consumption (channel usage time) while achieving high signal-to-noise ratio (SNR) by sweeping narrow beams is proposed. Unlike the aforementioned works, the probability of losing packets and retransmissions (packet-level FEC strategy) is considered, and the number of packets transmitted within the beam is optimized. In [30], an improved beamformed broadcast/multicast technology that builds on adaptive and robust beam management techniques is introduced. The proposed technology is especially suitable for mmWave bands, where large antenna arrays are deployed at 5G NR BS.

In [31], it is stated that the service specifics implicitly prioritize multicast sessions over unicast ones in NR. Therefore, an explicit prioritization mechanism is needed at the NR BS to achieve a trade-off between unicast and multicast session drop probabilities. Later, in [32], a framework that applies stochastic geometry and queuing theory to estimate the NR AP parameters in the system where multicast and unicast traffic are simultaneously supported is proposed. In [33], 5G NR Mixed Mode (MM) to enable the use of multicast in the 5G NR Release 17 is proposed. 5G NR MM provides a flexible, dynamic, and seamless switching between unicast and multicast or broadcast transmissions and traffic multiplexing under the same radio structures. Non-orthogonal multiple access (NOMA) system is considered in [34], where a cooperation strategy for both unicast and multicast users by sharing the same time/space/frequency resource is designed. However, NOMA deployment in the 5G NR mmWave is still under discussion by 3GPP [32].

Although previous works on mmWave communications focus on data transmission optimization using multicast links, energy efficiency aspects, which are crucial for the resource-constrained devices' communication, have received insufficient coverage by the research community. This paper aims to fill this gap by addressing throughput and energy efficiency maximization of the directional multicast scheme, determined by the sequence of beamformed transmissions to serve all multicast group members. For this purpose, we develop an algorithm for multicast transmission scheduling, which utilizes adaptive[2] beamforming antennas.

III. Problem Statement

In this section, we discuss the challenges of multicast scheduling in directional systems. We then describe the scenario under analysis and the network topology. For the reader's convenience, the notation utilized throughout this paper is summarized in Table I.

A. Problem at a Glance

The use of multicast transmissions in mmWave systems is much more challenging than in traditional networks, where omnidirectional transmissions are applied [36]. Indeed, directional mmWave transmissions suffer from the limited coverage caused by oxygen absorption and severe path loss, which significantly complicates multicasting.

While utilizing the widest possible beam at high frequency severely limits data rate and transmission range, the use of only fine beams steered toward each client in unicast fashion requires a long data transmission duration. Hence, to find a trade-off between the latency and data rate, the NR AP should properly partition the multicast group into multiple subsets and select an appropriate beam and data rate to serve each subset of users. Moreover, IoT and wearable devices' low battery capabilities pose additional challenges to multicast delivery in mmWave systems.

The study of the group-oriented directional communication between devices, especially resource-constrained (i.e., wearables), is the key issue we investigate in this paper.

B. Scenario of Interest

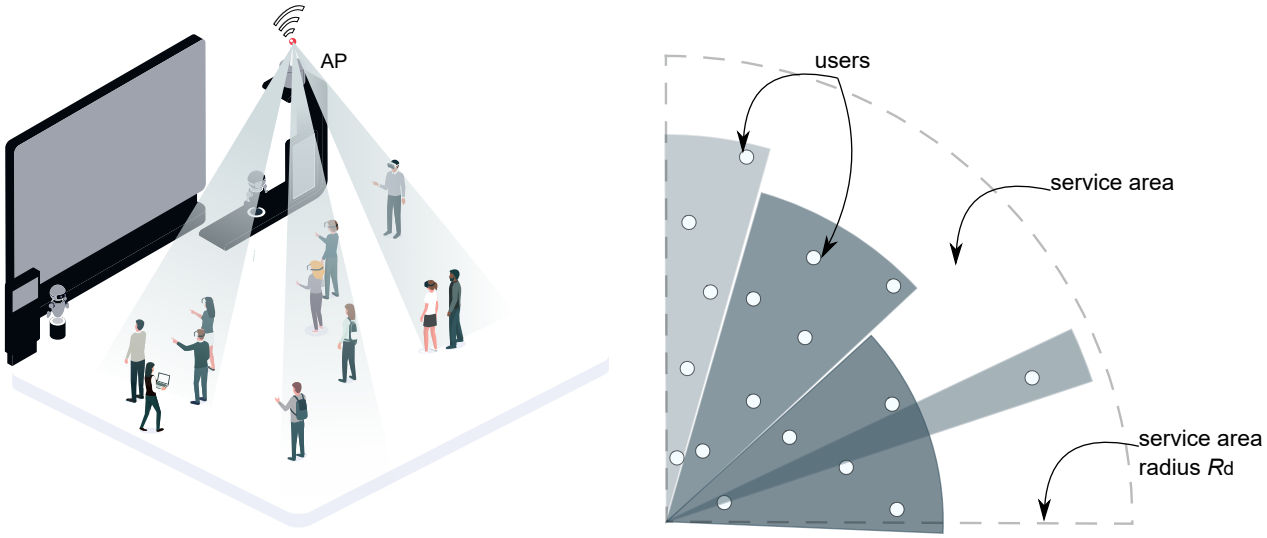
We consider a scenario where users engaged in XR applications are interested in receiving the same UHD video. XR multicasting is typical for 5G/B5G systems and generally exploits IoT terminals and wearable devices (e.g., visor wearable headsets, glasses, head-mounted displays, etc.). Furthermore, it requires high data rates and, hence, low energy consumption. Consequently, increasing the data rate while efficiently managing the device's battery life is the main objective we aim for in this paper.

All user equipment (UE) devices are provided with mmWave modules and served by B5G wireless NR AP, as

TABLE I
System modeling notation

Parameter	Definition
f	Carrier frequency
W	Bandwidth
θ	Half-power beamwidth
R_d	Radius of the area of interest
$G_{\text{rx},i}$	Receive antenna gain of device i
$G_{\text{tx},i}$	Transmit antenna gain of device i
D_0	Antenna directivity
α_i	Angular deviation from the antenna boresight of device i
$P_{\text{rx},i}$	Received power
P_{tx}	Transmit power
$M_{S,nB}$, $M_{S,B}$	Fading margins
M_I	Interference margin
r_i	Separation distance between device i and the NR AP
N	Number of users
$PL(r_i)$, $PL_{\text{dB}}(r_i)$	Path loss in linear and decibel scales
ς	Propagation coefficient
SNR_{thr}	SNR threshold
$\text{SNR}_{\max}$	Maximum SNR corresponding to choosing MCS
P_{noise}	Noise power
N_0	Power spectral density of noise
NF	Noise figure
D	Achievable data rate
P_{thr}	Receiver sensitivity
T^{DT}	Data duration transmission
B	Packet size
k	Total number of sequential beams
R	Maximum distance between a transmitter and a receiver
d	Vertical distance between two receivers
d_{drop}	Maximum vertical distance between two receivers
φ_i	Angle from OX axis for device i
φ_r	Angle from OX axis for the reference device

[2]Beamforming could be either adaptive (the beam patterns are computed on the fly based on channel feedback) or switched (precomputed beams covering 360° sequentially are used independently on users' locations) [35].



(a) Scenario of interest.

(b) Example of the beam structure with k=5 sequential multicast beams.

Fig. 1. Illustration of the system model components.

shown in Fig. 1(a). The NR AP operates in the 28 GHz band. We focus the analysis on the coverage area of a single antenna array.

In particular, the scenario under investigation is composed of a group of UEs uniformly distributed within a sector of 90° (see Fig. 1(b)). The UEs are the communication devices carried by people interested in video streaming services. The NR AP is located at the origin of the coordinates and transmits data to multiple users using a multicast wireless mmWave link. The NR AP has a coverage range of radius R_d. We assume that R_d allows all UEs inside its scope to perform reliable data transmission.

IV. System Model

In this section, we specify our system model. Namely, we describe the core components, including antenna and propagation related models, and outline the metrics of interest.

A. Antenna Model

We assume that devices transmit directionally and have the same antenna beam pattern, which is symmetrical w.r.t. the boresight [37]. Under this assumption, we mean that antennas have a unique beam shape in both elevation and azimuth planes, i.e., the antenna pattern is akin to a conical shape. For numerical tractability, we approximate the beamforming pattern as proposed in [37] with the following transmit/receive antenna gains

$$G_{\text{tx},i}G_{\text{rx},i} = D_0\rho(\alpha_i), \tag{1}$$

where D_0 corresponds to the maximum directivity along the antenna boresight, α_i is the angular deviation of the transmit/receive direction from the antenna boresight for receiver i, $i = 1,\dots,N$, N is the total number of users, $\rho(\alpha_i) \in [0;1]$ is a piece-wise linear function that scales the directivity D_0 [37][3].

[3] $\rho(\alpha) = 1$ corresponds to the antenna boresight in the case of the perfect alignment (e.g., the unicast transmission after beamforming procedure). In the case of the multicast transmission, each user deviates on angle α from the boresight of the transmitter.

B. Propagation and Blockage Models

Since we assume that, in our system, all entities transmit in directional mode, to generate a reliable model, we consider the path loss model with the inclusion of directional and beamforming antenna arrays. We also assume that pedestrians might temporarily occlude the line of sight (LoS) path between the UE and the NR AP (that is, human blockage).

To model the mmWave propagation, we utilize the 3GPP urban micro (UMi) street canyon model [38]. Accordingly, the path loss measured in dB is given by

$$PL_{\text{dB}}(r_i)=\begin{cases}32.4+21\log 10(r_i)+20\log 10(f), \text{non-blocked}\\ 47.4+21\log 10(r_i)+20\log 10(f), \text{blocked}\end{cases} \tag{2}$$

where f is the operating frequency in GHz, and r_i is the three-dimensional (3D) distance between the NR AP and the UE i.

The path loss in (2) can be rewritten in the linear scale using Ar_i^{ς}, where A and ς are propagation coefficients:

$$\begin{aligned} A_{nB} &= 10^{2\log_{10} f+3.24}M_{S,nB}M_I, \varsigma_{nB} = 2.1,\\ A_B &= 10^{2\log_{10} f+4.74}M_{S,B}M_I, \varsigma_B = 2.1. \end{aligned} \tag{3}$$

The total received signal power at the UE i is provided as

$$P_{\text{rx},i} = \frac{P_{\text{tx}}D_0\rho(\alpha_i)}{PL(r_i)M_IM_S}, \tag{4}$$

where P_{tx} is the transmit power (in Watt), $PL(r_i)$ is the linear path loss, M_I is the interference margin, and the effect of shadow fading is accounted for by using the shadow fading margins, $M_{S,B}$ and $M_{S,nB}$, for the LoS blocked and non blocked states, respectively [39]. Margins are in linear scale.

Then, the maximum achievable rate D_i of the Tx-Rx$_i$ link is expressed by Shannon Theorem:

$$D_i = W\log_2\left(1+\min\left(\frac{P_{\text{rx},i}}{P_{\text{noise}}}, \text{SNR}_{\max}\right)\right), \tag{5}$$

where P_{rx} incorporates both transmit and receive antenna gains after the beam refinement phase (BRP), $\text{SNR}_{\max}$ corresponds

to the SNR value at which the maximum MCS is selected, and the noise power, P_{noise}, is provided by

$$P_{\text{noise}} = W\, N_0\, \text{NF}, \tag{6}$$

where W is the bandwidth, N_0 is the power spectral density of noise per 1 Hz, and NF is the noise figure.

The data transmission duration (DT) T_i^{DT} of the single user i (the unicast Tx-Rx$_i$ link) is calculated by

$$T_i^{\text{DT}} = \frac{B}{D_i}, \tag{7}$$

where B is the packet size.

Since directional steerable antenna arrays and beamforming techniques are required to generate high antenna gains at EHF bands, a beam can cover only a portion of the users. As a consequence, multicast communication is performed in a sequential manner [17]. In what follows, we focus only on analog beamforming to analyze the sequential multicast performance in the TDMA fashion [17], which means that the NR AP can transmit through one beam at a time.

In the following, we indicate with the term *subgroup* the subset of users belonging to the multicast group served by the same beam. For a multicast subgroup $\mathcal{N}_j \subseteq \{1, ..., N\}$, the overall performance of the multicast transmission depends on the user with the worst channel condition. Hence, the achievable rate of the multicast transmission is

$$D_{\mathcal{N}_j} = W \log_2 \left(1 + \min_{i \in \mathcal{N}_j} \left(\frac{P_{\text{rx},i}}{P_{\text{noise}}}, 0 | P_{\text{rx},i} < P_{\text{thr}}\right)\right), \tag{8}$$

where P_{thr} represents the lower bound of the received power (receiver sensitivity) for the most robust data transmission (i.e., MCS 1). We note that the receiver sensitivity is a key parameter that impacts the system performance determining the weakest signals that can be successfully received.

The duration of the multicast data transmission of subgroup $\mathcal{N}_j$ is given by

$$T_{\mathcal{N}_j}^{\text{DT}} = \frac{B}{D_{\mathcal{N}_j}}. \tag{9}$$

C. Metrics of Interest

This paper mainly focuses on network performance improvements in terms of throughput, intending to reduce devices' energy consumption. We concentrate on the following four metrics of interest:

1) Network throughput: Network throughput (NT), or aggregate throughput, is the sum of data rates that are delivered to all terminals in the network. For sequential multicast (SM), network throughput can be written as

$$\text{NT}^{\text{SM}} = \frac{B \sum_{j=1}^{k} |\mathcal{N}_j|}{\sum_{j=1}^{k} T_{\mathcal{N}_j}^{\text{DT}}}, \tag{10}$$

where k is the number of sequential beams (multicast subgroups), $T_{\mathcal{N}_j}^{\text{DT}}$ is the data transmission duration of the subgroup $\mathcal{N}_j$ covered with beam j and corresponds to (9).

For sequential unicast fashion, we calculate NT as

$$\text{NT}^{\text{U}} = \frac{B\, N}{\sum_{i=1}^{N} T_i^{\text{DT}}}, \tag{11}$$

where T_i^{DT} is the DT duration of user i that corresponds to (7), and the denominator represents the total DT duration of all unicast users.

2) Energy consumption: Energy consumption (EC) is the amount of energy used during a given period of time. Then, the total EC of the sequential multicast data transmission is provided by

$$\text{EC}^{\text{SM}} = P_{\text{tx}} \sum_{j=1}^{k} T_{\mathcal{N}_j}^{\text{DT}}, \tag{12}$$

where energy consumption is expressed in Joules, power in Watts, and time in seconds.

For the unicast mode, EC due to the transmission is given by

$$\text{EC}^{\text{U}} = P_{\text{tx}} \sum_{i=1}^{N} T_i^{\text{DT}}. \tag{13}$$

According to Shannon's formulation, when the transmission power is constant, the better the channel quality, the higher the transmission rate. Then, more data is delivered in a given period. In contrast, in the multicast case, the faster the data transmission duration of all sequential beams, the lower the energy consumption becomes.

3) Energy efficiency: Energy efficiency (EF) is defined as the achieved network throughput divided by the consumed energy in bit/s/J [27] and evaluates how efficiently energy is used to provide a given network throughput. In the case of the sequential multicast, EF is calculated as

$$\text{EF}^{\text{SM}} = \frac{\text{NT}^{\text{SM}}}{\text{EC}^{\text{SM}}} = \frac{B \sum_{j=1}^{k} |\mathcal{N}_j|}{P_{\text{tx}} \left(\sum_{j=1}^{k} T_{\mathcal{N}_j}^{\text{DT}}\right)^2}. \tag{14}$$

For the unicast mode, EF can be then written as

$$\text{EF}^{\text{U}} = \frac{\text{NT}^{\text{U}}}{\text{EC}^{\text{U}}} = \frac{B\, N}{P_{\text{tx}} \left(\sum_{i=1}^{N} T_i^{\text{DT}}\right)^2}. \tag{15}$$

4) User throughput: Further, we concentrate on the user-side metric of interest, such as user throughput (THR). User throughput is a term used for the determination of the amount of data transferred from the NR AP to the user within a given time. For sequential multicast, user throughput is provided by

$$\text{THR}^{\text{SM}} = \frac{B}{\sum_{j=1}^{k} T_{\mathcal{N}_j}^{\text{DT}}}. \tag{16}$$

For the unicast transmission, user throughput can be expressed as

$$\text{THR}^{\text{U}} = \frac{B}{\sum_{i=1}^{N} T_i^{\text{DT}}}. \tag{17}$$

We note that, in the case of pure multicast, the number of sequential beams is constant and equal to one. Hence, for multicast (M), we calculate NT$^{\text{M}}$, EC$^{\text{M}}$, EF$^{\text{M}}$, and THR$^{\text{M}}$ according to (10), (12), (14), and (16), respectively, with $k = 1$.

V. Efficient Management of Multicast Traffic

In this section, we define the policies and potential thresholds for the switching between unicast and multicast transmissions. Then, we develop a dynamic resource management algorithm in directional mmWave networks that determines the number and resolution of the beams required to deliver the content to all interested users as, in mmWave, multicasting is performed through multiple sequential transmissions.

A. Beam Switching Policy

We consider the scenario where two devices are located at the edges of a line of length d that lays at distance r from the NR AP (see Fig. 2). We analyze the coverage area of the beam and investigate the maximum possible distance d (determined by d_{drop}) for multicast (as directional beams are limited in the coverage area) in pursuance of the required QoS characteristics in terms of the achievable data rate to define policies and potential thresholds for the switching between unicast and multicast transmissions.

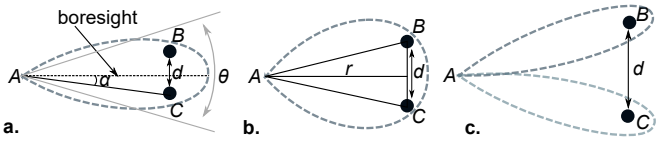


Fig. 2. Illustration of distances d and r for: (a) multicast with the narrow beam, (b) multicast with the wide beam, and (c) sequential unicast transmissions.

To this aim, we model the coverage area of a directional antenna as a *drop* and analytically define distance d_{drop} between the two edge devices at which the minimum required data rate can be guaranteed for a given MCS and distance r. More precisely, the drop periphery determines the zone wherein users can be deployed to receive the signal (see Fig. 3). In this case, distance d_{drop} is calculated from the drop's coordinates (see Algorithm 1) and acts as a tool for the NR AP to determine the maximum beam pattern θ to be swept toward the user at distance r.

Adhering introduced notations, we define the equation in rectangular coordinates, which describes the drop as

$$(x^2 + y^2)^2 = (x^2 - y^2)R^2. \tag{18}$$

In polar coordinates, the region boundary of the drop can then be written as

$$q = \sqrt{1 - 2\sin^2\theta}R, \tag{19}$$

where R is defined as

$$R = \left(\frac{P_{\text{tx}} D_0}{A_B M_I M_{S,B} P_{\text{thr}}}\right)^{\frac{1}{\zeta}}. \tag{20}$$

The reason for using the blocked link in (20) is to determine the service area radius within which no user experiences outage conditions when its LoS link is blocked.

In Algorithm 1 (named *Beam Switching Policy*), equation (19) is used to build the coverage area by taking the half power beam width (HPBW) of the antenna into account. Then, for the sake of simplicity, we proceed with the Cartesian coordinate dimensioning. After converting the coordinates, which define the drop surface, the algorithm calculates d_{drop} from y-coordinate of the drop by taking it twice (lines 7-9) and determines if beam switching is needed (lines 10-13). In the case when distance $d(r)$ exceeds the d_{drop} threshold, beam switching is required. Otherwise, the NR AP continues serving UEs without any modifications on the beam management.

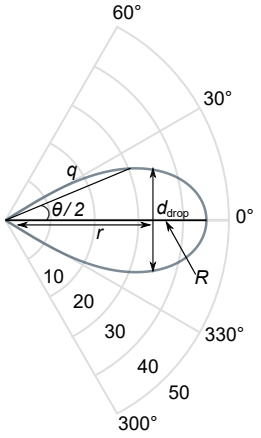


Fig. 3. Illustration of the drop-based coverage area approximation.

Algorithm 1: Beam Switching Policy

1 **Input:** θ, r
2 **Output:** decision on beam switching
3 **construct** the coverage area of a beam as a drop (19);
4 **convert** polar coordinates into Cartesian ones:
5 $x = q\cos(\theta/2)$;
6 $y = q\sin(\theta/2)$;
7 **for** *all x, which corresponds to r,* **do**
8 find $2y$;
9 set $d_{\text{drop}} \leftarrow 2y$
10 **match** d_{drop} with current d: ▷ $x = r$
11 **if** $d_{drop}(x = r) < d(r)$ **then**
12 perform the beam switching to *(i)* a wider beam (if r allows) or *(ii)* to a high number of narrow beams;
13 **end**

B. Multicast Radio Resource Management Policy

This section describes the proposed dynamic RRM algorithm for the multicast data delivery in directional mmWave networks that determines the number and resolution (i.e., width) of the beams required to optimize the performance of the multicast transmission. The number of beams k required to serve all users ranges from one to at most N beams, where N is the number of users. The pseudo-code of the proposed multicast RRM policy is presented in Algorithm 2.

We denote the set of users awaiting the multicast traffic as A. We select multicast subgroup $\mathcal{N}_j$ to be served by the NR AP with beam θ_j. We also use the distance vector $r_j = \{r_{j_1}, r_{j_2}, .., r_{j_i}, .., r_{j_N}\}$, each element thereof represents the

Algorithm 2: Multicast RRM Policy

1 **initialize** set of users waiting for the multicast data transmission: $A \leftarrow \{1, ..., N\}$;
2 **initialize** sequential beam counter: $j \leftarrow 0$;
3 **initialize** distance vector between the NR AP and devices: $r_j = \{r_{j_1}, r_{j_2}, .., r_{j_i}, .., r_{j_N}\}$;
4 **initialize** $\text{MAX}_Q \leftarrow 0$;
5 **initialize** $T^{\text{DT}}_{\mathcal{N}} \leftarrow 0$; ▷ $\mathcal{N} = \{1, ..., N\}$
6 **while** $A \neq \emptyset$ **do**
7 $j \leftarrow j + 1$:
8 find $r^{\max}_j(r_i)$, $i \in A$;
9 set r to $r^{\max}_j(r_i)$;
10 **for** $\theta^{\circ}_{\max}(d_{drop}, r)$ *to* $\theta^{\circ}_{\min}$ *increments by* δ° **do**
11 $\mathcal{N}_\theta = \varphi_r - \theta/2 \leq \varphi_i \leq \varphi_r + \theta/2$; ▷ *$\varphi_r$ is the angle from OX axis for the reference device, whereas φ_i is the angle from OX axis for device i*
12 calculate Q_θ by using (21) or (22);
13 **if** $MAX_Q < Q_\theta$ **then**
14 $\text{MAX}_Q \leftarrow Q_\theta$;
15 $\mathcal{N}_j \leftarrow \mathcal{N}_\theta$;
16 $\theta_j \leftarrow \theta$;
17 **end**
18 **end**
19 $A \leftarrow A \setminus \mathcal{N}_j$;
20 $T^{\text{DT}}_{\mathcal{N}} \leftarrow T^{\text{DT}}_{\mathcal{N}} + T^{\text{DT}}_{\mathcal{N}_j}$;
21 **end**
22 **return** j, $T^{\text{DT}}_{\mathcal{N}}$;
23 **calculate** NT (10), EC (12), EF (14), THR (16).

distance between the NR AP and user i, where i is the index of the user. The algorithm iteratively partitions users of set A into multiple subgroups, as indicated by line 6. The algorithm starts with choosing the farthest user from the set A; the distance between that user and the NR AP is denoted as r (lines 8-9).

In the RRM, we use adaptive beamforming, and, at any time instant, one beam pattern can be selected to transmit with a chosen MCS (i.e., transmission rate) depending on the user's location. In general, the directional beamwidth can be adjusted within the quasi-omnidirectional range, which is 180°. We highlight that not all values of θ (line 10) from 0° to 180° need to be analyzed because, for the MCS required to deliver data to the user at distance r, the NR AP guarantees the required QoS in terms of achievable data rate within a vertical distance d_{drop}, as discussed in Section V-A and shown in Algorithm 3. Therefore, by applying the *drop-based* approach, we reduce the computational complexity of the algorithm and, thus, the time required for the multicast beam sweeping as d_{drop} determines the maximum beamwidth to be steered toward users located at distance r.

Line 11 (Algorithm 2) collects all users covered by beam θ steered toward the device with distance r in the multicast subgroup $\mathcal{N}_\theta$.

The proposed RRM scheme has been designed in order to be flexible. In fact, the optimization function to be used (see line 12) can be selected according to the goal that the network operator aims to achieve. In this paper, we analyze the performance of the proposed RRM under the following optimization functions:

Algorithm 3: Maximum Beamwidth Determination

1 **Input:** separation distance r;
2 **Output:** $\theta^{\circ}_{\max}(d_{\text{drop}}, r)$;
3 **initialize** set of antenna patterns: $\Theta \leftarrow \theta^{\circ}_{\min} : \delta^{\circ} : \theta^{\circ}_{\max}$;
4 set $\theta^{\circ}_{\max}(d_{\text{drop}}, r) \leftarrow \theta^{\circ}_{\max}$;
5 calculate $d_{\text{drop}}(r)$ ▷ *by using Algorithm 1*
6 **if** $d_{drop}(r) > 0$ **then**
7 go to step 12;
8 **else**
9 set $\theta^{\circ}_{\max}(d_{\text{drop}}, r) \leftarrow \theta^{\circ}_{\max}(d_{\text{drop}}, r) - \delta^{\circ}$;
10 go to step 5;
11 **end**
12 **return** $\theta^{\circ}_{\max}(d_{\text{drop}}, r)$.

- *EF maximization*:

$$\begin{gathered} \text{maximize}\, Q_\theta = \frac{B\,|\mathcal{N}_\theta|}{T^{\text{DT}}_{\mathcal{N}_j}} \frac{1}{P_{\text{tx}} T^{\text{DT}}_{\mathcal{N}_j}}, \\ \text{subject to } \theta^{\circ}_{\min} \leq \theta \leq \theta^{\circ}_{\max}(d_{\text{drop}}, r), \\ r \leq R, \end{gathered} \tag{21}$$

  where $|\mathcal{N}_\theta|$ is the size of the multicast subgroup covered with beam θ.
- *NT maximization*:

$$\begin{gathered} \text{maximize}\, Q_\theta = |\mathcal{N}_\theta|\, D_\theta, \\ \text{subject to } \theta^{\circ}_{\min} \leq \theta \leq \theta^{\circ}_{\max}(d_{\text{drop}}, r), \\ r \leq R. \end{gathered} \tag{22}$$

Generally, wide beams can cover a larger angle range and may simultaneously serve more users. However, due to the lower antenna gain that wide beams provide, the supported transmission rate is limited. Inversely, narrow beams provide higher antenna gain and, thus, can support higher transmission rates. However, they are limited in the coverage in terms of the aperture angle and may not serve many users simultaneously. Since a data rate increase causes a decrease in the covered distance, a high data rate leads to a system throughput reduction as the number of multicast devices covered with the same beam decreases. We analyze the performance of different (narrow and wide) mmWave beams in Section VI.

Therefore, depending on the number of users, their locations, and density, different beam resolutions optimize the performance for multicast subgroup $\mathcal{N}^\theta$ according to the optimization function. Line 12 describes the Q_θ calculation for multicast subgroup $\mathcal{N}^\theta$. Lines 13-18 demonstrate the process of selecting the beam that maximizes Q_θ. The algorithm stops when all users have been served.

C. Complexity Analysis

The computational complexity of the proposed algorithm is given by

$$O(|A| \cdot |\Theta|),$$

where $|A|$ is the complexity due to the "while" cycle over all $|A|$ users in the worst case of the unicast transmission (lines 6-21). This means that each beam j covers only a single user. For the second component, which is inside the "while" cycle, $|\Theta|$ is the complexity due to the tree traversing from $\theta_{\max}$ to $\theta_{\min}$ with increments by δ (lines 10-18). In the worst case, traversal requires $|\Theta| = \theta_{\max}/\delta$ attempts (in the case, when $\theta_{\max} = 180°$, and $\theta_{\min}$ and δ are fixed). However, the upper limit on $\theta_{\max}$ is defined by the drop-based approach and, therefore, the number of attempts is significantly reduced at this "for" cycle. We note that the sequential execution of Algorithm 3 and lines 10-18 of Algorithm 2 has the following complexity: $O(|\Theta|) + O(|\Theta|) = O(|\Theta| + |\Theta|) = O(|\Theta|)$. As a result, in the worst case, the number of operations is in $O(|A| \cdot |\Theta|))$. However, in practice, the multicast scheme requires a significantly lower number of attempts as *(i)* one beam can serve more than one user at a time, and *(ii)* the drop scheme substantially decreases the number of beamwidths while searching for the optimal solution. We note that a reduced number of operations also means a lower execution time.

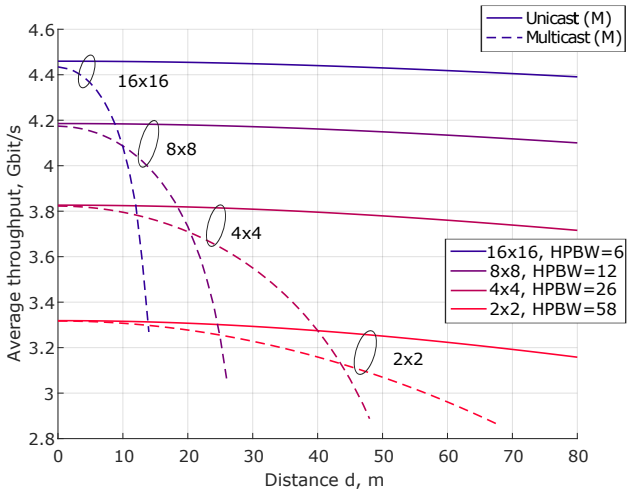


Fig. 4. Average throughput vs. distance d for different phased antenna arrays of antenna elements and HPBWs, i.e., 2x2 (HPBW=6°), 4x4 (HPBW=12°), 8x8 (HPBW=26°), and 16x16 (HPBW=58°), $r = 60$ m, MCS $= 1$. Matlab sensor array analyzer is used to generate the radiation patterns ('M').

VI. Performance Analysis

In this section, we analyze the performance of the drop-based approach exploited for implementing the beam switching policy and the proposed RRM scheme. To this aim, we develop a simulation environment in MATLAB that accepts the input parameters listed in Table II.

A. Analysis of the Drop-based Approach

We recall that we apply the *drop-based* approach to approximate the coverage area of a single antenna. Variables r and d represent, respectively, the distance between the NR AP and the *line* of users and the distance between the two edge devices of the line.

In Fig. 4, we show the performance in terms of the average throughput for unicast and multicast transmission modes under the increasing distance between users (i.e., d). In particular, we set $r = 60$ and consider only two users placed at the edges of the line. One may see that the unicast transmission mode is almost unaffected from d because the beams can be steered towards the users. Furthermore, the higher throughput values are achieved when narrower beams are used (see $\theta = 6°$ with respect to $\theta = 58°$). Differently, multicast users undergo a significant data rate reduction with low values of d, and this reduction is more significant for the narrower beams. This motivates our work to investigate when beam switching is required to prevent a decrease in network performance.

In Fig. 5, we analyze the correlation between the length of the line of devices (i.e., d on the y-axis) and its distance from the NR AP (i.e., r on the x-axis) as a function of θ for different beamwidths. We compare the drop-based approach (see Eq. 19) with the simulation results. By looking at the plot, we can appreciate that all curves follow the same trend: they start to rise, then reach a peak and fall slightly faster. This behavior can be explained by directional antenna properties. More precisely, the rise of distance d accounts for the coverage angle of the antenna. Until the antenna gain is sufficient to overcome the path loss in EHF bands, the curves show an increasing trend. When propagation properties cause high path loss, the curves start to fall and reach zero on the y-axis when the signal cannot be received.

Furthermore, in Fig. 5, we validate the behavior of the analytical model as per Eq. (19) by comparing achieved results to simulation when using *(i)* an isotropic antenna with no tapering generated by Matlab antenna array analyzer ('Sim (M)'), and *(ii)* a linear function of the beamwidth, as proposed in [37] ('Sim (L)'). As one may notice from Fig. 5, the analytical model exhibits a trend close to the simulation results but with a shift toward the right side. This displacement is small for low values of r, whereas it rises when the distance from the NR AP increases, especially for small values of HPBW. The gap between analysis and simulation curves can be explained by the form of the antenna. In practice, the

TABLE II
Default system parameters for numerical assessment

Notation	Parameter	*Value*
f	Carrier frequency	28 GHz
N	Number of users	20
h_A	Height of AP	4 m
h_B	Height of blocker	1.7 m
h_U	Height of UE	1.5 m
r_B	Blocker radius	0.4 m
λ_B	Density of blockers	0.1 bl./m^2
SNR_{thr}	SNR threshold	-9.478 dB
P_{tx}	Transmit power	33 dBm
$M_{S,nB}$, $M_{S,B}$	Fading margins	4/8.2 dB
M_I	Interference margin	3 dB
θ	Beamwidth	var
R_d	Radius of the area of interest	50 m
$\text{SNR}_{\max}$	SNR corresponding to choosing MCS15 (rate 948/1024)	20 dB
NF	Noise figure	7.6 dB
N_0	Power spectral density of noise	174 dBm/Hz
s_{PRB}	PRB size	1.44 MHz
B	Packet size	1 Gb

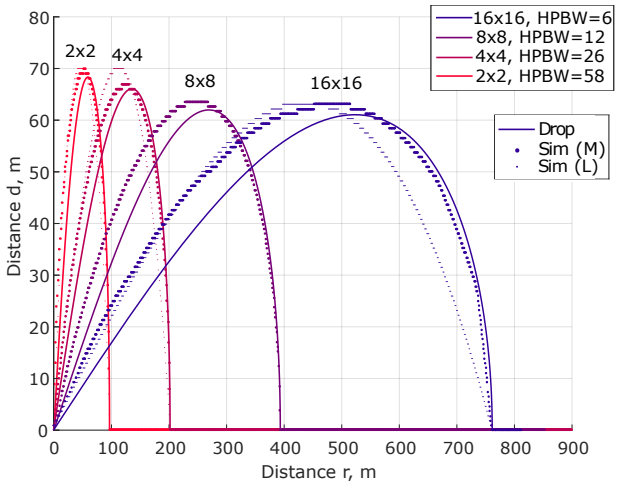


Fig. 5. Dependence between r and d for *(i)* the drop-shaped coverage area, *(ii)* simulation results for phased antenna arrays of 2x2 (HPBW=6°), 4x4 (HPBW=12°), 8x8 (HPBW=26°), and 16x16 (HPBW=58°) elements obtained by using Matlab antenna array analyzer ('M') and *(iii)* a piece-wise linear function of α misalignment ('L') [37].

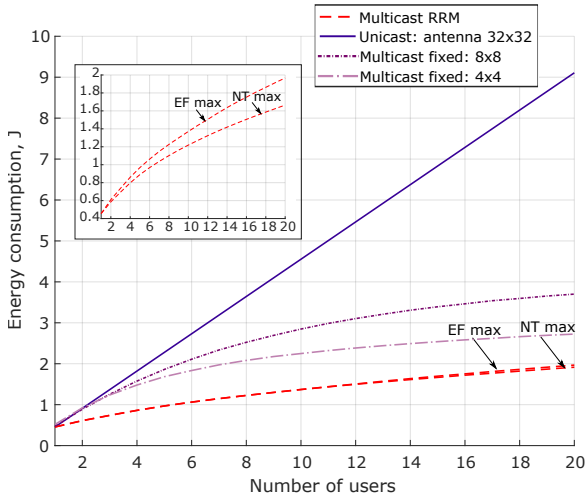


Fig. 6. Energy consumption vs. number of users for the proposed multicast RRM policy, sequential unicast, and sequential multicast with predefined beams.

antenna pattern has no concrete shape. Moreover, from the simulation curves, we can learn that the form of the coverage area of wide beams (e.g., $\theta = 58°$) resembles the drop, whereas it reminds a pencil in the case of narrow beams.

Fig. 5 highlights that the drop-based approximation perfectly matches with the simulation results ('M') when the distance d decreases after having reached the peak, and acts as a lower bound in the remaining parts. We emphasize that the type of antenna elements (e.g., isotropic, cosine) utilized with specialized properties for particular applications and the tapering, which is the manipulation of the amplitude contribution of an element to the overall antenna response [40] (e.g., Chebyshev, Taylor, etc.), among other parameters, affect the form of the coverage area of the directional antenna.

Finally, plots in Fig. 5 allow to determine, for a given value of r, the maximum value of d and investigate multicast in mmWave with reference to the coverage area. In particular, if current distance d, at a distance from the NR AP r, exceeds the threshold value depicted in Fig. 5 (e.g., d_{drop}), the network should: *(i)* create a wider beam (if any), or *(ii)* switch the beam to more narrower beams. For example, for $\theta = 6°$ and at $r = 100$ m, the distance d, based on the drop model, should not be higher than 16 m to support MCS 1.

B. Performance Analysis of the Proposed RRM Policy

We analyze the system performance according to the metrics of interest and evaluate them using a Montecarlo approach by running 10^6 simulations. We compare the proposed RRM scheme to sequential unicast and sequential multicast (with predefined beams) schemes. We randomly and uniformly distribute N users within a sector, as shown in Fig. 1(b). The radius of the service area R_d is 50 m.

In our scenario, we also assume that users can be blocked by the mobile crowd. Blockers are modeled as cylinders with constant base radius r_B and constant height h_B, whereas the attenuation due to the human blockage is assumed to be 15 dB. Other system parameters are provided in Table II. We analyze the results achieved by the proposed RRM using two different optimization functions described in subsection V-B, that is: *(i)* EF maximization and *(ii)* NT maximization. We emphasize that we apply the optimization function to each pure multicast transmission within the sequential multicast scheme.

To verify the effectiveness of the proposed method, we compare the results of RRM to baseline solutions of interest: multicast and unicast schemes. In the literature, the multicast scheme, where the most remote devices in each sector determine the data rates for covering one sector [35], [41] is referred to as the enhanced multicast (EM) scheme. As a benchmark solution, we consider the EM scheme with a slight modification on the beam angle. To cover a larger area, we assume a sector covered by four and eight beams of 26° and 12°, respectively. In the following, we refer to this approach as the multicast with fixed beams for the sake of convenience. In the unicast scheme, the NR AP serially transmits the data to each user using the fine beam with the resolution of $\theta = 2°$.

We emphasize that the difference in terms of the overhead among the compared schemes is small. Furthermore, the time overhead is significantly lower than the duration of the data transmission. Therefore, the additional time overhead for beam training has only a marginal impact on the overall throughput [26]. For these reasons, in this paper, we mainly evaluate transmission-related performance.

We start by comparing the energy consumption of the proposed RRM scheme under the two analyzed optimization functions with unicast and multicast with fixed beams, as illustrated in Fig. 6. As one may observe, EF and NT maximization strategies outperform benchmark multicast and unicast approaches and show lower energy consumption values. We recall that the number of sequential beams and their resolution has an impact on energy consumption. Note that the smaller number of antennas (wide HPBW) provides lower data transmission delay due to the sequentiality but, at the

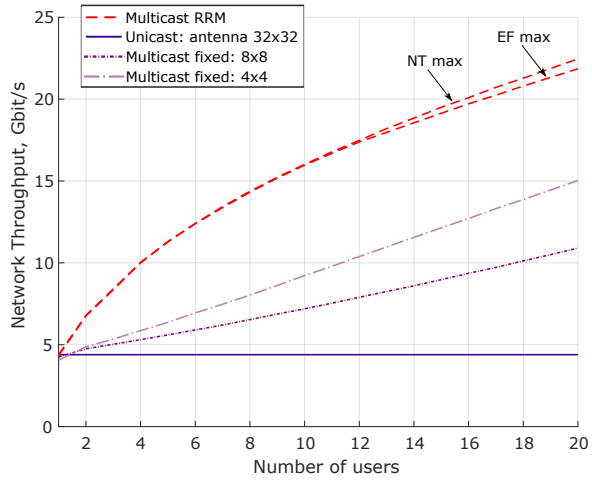


Fig. 7. Network throughput vs. number of users for the proposed multicast RRM policy, sequential unicast, and sequential multicast with predefined beams.

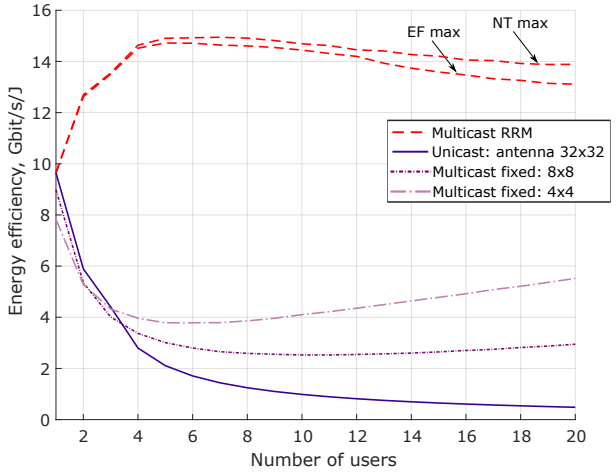


Fig. 8. Energy efficiency vs. number of users for RRM multicast, sequential unicast, and sequential multicast with predefined beams.

same time, it provides a less directional transmission, which leads to a smaller radius of the coverage area and lower data rate. Hence, the trade-off between the number of beams and their directivity gains is a challenging aspect. Analyzing the plot, we can learn that even though multicast with four fixed beams of 26° has a small number of sequential transmission, it provides lower data rates, significantly impacting the system energy consumption. Concerning the unicast transmission, due to its sequential nature and the use of highly directional narrow beams (e.g., $\theta = 2°$), it provides the highest EC value.

The plot illustrating the performance of the proposed and benchmarks schemes in terms of network throughput is provided in Fig. 7. Results are consistent with those reported in Fig. 6. As the use of wider beams ensures shorter total data transmission duration due to the reduced number of sequential operations, the strategies that utilize wide beams (e.g., NT and EF maximizations, multicast with fixed beams) show higher NT value compared to unicast. For NT and EF maximization, the proposed RRM policy prefers wider beams because they can serve more users at a time and provide high NT per beam. If the radius R_d of the considered area is too large, the gain of the wide beam might be insufficient to cover the most distant user. We also underline that the RRM algorithm allows capturing the trade-off between the number of transmissions and the SNR they provide, which is a beneficial feature compared to the benchmark approaches.

We continue by analyzing energy efficiency for the same group of transmission schemes, as depicted in Fig. 8. We recall that EF is defined as network throughput divided by the consumed energy. As one may observe, the proposed multicast RRM with NT and EF maximization outperforms all other schemes in terms of EF. Specifically, NT maximization demonstrates a high gain in the energy efficiency of up to 96.43%, 60.71%, and 78.57% over unicast and multicast with 4x4 and 8x8 antenna elements, respectively.

As discussed in subsection V-B, in our proposal, we search for the beam that optimizes our metrics of interest, using

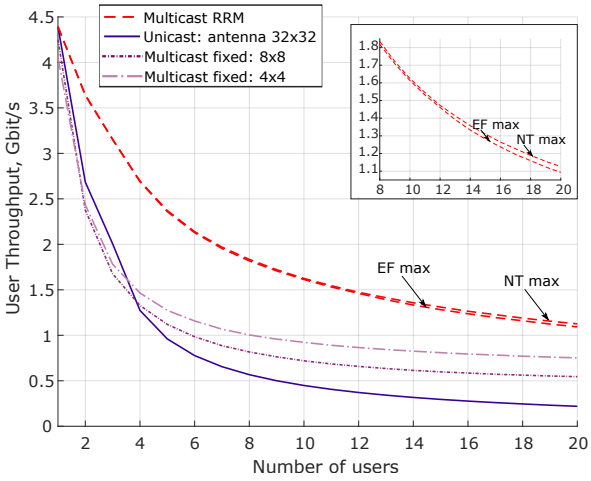


Fig. 9. User throughput vs. number of users for RRM multicast, sequential unicast, and sequential multicast with predefined beams.

reverse order from $\theta_{\max}(d_{\text{drop}}, r)$ to $\theta_{\min}$. To compare the proposed RRM with multicast schemes that exploit predefined beams, we select an area with a radius R_d such that all schemes can serve all users. As the instantaneous data rate is constrained by the selected MCS ($\text{SNR}_{\max}$), different beams can provide the same data rate (e.g., for lower R_d values). In this case, the RRM algorithm picks the first beam in the list if all beams guarantee the same performance. We emphasize that the tree traversal order does not produce any difference in NT/EF maximization.

Now we analyze a user-side metric, namely, user throughput (see Fig. 9). As one may notice, the RRM with NT/EF optimization functions reveals the highest user throughput. Such dominated behavior can be explained by the number of sequential beams used for the transmission. The less number of beams the NR AP uses, the better the user-perceived throughput. We emphasize that the proposed algorithm outperforms both unicast and multicast benchmark schemes for

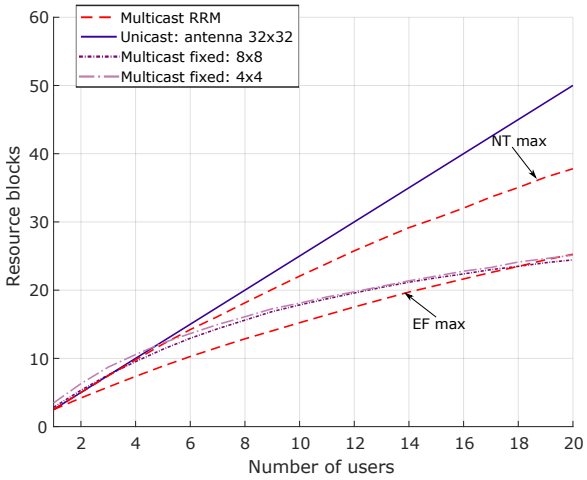


Fig. 10. Number of resource units vs. number of users for RRM multicast, sequential unicast, and sequential multicast with predefined beams.

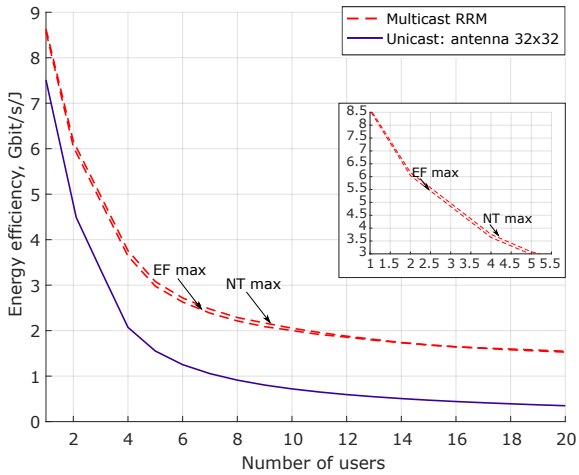


Fig. 11. Energy efficiency vs. number of users for RRM multicast and sequential unicast, $R_d = 250\,\text{m}$.

all considered metrics of interest, which we examine to ensure energy-efficient communication between users' devices.

C. Resource Utilization Analysis

We proceed by comparing the proposed schemes with the benchmark ones from the resource utilization point of view. To this end, we assume that the session requires constant bitrate R_s, i.e., 20 Mbps. Technically, to determine the number of resources required from NR AP to serve a session with the chosen bitrate, we have to know the channel quality indicator (CQI), MCS, and spectral efficiency (SE) values as well as SNR, SE to CQI mapping. As these parameters are usually vendor-specific, in our study, we use MCS mappings from [42]. Then, the number of physical resource blocks (PRBs) for multicast subgroup $\mathcal{N}_j$ is calculated as $\text{PRB} = \frac{R_s}{\text{SE}_{\mathcal{N}_j}\, s_{\text{PRB}}}$, where s_{PRB} is the bandwidth of the PRB and $\text{SE}_{\mathcal{N}_j}$ is the spectral efficiency.

The results provided in Fig. 10 indicate that, for given system parameters, the EF maximization strategy shows the best system resource utilization and thereby outruns all schemes when the number of users is lower than 17. Further, analyzing the plot, we can learn that unicast transmissions always demand a higher number of resources compared to all other approaches. The reason is that, for multicast, the SNR of each subgroup and the number of subgroups affect the amount of resources delivered to all multicast users. In contrast, in the case of unicast, the SNR value of each user affects the total amount of requested resources. However, as the radius of the service area is chosen in such a way as to cover all users with fixed beams in a multicast fashion, all multicast users experience at least the minimum SNR value needed for the multicast reception. This leads to the assignment to all multicast members of the lower number of resources required by the multicast user with the worst channel conditions. Here, the sequentiality of the unicast transmission has a substantial impact on the number of resources, despite the highest quality of the channel. We also highlight that both multicast transmission schemes with fixed beams demonstrate similar behavior and yield comparable with the EF maximization results.

D. Impact of System Parameters on RRM Performance

Finally, we assess the performance of the proposed algorithm in a larger coverage area w.r.t. the one under consideration in the previous analysis (i.e., R_d is set to 250 m rather than to 50 m). We recall that the lower the angle of the directional antenna, the larger the covered distance. In Fig. 11, we demonstrate the performance achieved in terms of energy efficiency. We can appreciate that the benefit of the proposed multicast RRM schemes with respect to unicast is reduced compared to the case of the small coverage area (see Fig. 8). This behavior is reasonable since to cover users located close to the border of the large area, the NR AP needs to sweep narrow beams. This, in turn, leads to a higher required number of sequential beams (similar to unicast). However, we highlight that the proposed RRM under NT maximization still performs better than unicast up to 80%.

We can conclude that NT/EF maximization offers higher energy efficiency compared to baseline schemes. We also point out that the network throughput of each beam plays a significant role in reducing the overall energy consumption. The reason behind the fact that NT and EF maximization functions achieve comparable performance in most of the presented results lies in followed approach towards optimization. More precisely, we apply EF/NT maximization to each beam (multicast subgroup) sequentially. In the case of NT, the algorithm preferably chooses wide beams, which leads to a higher throughput as the beam serves more users, providing a lower total energy consumption and, hence, higher energy efficiency. Conversely, EF indicates how effectively energy is used to achieve a given network throughput. Also in this case, the employed algorithm tends to preferably select wide beams. Thus, we can conclude that sequential optimization affects the results and leads to similar performance for both EF and NT maximization schemes.

The presented results show that that NT optimization function

is a better choice for all considered metrics of interest (that is, energy efficiency, energy consumption, network and user throughput), whereas EF demonstrates a better performance in terms of resource utilization.

VII. Conclusions and Future Work

Supporting multicast traffic and ensuring energy-efficient communications in mmWave systems is a crucial challenge to be faced in 5G systems. Indeed, multicast and broadcast are identified as one of the essential features in future systems to support efficient multimedia distribution due to their ability to increase coverage and support flexible resource multiplexing among users. Moreover, limiting the energy consumption of devices is essential for future advanced communications. In this work, we developed a solution that provides optimal radio resource management for multicast traffic delivery in mmWave networks while reducing devices' energy consumption.

Achieved results provide three practical outcomes. First, the drop-based approach offers a good approximation of the coverage area of the antenna and can serve as a tool to determine when the beam switching is required for multicast transmissions. Second, the proposed RRM scheme works for any area of interest in terms of the distance and coverage angle, whereas multicast with wide predefined beams fails to cover users located far from the NR AP (due to the propagation loss). Finally, we state that the network throughput of every single beam in sequential multicast transmissions should be maximized to ensure energy-efficient communication desirable for battery-limited devices, such as IoT and wearable.

An interesting topic for future research is hybrid or digital beamforming, which allows transmitting to more than one user at a time over multiple beams. However, when multiple RF chains are available, along with new opportunities, several challenges appear. For example, the shape of numerous beams to be swept simultaneously under the total transmission power constraint per antenna has to be properly selected. This means that, in addition to the beam orientation and the beam resolution (beamwidth) adjustment, the transmit power has to be split among beams, which is of particular interest for our future work.

Acknowledgment

The authors gratefully acknowledge funding from European Union's Horizon 2020 Research and Innovation programme under the Marie Skłodowska Curie grant agreement No. 813278 (A-WEAR: A network for dynamic wearable applications with privacy constraints, http://www.a-wear.eu/).

References

[1] G. H. Sim, A. Loch, A. Asadi, V. Mancuso, and J. Widmer, "5G Millimeter-Wave and D2D Symbiosis: 60 GHz for Proximity-Based Services," *IEEE Wireless Communications*, vol. 24, no. 4, pp. 140–145, 2017.

[2] G. Sanfilippo, O. Galinina, S. Andreev, S. Pizzi, and G. Araniti, "A Concise Review of 5G New Radio Capabilities for Directional Access at mmWave Frequencies," in *Internet of Things, Smart Spaces, and Next Generation Networks and Systems*, pp. 340–354, Springer, 2018.

[3] G. T. 38.808, "Study on Supporting NR from 52.6 GHz to 71 GHz (Release 17)," 2020.

[4] P. Zhou, K. Cheng, X. Han, X. Fang, Y. Fang, R. He, Y. Long, and Y. Liu, "IEEE 802.11 ay-based mmWave WLANs: Design Challenges and Solutions," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 3, pp. 1654–1681, 2018.

[5] M. R. Akdeniz, Y. Liu, M. K. Samimi, S. Sun, S. Rangan, T. S. Rappaport, and E. Erkip, "Millimeter Wave Channel Modeling and Cellular Capacity Evaluation," *IEEE journal on selected areas in communications*, vol. 32, no. 6, pp. 1164–1179, 2014.

[6] X. Lu, E. Sopin, V. Petrov, O. Galinina, D. Moltchanov, K. Ageev, S. Andreev, Y. Koucheryavy, K. Samouylov, and M. Dohler, "Integrated Use of Licensed-and Unlicensed-band mmWave Radio Technology in 5G and Beyond," *IEEE Access*, vol. 7, pp. 24376–24391, 2019.

[7] Y. Ghasempour, C. R. da Silva, C. Cordeiro, and E. W. Knightly, "IEEE 802.11 ay: Next-generation 60 GHz Communication for 100 Gb/s Wi-Fi," *IEEE Comm. Magazine*, vol. 55, no. 12, pp. 186–192, 2017.

[8] K. Sakaguchi, T. Haustein, S. Barbarossa, E. C. Strinati, A. Clemente, G. Destino, A. Pärssinen, I. Kim, H. Chung, J. Kim, *et al.*, "Where, when, and How mmWave is Used in 5G and Beyond," *IEICE Transactions on Electronics*, vol. 100, no. 10, pp. 790–808, 2017.

[9] K. Zeman, M. Stusek, J. Pokorny, P. Masek, J. Hosek, S. Andreev, P. Dvorak, and R. Josth, "Emerging 5G Applications over mmWave: Hands-on Assessment of WiGig Radios," in *2017 40th International Conference on Telecommunications and Signal Processing (TSP)*, pp. 86–90, IEEE, 2017.

[10] S. Paasovaara and T. Olsson, "Proximity-Based Automatic Exchange of Data in Mobile Gaming: Studying the Experiences of Streetpass Users," in *Proceedings of the 9th Nordic Conference on Human-Computer Interaction*, pp. 1–10, 2016.

[11] O. Vikhrova, S. Pizzi, A. Molinaro, A. Iera, K. Samuylov, and G. Araniti, "Group-based Delivery of Critical Traffic in Cellular IoT Networks," *Computer Networks*, p. 107563, 2020.

[12] F. Rinaldi, S. Pizzi, A. Orsino, A. Iera, A. Molinaro, and G. Araniti, "A Novel Approach for MBSFN Area Formation Aided by D2D Communications for eMBB Service Delivery in 5G NR Systems," *IEEE Trans. on Vehicular Technology*, vol. 69, no. 2, pp. 2058–2070, 2019.

[13] A. De La Fuente, J. J. Escudero-Garzás, and A. García-Armada, "Radio Resource Allocation for Multicast Services Based on Multiple Video Layers," *IEEE Trans. on Broadcasting*, vol. 64, no. 3, pp. 695–708, 2017.

[14] S. Pizzi, C. Suraci, A. Iera, A. Molinaro, and G. Araniti, "A Sidelink-Aided Approach for Secure Multicast Service Delivery: From Human-Oriented Multimedia Traffic to Machine Type Communications," *IEEE Transactions on Broadcasting*, 2020.

[15] W. Wu, N. Cheng, N. Zhang, P. Yang, W. Zhuang, and X. Shen, "Fast mmWave Beam Alignment via Correlated Bandit Learning," *IEEE Trans. on Wireless Comm.*, vol. 18, no. 12, pp. 5894–5908, 2019.

[16] S. Yang, C. Xu, L. Zhong, J. Shen, and G.-M. Muntean, "A QoE-Driven Multicast Strategy With Segment Routing—A Novel Multimedia Traffic Engineering Paradigm," *IEEE Transactions on Broadcasting*, vol. 66, no. 1, pp. 34–46, 2019.

[17] A. Biason and M. Zorzi, "Multicast via Point to Multipoint Transmissions in Directional 5G mmWave Communications," *IEEE Communications Magazine*, vol. 57, no. 2, pp. 88–94, 2019.

[18] H. Park and C.-H. Kang, "A Group-aware Multicast Scheme in 60GHz WLANs," *TIIS*, vol. 5, no. 5, pp. 1028–1048, 2011.

[19] W. Kim, T. Song, and S. Pack, "Rate Adaptation for Directional Multicast in IEEE 802.11 ad Networks," in *2012 IEEE international conference on consumer electronics (ICCE)*, pp. 364–365, IEEE, 2012.

[20] H. Park, S. Park, T. Song, and S. Pack, "An Incremental Multicast Grouping Scheme for mmWave Networks with Directional Antennas," *IEEE Communications Letters*, vol. 17, no. 3, pp. 616–619, 2013.

[21] J.-H. Kwon and E.-J. Kim, "Asymmetric Directional Multicast for Capillary Machine-to-Machine Using mmWave Communications," *Sensors*, vol. 16, no. 4, p. 515, 2016.

[22] O. Galinina, A. Pyattaev, K. Johnsson, S. Andreev, and Y. Koucheryavy, "Analyzing Effects of Directional Deafness on mmWave Channel Access in Unlicensed Bands," in *2017 IEEE Globecom Workshops (GC Wkshps)*, pp. 1–7, IEEE, 2017.

[23] W. Feng, Y. Li, Y. Niu, L. Su, and D. Jin, "Multicast Spatial Reuse Scheduling over Millimeter-Wave Networks," in *2017 13th International Wireless Communications and Mobile Computing Conference (IWCMC)*, pp. 317–322, IEEE, 2017.

[24] S. Naribole and E. Knightly, "Scalable Multicast in Highly-Directional 60-GHz WLANs," *IEEE/ACM Transactions on Networking*, vol. 25, no. 5, pp. 2844–2857, 2017.

[25] M. Makolkina, A. Vikulov, and A. Paramonov, "The Augmented Reality Service Provision in D2D Network," in *International Conference on*

Distributed Computer and Communication Networks, pp. 281–290, Springer, 2017.
[26] Y. Niu, Y. Liu, Y. Li, X. Chen, Z. Zhong, and Z. Han, "Device-to-Device Communications Enabled Energy Efficient Multicast Scheduling in mmWave Small Cells," *IEEE Transactions on Communications*, vol. 66, no. 3, pp. 1093–1109, 2017.
[27] Y. Niu, L. Yu, Y. Li, Z. Zhong, and B. Ai, "Device-to-Device Communications Enabled Multicast Scheduling for mmWave Small Cells Using Multi-level Codebooks," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 3, pp. 2724–2738, 2018.
[28] G. Hong, M. Mousavi, L. Wang, A. Klein, M. Hollick, *et al.*, "Joint Relaying and Spatial Sharing Multicast Scheduling for mmWave Networks," *arXiv preprint arXiv:1907.13085*, 2019.
[29] A. Biason and M. Zorzi, "Multicast Transmissions in Directional mmWave Communications," in *European Wireless 2017; 23th European Wireless Conference*, pp. 1–7, VDE, 2017.
[30] M. Fallgren, T. Abbas, S. Allio, J. Alonso-Zarate, G. Fodor, L. Gallo, A. Kousaridas, Y. Li, Z. Li, Z. Li, *et al.*, "Multicast and Broadcast Enablers for High-Performing Cellular V2X Systems," *IEEE Transactions on Broadcasting*, vol. 65, no. 2, pp. 454–463, 2019.
[31] R. Kovalchukov, D. Moltchanov, A. Pyattaev, and A. Ometov, "Evaluating Multi-connectivity in 5G NR Systems with Mixture of Unicast and Multicast Traffic," in *International Conference on Wired/Wireless Internet Communication*, pp. 118–128, Springer, 2019.
[32] A. Samuylov, D. Moltchanov, R. Kovalchukov, R. Pirmagomedov, Y. Gaidamaka, S. Andreev, Y. Koucheryavy, and K. Samouylov, "Characterizing Resource Allocation Trade-offs in 5G NR Serving Multicast and Unicast Traffic," *IEEE Trans. on Wireless Communications*, 2020.
[33] E. Garro, M. Fuentes, J. Carcel, H. Chen, D. Mi, F. Tesema, J. Gimenez, and D. Gomez-Barquero, "5G Mixed Mode: NR Multicast-Broadcast Services," *IEEE Trans. on Broadcasting*, vol. 66, no. 2 Part II, 2020.
[34] L. Yang, J. Chen, Q. Ni, J. Shi, and X. Xue, "NOMA-enabled Cooperative Unicast–Multicast: Design and Outage Aalysis," *IEEE Transactions on Wireless Communications*, vol. 16, no. 12, pp. 7870–7889, 2017.
[35] K. Sundaresan, K. Ramachandran, and S. Rangarajan, "Optimal Beam Scheduling for Multicasting in Wireless Networks," in *Proceedings of the 15th annual international conference on Mobile computing and networking*, pp. 205–216, 2009.
[36] S. Pizzi, M. Condoluci, G. Araniti, A. Molinaro, A. Iera, and G.-M. Muntean, "A unified approach for efficient delivery of unicast and multicast wireless video services," *IEEE Transactions on Wireless Communications*, vol. 15, no. 12, pp. 8063–8076, 2016.
[37] O. Chukhno, N. Chukhno, O. Galinina, Y. Gaidamaka, S. Andreev, and K. Samouylov, "Analysis of 3D Deafness Effects in Highly Directional mmWave Communications," in *2019 IEEE Global Communications Conference (GLOBECOM)*, pp. 1–6, IEEE, 2019.
[38] J. Meredith, "Study on Channel Model for Frequency Spectrum Above 6 GHZ," *3GPP TR 38.900, Jun, Tech. Rep.*, 2016.
[39] "Study on Channel Model for Frequencies from 0.5 to 100 GHz (Release 14)," tech. rep., 3GPP TR 38.901 V14.1.1, July 2017.
[40] A. W. Doerry, "Catalog of Window Taper Functions for Sidelobe Control," *Sandia National Laboratories Report SAND2017-4042, Unlimited Release*, vol. 10, p. 1365510, 2017.
[41] S. Sen, J. Xiong, R. Ghosh, and R. R. Choudhury, "Link Layer Multicasting with Smart Antennas: No Client Left Behind," in *2008 IEEE International Conf. on Network Protocols*, pp. 53–62, IEEE, 2008.
[42] J. Fan, Q. Yin, G. Y. Li, B. Peng, and X. Zhu, "MCS Selection for Throughput Improvement in Downlink LTE Systems," in *2011 Proceedings of 20th international conference on computer communications and networks (ICCCN)*, pp. 1–5, IEEE, 2011.

Nadezhda Chukhno is an Early Stage Researcher at A-WEAR and Doctoral Researcher at Mediterranea University of Reggio Calabria, Italy and Jaume I University, Spain. She graduated from RUDN University, Russia, and received her B.Sc. in Business Informatics (2017) and M.Sc. in Fundamental Informatics and Information technologies (2019). Her current research activity mainly focuses on wireless communications, 5G networks, D2D, and wearable technologies.

Olga Chukhno is an Early Stage Researcher within H2020 MCSA ITN/EJD A-WEAR project and a PhD student at Mediterranea University of Reggio Calabria, Italy and Tampere University, Finland. She received M.Sc. (2019) in Fundamental Informatics and Information Technologies and B.Sc. (2017) in Business Informatics from RUDN University, Russia. Her current research interests include wireless communications, social networking, edge computing, and wearable applications.

Sara Pizzi is an assistant professor in telecommunications at University Mediterranea of Reggio Calabria, Italy. From the same university she received the 1st (2002) and 2nd (2005) level Laurea Degree, both cum laude, in Telecommunication Engineering and the Ph.D. degree (2009) in Computer, Biomedical and Telecommunication Engineering. In 2005, she received a Master's degree in IT from CEFRIEL/Politecnico di Milano. She was a visiting PhD student at the Department of Computer Science of Alma Mater Studiorum-University of Bologna in 2008. Her current research interests focus on radio resource management for multicast service delivery, Device-to-Device and Machine Type Communications over 5G networks, integration of Non-Terrestrial Networks in the Internet of Things.

Antonella Molinaro graduated in Computer Engineering (1991) at the University of Calabria, received a Master degree in Information Technology from CEFRIEL/Polytechnic of Milano (1992), and a Ph.D. degree in Multimedia Technologies and Communications Systems (1996). She is currently an associate professor of telecommunications at the University Mediterranea of Reggio Calabria, Italy. Her research activity mainly focuses on wireless and mobile networking, vehicular networks, and future Internet.

Antonio Iera graduated in computer engineering from the University of Calabria in 1991, and received a Master's degree in IT from CEFRIEL/Politecnico di Milano in 1992 and a Ph.D. degree from the University of Calabria in 1996. From 1997 to 2019 he has been with the University Mediterranea, Italy, and currently holds the position of full professor of Telecommunications at the University of Calabria, Italy. His research interests include next generation mobile and wireless systems, and the Internet of Things.

Giuseppe Araniti (Senior Member, IEEE) received the Laurea degree and the Ph.D. degree in electronic engineering from the University Mediterranea of Reggio Calabria, Italy, in 2000 and 2004, respectively. He is currently an Assistant Professor of telecommunications with the University Mediterranea of Reggio Calabria. His major area of research is on 5G/6G networks and it includes personal communications, enhanced wireless and satellite systems, traffic and radio resource management, multicast and broadcast services, device-to-device (D2D), and machine-type communications (M2M/MTC).