Evaluating multi-connectivity in 5G NR systems with mixture of unicast and multicast traffic

Evaluating Multi-Connectivity in 5G NR Systems with Mixture of Unicast and Multicast Traffic

Roman Kovalchukov, Dmitri Moltchanov,
Alexander Pyattaev, and Aleksandr Ometov

Tampere University, Tampere, Finland
aleksandr.ometov@tuni.fi

Abstract. The future 5G New Radio (NR) systems are expected to support both multicast and unicast traffic. However, these traffic types require principally different NR system parameters. Particularly, the area covered by a single antenna configuration needs to be maximized when serving multicast traffic to efficiently use system resources. This prevents the system from using the maximum allowed number of antenna elements decreasing the inter-site distance between NR base stations. In this paper, we formulate a model of NR system with multi-connectivity capability serving a mixture of unicast and multicast traffic types. We show that multi-connectivity enables a trade-off between new and ongoing session drop probabilities for both unicast and multicast traffic types. Furthermore, supporting just two simultaneously active links allows to exploit most of the gains and the value of adding additional links is negligible. We also show that the service specifics implicitly prioritize multicast sessions over unicast ones. If one needs to achieve a balance between unicast and multicast session drop probabilities, explicit prioritization mechanism is needed at NR base stations.

Keywords: New Radio · 5G cellular systems · Multicasting · Multi-connectivity · Session drop probabilities · Resource utilization.

1 Introduction

The Third Generation Partnership Project (3GPP) is currently in the process of specifying a new 5^{th} generation (5G) radio interface widely known as 5G NR [1]. The first phase of the system is expected to be the "ready" in 2020, and the standartization is pacing its' way fast [2]. Industry and academia show great interest in 5G NR technology providing the initial performance evaluation not only by simulations but also with testbeds in the field and laboratory environments [3–5]. 5G NR is generally aiming at bringing high rates and reducing latencies to the air interface at the same time [6], thus, enabling an entirely new range of bandwidth-hungry real-time applications such as HD streaming, augmented- and virtual reality [7], etc.

The use of millimeter wave (mmWave) band for NR operation brings specific challenges to system designers. These include unique propagation properties, dynamic human-body blockage environment with the mobility of users and blockers, beam steering with large antenna arrays, the mobility of users [8–10], etc. Accordingly, performance analysis of unicast traffic support in NR systems has received considerable attention over the last few years. The authors in [11] quantify the effect of antenna pattern directivity and human-body blockage on mean interference and signal-to-interference (SIR) ratio in NR systems. Corresponding Laplace transforms of these metrics have been reported in [12]. Three-dimensional deployments have been considered in [13]. The engineering studies focused on assessing the effects of mitigating blockage techniques, such as multi-connectivity and guard capacity, recently started to appear [14, 15].

Most of the studies performed so far concentrated solely on the unicast type of traffic. However, similarly to other radio access technologies NR should also support multicast sessions [16]. Unfortunately, only little is known regarding the support of multicast service in 5G NR systems. In [17], the authors propose a dynamic grouping scheme for mobile users having the same multicast session based on their proximity. The associated algorithm is based on consecutive testing of different antenna half-power beamwidths (HPBW) maximizing the sum rate of the system. A similar approach is proposed in [18]. In addition to the use of dynamic HPBW, the optimization framework developed in [19] accounts for non-equal power-sharing between beams.

The only study, where the mixture of unicast and multicast traffic is addressed is due to Samuilov et al. [20]. Among other conclusions, the authors demonstrated that there exist a specific parameters for base station (BS) and user equipment (UE) inducing BR BS coverage R_M with session drop probabilities (p_M, p_U) for a given arrival rates of multicast and unicast sessions (λ_M, λ_U). Thus, there is a principal trade-off between system characteristics required for unicast and multicast traffic affecting the inter-site distance (ISD) between NR BS: while unicast traffic requires small HPBW to extend ISD and decrease deployment cost, multicast traffic needs higher HPBW to form larger multicast groups and thus decrease the load imposed on NR BS.

One of the consequences of increased ISD in NR deployment is potential session drops caused by human-body blockage phenomenon or insufficient resources available at the NR BS to support a session changing its state from non-blocked to blocked [15, 21]. To alleviate these effects, 3GPP has recently proposed so-called multi-connectivity operation (also known as macro-diversity), where UE is allowed to simultaneously support connections to multiple NR BSs and switch between them in case of outage events [22].

In this paper, we investigate the effect of recently standardized 3GPP multi-connectivity operation on performance metrics of a mixture of unicast and multicast traffic in the cellular deployment of NR systems. Based on the developed system level modeling framework we analyze user- and system-centric performance metrics provided to multicast and unicast sessions in the presence of multi-connectivity capabilities.

The paper is organized as follows. In Section 2, the system model is described. We introduce out system level modeling framework in Section 3. Numerical results are presented in Section 4. The last section concludes the paper.

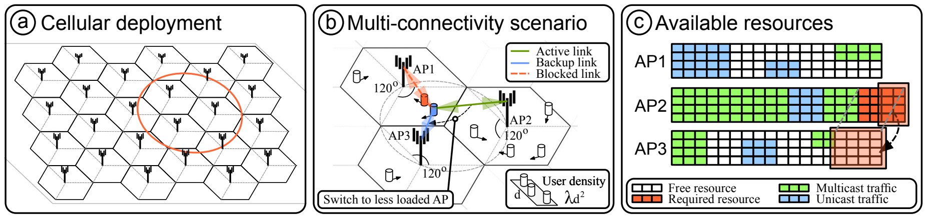


Fig. 1. An illustration of the considered NR cellular deployment.

2 System model

The system model is introduced in this section by specifying the corresponding components including deployment, antenna, propagation, blockage, traffic, connectivity, and service models. Finally, we specify the metrics of interest.

Deployment, Traffic, and Blockage Models We consider a standard NR cellular deployment illustrated in Fig. 1(a). The height of all NR BSs is assumed to be h_A. The ISD is assumed to be D.

In the considered deployment, moving pedestrians, represented by cylinders with constant height and base radius, h_B and r_B, respectively, may block the line-of-sight (LoS) path between BS and UE, see Fig. 1(b). Pedestrians, also acting as blockers, are assumed to move according to random direction model (RDM) with constant speed v_B, and run the length of τ_B s [23]. Using the property of RDM model, at each instant of time blockers organize a Poisson point process (PPP) in $\Re^2$. The density of blockers is assumed to be λ_B blockers/m^2. According to [24], the blockage of human-body in NR bands result in additional degradation of $15-40\,$dB. In this study, we assume 20 dB losses.

The session arrival rate is assumed to be λ session/m^2. With probability p_M, the session is assumed to be of a multicast type. A session is classified as multicast with complementary probability $1-p_M$. Unicast and multicast sessions are characterized by constant rate requirements of T_U and T_M Mbps, respectively. The corresponding session durations are exponentially distributed with parameters μ_U and μ_M.

Propagation and Antenna Models In this work, we use 3GPP urban micro (UMi) street canyon propagation model [25] providing the following path loss at the three-dimensional distance x between NR BS and UE

$$L(x) = \begin{cases} 52.4 + 21.0\log x + 20\log f_c, \text{ LoS blocked}, \\ 32.4 + 21.0\log x + 20\log f_c, \text{ LoS non-blocked}, \end{cases} \quad (1)$$

where f_c is the carrier frequency measured in GHz.

The selected model specifies three zones around BS. Up to the distance R_B, UE is never in outage conditions even when LoS path is blocked. In the interval

(R_B, R_O), UE is in outage only when the LoS is blocked. Finally, R_O defines the distance when UE is in outage even when LoS is not blocked.

Linear antenna array is assumed to be utilized at BSs an UEs. The gain at transmit and receive sides is computed as [26]

$$G = \frac{1}{\theta_{3db}^{+} - \theta_{3db}^{-}} \int_{\theta_{3db}^{-}}^{\theta_{3db}^{+}} \frac{\sin(K\pi\cos(\theta)/2)}{\sin(\pi\cos(\theta)/2)} d\theta, \tag{2}$$

where K is a number of elements in a linear array.

Connectivity and Service Models We assume that the UE has the 3GPP multi-connectivity functionality [22], where each node maintains the links to M neighboring BSs. M is referred to as "degree of multi-connectivity". According to the multi-connectivity operation, the UE can switch to one of the available BSs in case current link becomes unavailable.

Upon arrival, a unicast session is assumed to establish M active connections with M nearest NR BSs based on time-averaged SNR metric. A connection with the nearest BS is first tried. Depending on the distance between BS and UE the required rate T_U is translated into the bandwidth requirements B_U using the set of NR modulation and coding schemes (MCS, [27]). If there is an insufficient amount of resources at this BS, BS with the next nearest distance is used. If none of those BSs have a sufficient amount of resources, a unicast session is dropped. Upon blockage event, SNR drops by 20 dB and session changes its resource requirements. If there are no resources available to support these new bandwidth requirements, BS with the nearest distance is tried next. If there are no BSs out of M available that may support these new bandwidth requirements, a session is dropped during service.

The connectivity process of multicast sessions is similar to described above for unicast sessions. The principal difference is in the resource allocation at BSs. In particular, upon arrival of the multicast session, the nearest BS tests whether there is an ongoing multicast service and, if yes, what kind of MCS it uses. If there is an ongoing multicast service, a new session is accepted to the system whenever the MCS order is higher than the one used for ongoing multicast sessions. If the MCS order is lower, the BS tests whether it can support multicast service at this lower MCS. If the conclusion is positive, session is accepted to the system and the current MCS of the multicast service is lowered. Otherwise, other $M-1$ is descending order of distances are tested using the same procedure. Finally, if there is no ongoing multicast service, the session acceptance logic is similar to unicast sessions. At the blockage time instant, the procedure described above is tested first for the currently active BS and then for the remaining $M-1$ BSs. If none of those may support an ongoing session in blockage state, a multicast session is dropped.

Metrics of Interest In the considered system model, we are interested in new and ongoing multicast, $p_{N,M}$ and $p_{O,M}$, and unicast, $p_{N,U}$ and $p_{O,U}$, session drop probabilities. The new session drop probability is defined as the probability that at the moment of new session arrival there are no resources to accept the session for service. Ongoing session drop probability is defined as the probability that a session accepted for service is then dropped at the moment of blockage due to insufficient resources available. As a system-centric performance metric, we concentrate on system resource utilization averaged over all NR BSs.

3 Performance Evaluation Framework

To characterize the system performance, we have developed a system level simulation framework (SLS) is based on discrete-event simulation (DES) with multi-thread optimization in Java.

To achieve reliable results, we have executed the simulations for 10^6 seconds of the simulated system time. Since all stochastic processes in our system are stationary, the system of interest eventually converges to the steady-state. The beginning of the steady-state is detected using the exponentially-weighted moving average (EWMA) technique with the weighting parameter 0.1 [28]. The average duration of the transient period has been found to be 134 seconds.

During the simulation run, the data is collected during the steady-state period only. On top of sampling the system-state every 10 seconds, the batch means strategy is used to exclude residual correlation in the statistical data. According to it, batches of 10^3 observations were collected first. The means of these batches are considered as independent identically distributed (iid) observations. The resulting iid samples were processed using conventional statistical methods. Due to the large sample sizes, in what follows, only point estimates are demonstrated. The reason is that the interval estimates do not differ by more than ± 0.01 from the point estimates under the level of significance $\alpha = 0.1$.

4 Selected Numerical Results

In this section, we numerically assess the performance gains provided by multi-connectivity operation at BSs serving mixtures of unicast and multicast traffic. The default input system parameters used in this section are provided in Table 1.

First, we consider the new session drop probability for multicast and unicast sessions with different degrees of multi-connectivity as a function of the fraction of multicast sessions, p_M, illustrated in Fig. 2. As one may observe, the presence of multi-connectivity has a profound negative effect on the new session drop probabilities for both types of traffic. Analyzing the presented data further, we may observe that the new session drop probabilities of both type of traffic decrease as the fraction of multicast sessions increases. This behavior is logical for the unicast sessions as their fraction decreases as p_M increases. In cases, of the multicast session, this effect is explained by the nature of the service process. Indeed, as multicast sessions do not always require a new set of resources upon their arrival. If there is at least one multicast session in the system, they may change the amount of resources required for service. Thus, when p_M is not

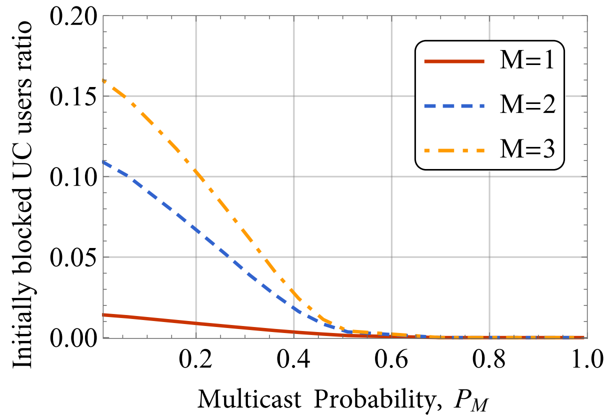
Table 1. Main parameters for numerical assessment.

System parameter	**Value**
User density, λ	0.5 users/m^2
User velocity, v_B	4 m/s
BS height, h_A	4 m
UE height, h_U	1.5 m
Blocker height, h_B	1.7 m
Blocker radius, r_B	0.3 m
Multi-connectivity degree, M	$(1, 2, 3)$
Number of NR BSs, N	3
Transmit power, P_T	0.2 W
Planar antenna elements at BS, K_A	32
Planar antenna elements at UE, K_U	4
Carrier frequency, f_c	28 GHz
Bandwidth at each NR BS, B	1 GHz
LoS blockage loss, L_B	20 dB
Mean session time, $1/\mu$	20 s
Session initiation probability, p_A	3.14×10^{-4}
Session rate, R	100 Mbps

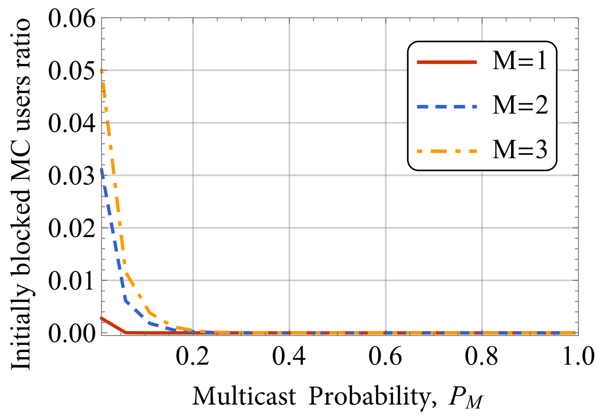
negligible, each NR BSs almost always have an active multicast session in the system drastically decreasing the multicast session drop probability.

Consider now the effect of multi-connectivity on ongoing session drop probability illustrated in Fig. 3 for a range of values of a fraction of multicast sessions, p_M. As one may observe, there is a noticeable positive effect of multi-connectivity on both considered types of traffic. Quantitatively, it is essential to note that the gains of the system with $M = 3$ compared to $M = 2$ is much milder than that of $M = 2$ compared to no multi-connectivity at all, $M = 1$. This is essential observation as maintaining simultaneously active links is expensive from the control overhead point of view [22].

Analyzing the data corresponding to multicast service, illustrated in Fig. 2(b) and Fig. 3(b), one may notice that much sharper decays characterize the new and ongoing session drop probabilities compared to unicast traffic. As the service process of this traffic is drastically different the interplay between the unicast and multicast type of traffic in NR systems is of particular interest. Notably, for non-negligible values of p_M this type of traffic has implicit priority over unicast sessions. This conclusion is supported by the absolute values of new and ongoing session drop probabilities in Fig. 2(b) and Fig. 3(b). Thus, NR system may need some explicit mechanism to prioritize unicast traffic.



(a) Unicast



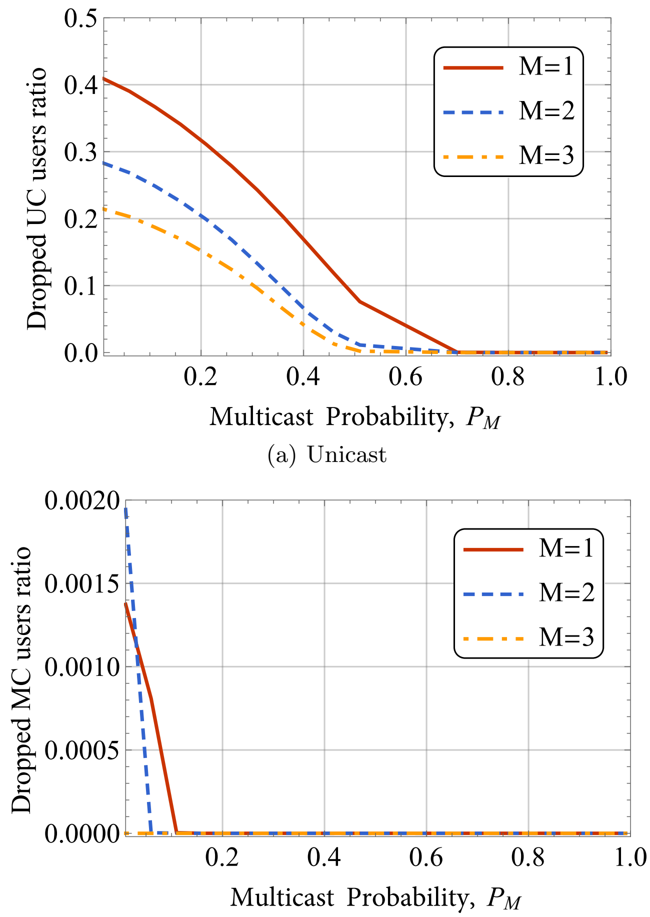
(b) Multicast

Fig. 2. New session drop probabilities for unicast and multicast traffic.

Finally, we assess the system-centric metric of interest – resource utilization averaged across NR BSs, illustrated in Fig. 4. As one may observe, the system with multi-connectivity operation enabled, $M = 2$ or $M = 3$, is characterized by much better performance compared to the system with $M = 1$ in overloaded regime corresponding to small values of p_M. When the fraction of multicast sessions increases, the systems switch to the underloaded regime, where the performance of systems with different degrees of multi-connectivity coincide.

5 Conclusion

Motivated by the need to support multicast traffic in prospective 5G NR systems, we formulated a model of cellular NR deployment serving a mixture of unicast



(a) Unicast

(b) Multicast

Fig. 3. Ongoing session drop probabilities for unicast and multicast traffic.

and multicast traffic in this paper. Owing to the complexity of the systems, the developed model is based on system level simulations with an analytical model of the LoS blockage process. The proposed framework can be used to assess user- and system-centric performance metrics of interest.

We have demonstrated that the multi-connectivity operation negatively affects new session drop probabilities for both multicast and unicast traffic. However, it allows to significantly improve the service performance of sessions already accepted to the system by drastically reducing the ongoing session drop probability for both types of traffic. Varying the degree of multi-connectivity one may achieve the desired trade-off between new and ongoing session drop probability. Quantitatively, the significant impact on both user- and system-centric performance metrics is produced by supporting just two simultaneously active links,

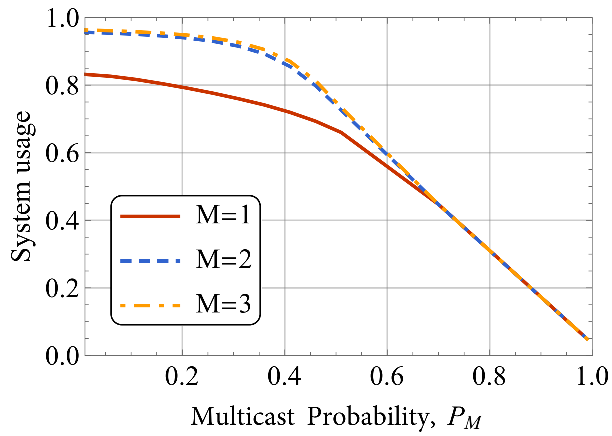


Fig. 4. Resource utilization averaged across NR BSs.

so-called dual connectivity operation. The gains of adding additional links are much milder.

Finally, we would like to note that the specifics of the service process of multicast traffic profoundly affects unicast traffic performance. In particular, if the fraction of multicast session is non-negligible, there is resource capturing effect, when multicast sessions exclusively occupy a fraction of resources. Thus, if one needs to achieve a balance between unicast and multicast new and ongoing session drop probabilities, explicit prioritization mechanism is needed.

References

1. S. Parkvall, E. Dahlman, A. Furuskar, and M. Frenne, "NR: The New 5G Radio Access Technology," *IEEE Communications Standards Magazine*, vol. 1, no. 4, pp. 24–30, 2017.
2. 3GPP, "Multiplexing and Channel Coding, 3rd Generation Partnership Project (3GPP)," TS 38.212, v15.4.0, Release 15., December 2018.
3. Qualcomm, "mmWave 5G NR prototype demo video." [Online] `https://www.qualcomm.com/videos/mmwave-5g-nr-prototype-demo-video`, May 2018.
4. B. Halvarsson, K. Larsson, M. Thurfjell, K. Hiltunen, K. Tran, P. Machado, D. Juchnevicius, and H. Asplund, "5G NR Coverage, Performance and Beam Management Demonstrated in an Outdoor Urban Environment at 28 GHz," in *Proc. of IEEE 5G World Forum (5GWF)*, pp. 416–421, 2018.
5. M. Morant, A. Trinidad, E. Tangdiongga, T. Koonen, and R. Llorente, "Experimental Demonstration of mm-Wave 5G NR Photonic Beamforming Based on ORRs and Multicore Fiber," *IEEE Transactions on Microwave Theory and Techniques*, 2019.
6. A. Ghosh, "5G New Radio (NR): Physical Layer Overview and Performance," in *Proc. of IEEE Communication Theory Workshop*, pp. 1–38, 2018.

7. A. Prasad, A. Benjebbour, O. Bulakci, K. I. Pedersen, N. K. Pratas, and M. Mezzavilla, "Agile Radio Resource Management Techniques for 5G New Radio," *IEEE Communications Magazine*, vol. 55, no. 6, pp. 62–63, 2017.
8. D. Moltchanov and A. Ometov, "On the Fraction of LoS Blockage Time in mmWave Systems with Mobile Users and Blockers," in *Proc. of International Conference on Wired/Wireless Internet Communication*, pp. 183–192, Springer, 2018.
9. K. Zeman, M. Stusek, P. Masek, and J. Hosek, "Improved NLOS Propagation Models for Wireless Communication in mmWave bands," in *Proc. of 8th International Conference on Localization and GNSS (ICL-GNSS)*, pp. 1–6, IEEE, 2018.
10. D. Moltchanov, A. Ometov, S. Andreev, and Y. Koucheryavy, "Upper Bound on Capacity of 5G mmWave Cellular with Multi-connectivity Capabilities," *Electronics Letters*, vol. 54, no. 11, pp. 724–726, 2018.
11. V. Petrov, D. Moltchanov, and Y. Koucheryavy, "On the Efficiency of Spatial Channel Reuse in Ultra-Dense THz Networks," in *Proc. of IEEE Globecom*, Dec. 2015.
12. S. Singh, M. N. Kulkarni, A. Ghosh, and J. G. Andrews, "Tractable Model for Rate in Self-backhauled Millimeter Wave Cellular Networks," *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 10, pp. 2196–2211, 2015.
13. R. Kovalchukov, D. Moltchanov, A. Samuylov, A. Ometov, S. Andreev, Y. Koucheryavy, and K. Samouylov, "Evaluating SIR in 3D Millimeter-Wave Deployments: Direct Modeling and Feasible Approximations," *IEEE Transactions on Wireless Communications*, vol. 18, no. 2, pp. 879–896, 2019.
14. V. Petrov, M. A. Lema, M. Gapeyenko, K. Antonakoglou, D. Moltchanov, F. Sardis, A. Samuylov, S. Andreev, Y. Koucheryavy, and M. Dohler, "Achieving End-to-End Reliability of Mission-Critical Traffic in Softwarized 5G Networks," *IEEE Journal on Selected Areas in Communications*, vol. 36, no. 3, pp. 485–501, 2018.
15. R. Kovalchukov, D. Moltchanov, V. Begishev, A. Samuylov, S. Andreev, Y. Koucheryavy, and K. Samouylov, "Improved Session Continuity in 5G NR with Joint Use of Multi-Connectivity and Guard Bandwidth," in *Proc. of Global Communications Conference (GLOBECOM)*, pp. 1–7, IEEE, 2018.
16. 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.
17. 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.
18. W. Feng, Y. Li, Y. Niu, L. Su, and D. Jin, "Multicast Spatial Reuse Scheduling over Millimeter-Wave Networks," in *Proc of 13th International Wireless Communications and Mobile Computing Conference (IWCMC)*, pp. 317–322, IEEE, 2017.
19. K. Sundaresan, K. Ramachandran, and S. Rangarajan, "Optimal Beam Scheduling for Multicasting in Wireless Networks," in *Proc. of the 15th Annual International Conference on Mobile Computing and Networking*, pp. 205–216, ACM, 2009.
20. A. Samuvlov, D. Moltchanov, A. Krupko, R. Kovalchukov, F. Moskaleva, and Y. Gaidamaka, "Performance Analysis of Mixture of Unicast and Multicast Sessions in 5G NR Systems," in *Proc. of 10th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT)*, pp. 1–7, IEEE, 2018.
21. D. Moltchanov, A. Samuylov, V. Petrov, M. Gapeyenko, N. Himayat, S. Andreev, and Y. Koucheryavy, "Improving Session Continuity with Bandwidth Reservation

in mmWave Communications," *IEEE Wireless Communications Letters*, vol. 8, no. 1, pp. 105–108, 2019.
22. 3GPP, "NR; Multi-connectivity; Overall description (Release 15)," 3GPP TS 37.340 V15.2.0, June 2018.
23. P. Nain, D. Towsley, B. Liu, and Z. Liu, "Properties of Random Direction Models," in *Proc. of 24th Annual Joint Conference of the IEEE Computer and Communications Societies*, vol. 3, pp. 1897–1907, IEEE, 2005.
24. K. Haneda *et al.*, "5G 3GPP-like Channel Models for Outdoor Urban Microcellular and Macrocellular Environments," in *Proc. of IEEE Vehicular Technology Conference (VTC 2016-Spring)*, May 2016.
25. 3GPP, "Study on channel model for frequencies from 0.5 to 100 GHz (Release 15)," 3GPP TR 38.901 V15.0.0, June 2018.
26. C. A. Balanis, *Antenna Theory: Analysis and Design.* John wiley & sons, 2016.
27. 3GPP, "NR; Physical channels and modulation (Release 15)," 3GPP TR 38.211, Dec 2017.
28. H. Perros, "Computer Simulation Techniques," *The definitive introduction. North Carolina State University*, 2009.