# Performance Modelling for Social VR Conference Applications in Beyond-5G Radio Access Networks

João Morais
*Radio Systems Group*
*Instituto Superior Técnico*
Lisbon, Portugal
joao.de.morais@tecnico.pt

Sjors Braam
*Department of Networks*
*TNO*
The Hague, The Netherlands
sjors.braam@tno.nl

Remco Litjens
*Department of Networks*
*TNO*
The Hague, The Netherlands
*Faculty of EEMCS*
*TU Delft*
Delft, The Netherlands
remco.litjens@tno.nl

Hans van den Berg
*Department of Networks*
*TNO*
The Hague, The Netherlands
*Faculty of EEMCS*
*University of Twente*
Enschede, The Netherlands
j.l.vandenberg@tno.nl

***Abstract*— One of the most challenging applications targeted by evolving (beyond-)5G technology is virtual reality (VR). Particularly, 'Social VR' applications provide a fully immersive experience and sense of togetherness to users residing at different locations. To support such applications the network must deal with enormous traffic demands, while keeping end-to-end latencies low. Moreover, the radio access network must deal with the volatility and vulnerability of mm-wave radio channels, where even small movements of the users may cause line-of-sight blockage, causing severe throughput reductions and hence Quality of Experience (QoE) degradation or even lead to loss of connectivity. In this paper we present an integral modelling approach for feasibility assessment and performance optimisation of the radio access network for Social VR applications in indoor office scenarios. Such modelling enables us to determine the performance impact of e.g. 'natural' human behaviour, the positions and configurations of the antennas and different resource management strategies. Insights into these issues are a prerequisite for setting up guidelines for network deployment and configuration as well as for the development of (potentially AI/ML-based) methods for dynamic resource management and tuning of radio access parameters to best support Social VR applications.**

***Keywords—Social XR, VR, 5G, modelling***

## I. INTRODUCTION

It is widely recognized that emerging 5G network technology [1] will boost the development of new, highly innovative applications in virtually all domains. One of the most challenging application classes targeted by (beyond-)5G networks is augmented/virtual reality (AR/VR), in particular scenarios with multiple distant users who are able to interact fluently with each other through all human senses [2][3]. Transporting one's social and functional self to any place on earth is an exciting idea that will save travelling time and costs (and reduce carbon footprint) by e.g. enabling virtual meetings, exploitation of remote expertise or skills in smart industry or other contexts in the form of collaborative working, remote inspection or maintenance, and remote education and training.

To support such so-called 'Social VR' applications with compelling visual, haptic, audio and possibly even olfactory experiences for remote users, the networking infrastructure must be able to handle extremely high-bandwidth streams while keeping the end-to-end latency low [3][4][5]. Furthermore, the infrastructure should provide powerful (in-network) processing capabilities, reducing the need for heavy computational capabilities at the client devices, thus increasing the clients' flexibility and mobility. Actually, natural Social VR experiences can only be realised when all components in the end-to-end capture, transmission and rendering pipeline optimally work together [5][7][8]. Optimisation of the Quality of Experience (QoE) requires dynamic orchestration of these components, allocation of resources in the networking and compute infrastructure, and an optimised exploitation of the connect-compute trade-off, noting that the end-to-end latency budget can be flexibly distributed over computational and transport tasks. In particular, dynamic orchestration and resource management complemented with e.g. multi-connectivity should also be able to cope with the intermittent nature of high-frequency (mm-wave) radio channels used in 5G and beyond to satisfy the huge capacity and throughput requirements [5][9][10].

Other (beyond-)5G Radio Access Network (RAN) features and capabilities to support extremely demanding wireless applications like Social VR, are the use of massive MIMO-based beamforming to overcome the associated attenuation challenges, the flexible OFDM numerology, the use of scheduling mini-slots and self-contained subframes, and the use of channel-adaptive latency-based packet schedulers to efficiently achieve high data rates while satisfying latency requirements.

This paper focuses on radio access challenges for supporting Social VR applications. In particular, we provide an integral modelling framework covering all relevant aspects needed for an adequate feasibility assessment and performance optimisation of 'Social VR' applications in indoor office scenarios. Studies presented in literature mostly focus on individual aspects, particularly blockage of mm-wave channels and possible ways to mitigate such blockage effects [10][11]. Comprehensive studies of VR performance in the RAN accounting for multiple relevant aspects (and their dependencies) simultaneously are lacking, while there is a clear need for such an integral view.

The modelling aspects covered in our study include the specifics of (*i*) the use case scenario: the office layout, application's traffic characteristics and service requirements; (*ii*) the network deployment, incl. the deployed antennas; (*iii*) the propagation environment; and (*iv*) key 5G traffic handling/resource management mechanisms. The integration and implementation of the developed models in a simulation tool enables e.g. the optimisation of antenna deployments and selection of (sub- or sup-6 GHz) frequency bands, an

assessment of the use of MU/D-MIMO-based (multi-user/distributed MIMO) beamforming, latency-based packet scheduling and other latency-oriented scheduling features, as well as an investigation into the feasibility and performance impact of use case aspects such as the number of physical/virtual meeting participants and their behaviour in term of typical head/body movements. Such insights are particularly useful, a prerequisite even, for setting up guidelines for local network deployment as well as for the development of (potentially AI/ML-based) methods for dynamic resource management and tuning of radio access parameters to best support Social VR applications in indoor office scenarios. The envisioned ultimate goal is then to use these results for the design of a fully automated 'Social VR network slice', i.e. a self-configuring and autonomously managed 5G (radio) network slice to support this type of Social VR applications.

The remainder of the paper is organized as follows. First, in Section II an overview of related work is given. Next, in Sections III to VI, we present our integral modelling framework for the performance assessment and optimisation of Social VR scenarios. In particular, Section III describes all relevant aspects of a Social VR office scenario, including location (meeting room) specifics, the 'mix' of involved physical and virtual meeting participants, participant behaviour and VR application characteristics and Quality of Service (QoS) requirements. Next, Section IV discusses different options for network deployment (number of antennas, their positions and characteristics). Section V then describes the modelling of the propagation aspects that need to be taken into account. Section VI discusses the modelling of various traffic handling mechanisms (e.g. the definition of a Grid of Beams, beam pair establishment and QoS-oriented packet scheduling) which are crucial for efficiently supporting the highly demanding VR application. Finally, Section VII provides concluding remarks and an outlook for further research exploiting the proposed modelling framework.

## II. Related work

There is a large body of research addressing virtual reality systems in a wireless setting. In [12], a framework is presented that analyses the performance of VR services over wireless networks. The framework captures the tracking accuracy, transmission delay, and processing delay, but most radio characteristics such as frequency-selective fading, antenna configurations and blockage effects are not considered. In [13], the authors study the impact of blockage by hand, head and body on wireless mm-wave links, and suggest an algorithm to overcome the corresponding challenges. The proposed solution uses a fixed relay to increase robustness against blocking and is assessed in an experimental setup. The attainable gains strongly depend on numerous assumptions and deployment configurations which are not described in any detail. In [10], the challenges and enablers for ultra-reliable and low-latency VR are discussed. The authors state that for a truly immersive VR user experience, three aspects should be designed and orchestrated together: communication, computing and caching. A case study has been worked out where multi-connectivity is used to mitigate the blockage and disturbance caused by several impulse actions. Although the results of this study are quite extensive, a lot of modelling aspects and choices are not covered.

Recently, 3GPP also showed interest in supporting Extended Reality (XR) over 5G networks. An XR 360º conference meeting has been defined in [4]. However, QoS requirements for such use cases are not clearly defined, likely because the QoE assumptions vary so tremendously. 3GPP considers 50-100 Mb/s [4], while other sources mention bit rates in the order of Gb/s [5][10]. Despite lacking concreteness, there is more of an agreement on latency requirements, in the sense that latency should be lower than for current real-time applications.

To the best of our knowledge, there is no study available in literature providing a comprehensive and integral framework of all relevant modelling aspects for the considered type of scenario. This paper aims to provide such an integral modelling overview and fill in any gaps in substantial modelling detail, enabling a variety of sensitivity analyses, derivations of deployment guidelines and the development and assessment of radio network management solutions.

## III. Application scenario

In this first modelling-oriented section we address the different aspects characterising the application scenario. We consider a use case of running a *social VR application* in an indoor office scenario. More specifically, we consider a square meeting room with a large round conference table placed in the centre of the room (see Figure 1). Uniformly spread around the conference table are $N = N_P + N_V$ chairs, seating in this example $N_P = 4$ participants that are physically present in the room, and $N_V = 4$ participants located elsewhere and virtually attending the conference meeting. The physical/virtual participants are assumed to be seated in an alternate fashion (P – V – P – …).

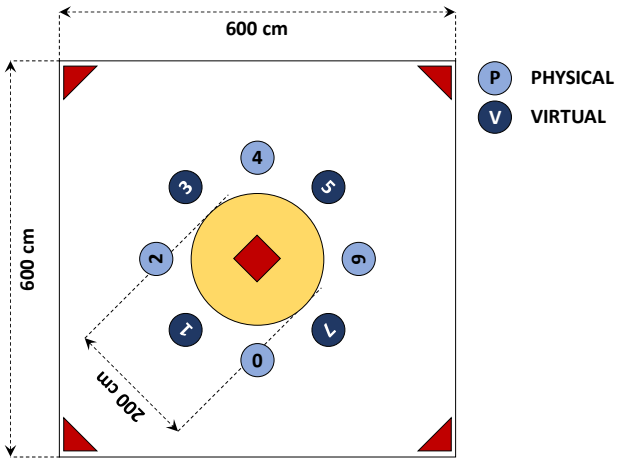


*Figure 1: Visualisation of the 'Social VR' application scenario with eight physically or virtually present meeting participants.*

Since we concentrate on the communication challenges in the considered conference room, for some modelling aspects we specifically focus on the participants that are *physically* present. Each such participant wears a VR headset (Head-Mounted Display; HMD) towards which downlink images are transferred. Furthermore, two on-table video capture devices are installed and pointed to the participant to capture images for uplink transfer (see Figure 2). Although besides video also audio information is captured and transferred, the traffic modelling covers only the video component considering its substantially more demanding nature in terms of offered traffic load and QoS requirements.

Each physically present participant is and remains seated in his chair, but the participant's head is modelled to exhibit some degree of randomised *head motion* in terms of changes in position and orientation [14][15]:

- The change of a head's *position* can be modelled according to a modified random waypoint model in 3D. More specifically, a participant's default (or average) head position has coordinates $(x_0,y_0,z_0)$, with $(x_0,y_0)$ given by its appointed chair and $z_0$ fixed to e.g. 1.2 m. The initial head position is sampled according to a three-dimensional Gaussian distribution with mean vector $(x_0,y_0,z_0)$, standard deviations $\sigma_x$, $\sigma_y$, $\sigma_z$, and all covariances set to 0 $m^2$. From there, the head is modelled to move linearly to its next position, sampled from the same distribution, at a speed of $0.1 \times \mu$ m/s, where $\mu$ is the head motion index in the range [0,10]. Once there, the process repeats itself, ad infinitum.
- The head's *orientation* is given by a randomly sampled 3D vector of angular offsets around a straight line aimed at the current speaker (the speaker itself is assumed to aim his view on the previous speaker). Offsets are independently sampled w.r.t. the 'yes', 'no' and 'may be' axes perpendicularly cutting through the head, assuming uniform angular distributions on the ranges $[-\alpha_Y, \alpha_Y]$, $[-\alpha_N, \alpha_N]$ and $[-\alpha_M, \alpha_M]$, respectively. Whenever the speaker changes, such offsets are sampled and the corresponding change in orientation is modelled to take $1.5 - 0.15 \times \mu$ seconds, with $\mu$ as defined above. Once the targeted orientation is reached, a new offset is sampled and pursued, and so on, until the speaker changes once again.

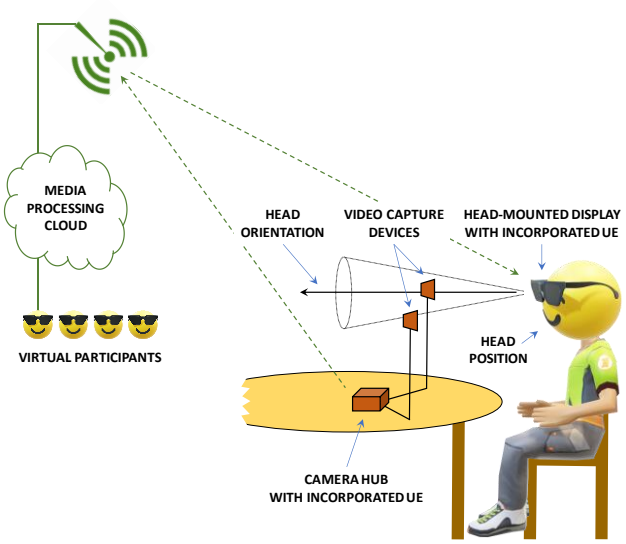


*Figure 2: Illustration of user setting and end-to-end architecture.*

Aside from substantial changes in the head orientation that are due to speaker changes, the randomisations of the head position and its orientation effectively constitute some degree of 'head wobbling' in 3D, with the intensity configured by head motion index $\mu$. In general, the relevance of modelling the head motions lies therein that the antennas used for downlink reception are fixed to the VR headset and hence any realistic motion of a participant's head directly impacts the position and orientation of the receive antennas.

The *meeting dynamics* are modelled by traversing through a pre-generated speaker list, letting each designated participant speak for a deterministic or randomly distributed time of $\tau_S$ seconds, and then switch to the next speaker on the list in some random or predetermined order. For the above-mentioned scenario with $N = N_P + N_V = 4 + 4 = 8$ participants (visualised in Figure 1), the speaker list could e.g. be given by '0 – 5 – 2 – 7 – 4 – 1 – 6 – 3', where we choose to alternate between physically and virtually present users.

The *VR application* is modelled as a persistent real-time video streaming application [16], where each up- or downlink stream is a sequence of GoPs ('Groups of Pictures') and each GoP comprises a single I ('intra-coded') and $M_P$ P ('predicted') video frames. The size of the relatively large fresh I and the differential P frames is denoted $S_I$ and $S_P < S_I$, respectively. A typical frame generation rate of $R_F$ fps (frames per second) is assumed, which is noted to apply at the source. For an example case inspired by [17] and visualised in Figure 3, we have $M_P = 5$, $S_I = 1250$ kB, $S_P = 250$ kB and $R_F$ 30 fps, the application-level bit rate is equal to $R_B = (8 / 1000000) \times R_F \times (S_I + M_P S_P) / (1 + M_P) = 100$ Mb/s, which in [17] is labelled as an 'entry-level VR' quality providing a 2048 × 2048-pixel 2D visual field at 30 frames per second. As said before, given the relative negligibility of the audio component in terms of required bit rates, we limited our modelling to include the transfer of visual content only. In a downlink transfer, this video stream is noted to be a (potentially) processed) aggregation of the distinct uplink video streams captured for the different meeting participants.

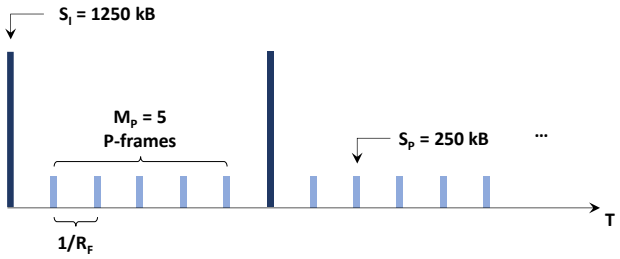


*Figure 3: The VR application traffic generation model is characterised by $R_F$ = 30 fps, $M_P$ = 5, $S_I$ = 1250 kB and $S_P$ = 250 kB, yielding an application-level bit rate of $R_B$ = 100 Mb/s.*

For the considered indoor office scenario, in the uplink the sources generating the video frames are two video capture devices per person, with synchronised aggregation done in a camera hub before the uplink transfer (see Figure 2). The source of the downlink transfers is a so-called multi-point control unit [7] residing in the media processing cloud, which receives, aggregates and potentially processes all uplink video streams before generating a downlink video stream to a given UE. Generated as application-level frames at the source (see Figure 3), the video information arrives in the form of 1500-byte IP packets at the uplink (User Equipment; UE) or downlink (Base Station; BS) transmission buffer. For the downlink streams, the packets are dispersed in time due to the variabilities of packet latencies on the path from source to transmit buffer. For the uplink streams, with the video capture devices directly connected to a camera hub incorporating a 5G UE, no such time dispersion is assumed, and the IP packets are assumed to be available in the UE's transmit buffer immediately upon frame generation. The time dispersion for the downlink streams is modelled as follows. Upon generation of a video frame, the frame is segmented into IP packets at a

'packet extraction bit rate' given by $R_B / (1 - \beta)$, with $R_B$ the application-level bit rate as introduced above and $\beta \in [0,1)$ the configured burstiness parameter. Note that for the extreme case of $\beta = 0$ the IP packets are maximally dispersed (effectively yielding a constant bit rate flow), while for the case of $\beta \rightarrow 1$ each video frame is upon generation instantaneously segmented into IP packets, reflecting the maximum degree of burstiness. Besides this time dispersion, a fixed delay of $\tau_P$ seconds is imposed to each downlink IP packet to model the delay incurred on the way from the source to the transmit buffer in the BS.

The QoS requirement of the VR application is given by a maximum tolerable end-to-end one-way frame-level latency of e.g. $\Delta_{E2E} = 15$ ms [17][1], where any VR frame delivered with a latency exceeding $\Delta_{E2E}$ is considered useless. The RAN segment of the end-to-end path is assigned a packet-level budget of $\Delta_{RAN} < \Delta_{E2E}$ ms, acknowledging that the RAN handles IP packets and is unaware of application-level video frames [17]. Note that while $\Delta_{E2E}$ is determined based on application-level perceived quality considerations, $\Delta_{RAN}$ is essentially a configuration parameter that needs to be optimised and is used by the packet scheduler to give due priority to packets approaching the associated transmission deadline and potentially even drop packets that (are estimated to) exceed the deadline (see also Section VI). The optimal setting of $\Delta_{RAN}$ is not trivial. If we set $\Delta_{RAN}$ too low the scheduler may end up dropping too many packets that could potentially have met their end-to-end latency requirement. Alternatively, if we set $\Delta_{RAN}$ too high the scheduler may not drop any packet and too many packets that end up not meeting their end-to-end deadlines have fruitlessly consumed and hence wasted precious radio resources. In an end-to-end network context, a cross-layer management mechanism may be used to dynamically adapt $\Delta_{RAN}$ in order to optimise the 'connect-compute trade-off', thereby intelligently distributing the end-to-end latency budget over a.o. one or more media processing clouds and the radio access network.

## IV. Network Aspects

The meeting room is adequately equipped to support social VR meetings, including, besides audio/video capturing and audio playback devices, also advanced wireless networking hardware. More specifically, we assume that $K_{BS} \in \{1, 2, 4, 5\}$ massive MIMO antenna arrays are deployed, where the different values of $K_{BS}$ indicate distinct deployment scenarios. For $K_{BS} = 1$, we assume a single antenna mounted at the ceiling centre and pointed downwards; for $K_{BS} = 2$ we assume two downtilted antennas mounted in adjacent (scenario '2a') or in opposite (scenario '2o') ceiling corners; for $K_{BS} = 4$ we assume four downtilted antennas mounted in all four ceiling corners; lastly, for $K_{BS} = 5$ we assume a combination of the antennas at the centre and in the corners of the ceiling. The different deployment scenarios are visualised in Figure 4.

Each BS antenna is assigned a single TDD carrier of a given bandwidth. In this paper we will consider carriers in either the 3.5 GHz or the 26 GHz (mm-wave) band, which constitute distinct assessment scenarios. To enable a fair comparison the band-specific assumptions regarding the design of the antenna arrays at the BS are limited by a common maximum form factor of e.g. 20 × 20 cm (see Figure 5). For the 3.5 GHz deployment scenario, this allows an antenna array comprising 4 × 4 pairs of cross-polarised AEs (Antenna Elements) with a half-wavelength inter-AE spacing in both dimensions. No subarray structure is assumed for this antenna array, hence all AEs of the 32T32R antenna are fed with independent RF (Radio Frequency) chains and fully digital beamforming is applied. For the 26 GHz scenario, the form factor limitation allows an antenna array comprising 32 × 32 pairs of cross-polarised AEs with half-wavelength inter-AE spacings. Given the very large number of AEs [18], a subarray structure may be applied to limit the number of RF chains and apply hybrid beamforming. (Only) for the corner antenna deployments a mechanical azimuth and downtilt is applied to aim the antennas at the centre of the room. The maximum transmit power and receiver noise figure of the BS are denoted by $p_{MAX,BS}$ and $NF_{BS}$, respectively. Typical values are $p_{MAX,BS} = 30$ dBm and $NF_{BS} = 2$dB.

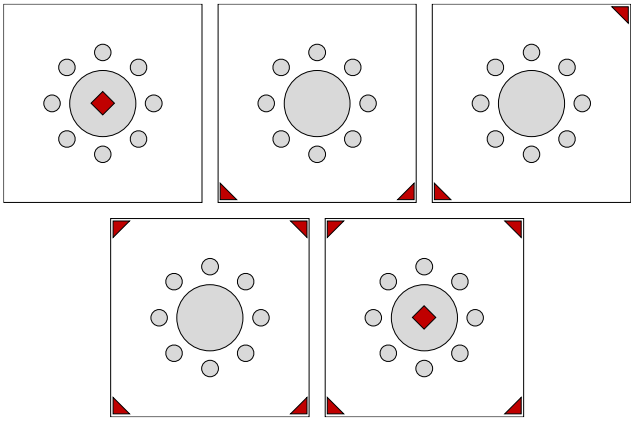

*Figure 4: Network deployment scenarios: line-wise from left to right we have scenarios with $K_{BS}$ = 1, 2 ('2a'), 2 ('2o'), 4 and 5.*

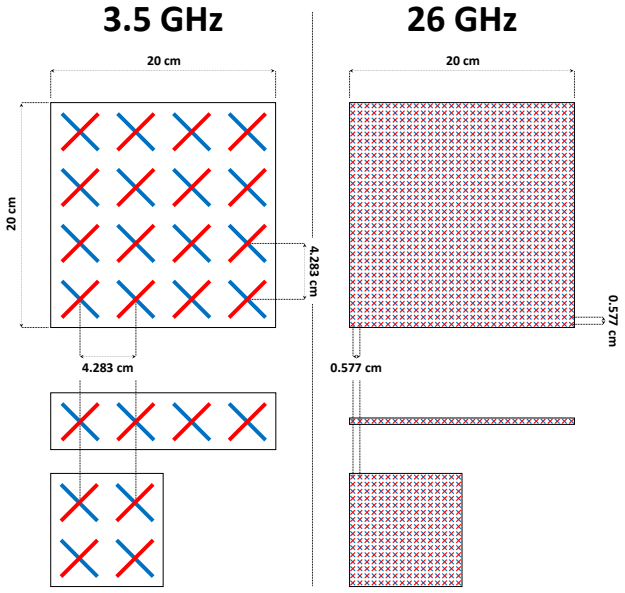


*Figure 5: From top to bottom: assumed BS and UE (HMD, CAM) antennas for the 3.5 GHz (left) and 26 GHz (right) scenarios.*

At the UE side, a distinction is made between *(i)* the UEs incorporated into the HMD, whose communication role is downlink- and hence reception-oriented; and *(ii)* the UEs incorporated in the camera hub ('CAM') in front of each

[1] Based on the 100 Mb/s 'entry-level VR' quality level taken as an example before.

physically present participant on the meeting table, which are primarily taking care of uplink transmissions. The former type of UEs are assumed to be equipped with a linear array of a common length of, say, 20 cm, while for the latter type of UEs a planar array of e.g. 10 × 10 cm can be assumed. For both UE types the assumed array is assumed to be 'filled' with pairs of cross-polarised AEs at half-wavelength spacing (see Figure 5). The maximum transmit power and receiver noise figure of the UE are denoted by $p_{MAX,UE}$ and $NF_{UE}$ with typical values 23 dBm and 8 dB, respectively.

## V. Propagation environment

The characteristics of the propagation environment in terms of path loss, shadowing and multipath fading, are taken from the 'indoor office' scenario specified by 3GPP [19] with implementations provided by Quadriga [20]. Quadriga does not consider any blockage models, but there are several widely adopted blockage models available. 3GPP [19] provides two distinct blockage models: the first model adopts a stochastic method and can be applied if a computationally efficient blockage model is required; the second model adopts a geometric method and can be applied in case a more realistic blockage model is required. Ju et al. [21] propose a four-state Markov model to simulate shadowing due to dynamic human blockage, based on measurements of pedestrians. The model divides a blockage event into four stages, viz. 'unshadowed', 'decay', 'shadowed' and 'rising'. Simpler models are also available. E.g. in [22] a fixed attenuation of 20 dB is introduced if the line-of-sight (LOS) path between the BS and the UE is blocked by a human blocker [23].

As an illustrative example, Figure 6 shows the received power for a UE that is moving through the environment on both the 3.5 GHz and the 26 GHz carrier frequency. Additionally, the impact of two events are demonstrated in the figure: a sudden change in UE orientation (starting at 2 seconds) and the appearance of a human blocker (starting at 4 seconds).

*Figure 6: Impact of head orientation and blockage on the received power level for both a 3.5 GHz and a 26 GHz carrier.*

## VI. Traffic handling

The up/downlink (UL/DL) social VR traffic is handled over a 5G RAN utilising the 3.5/26 GHz carrier and the BS/UE antennas as described above. At each BS antenna a *Grid of Beams* (GoB) comprises a set of pre-defined transmission beams (wideband precoders) in a three-dimensional angular grid covering an angular span of 120° in both the azimuth and elevation planes. Since we assume a persistently on-going VR-aided meeting, we are not concerned with the more coarsely defined access-oriented SS/PBCH-based (Synchronisation Signal / Physical Broadcast CHannel) GoB, but rather model the traffic-oriented CSI-RS (Channel State Information Reference Signals) beams in a more finely granular GoB. Specifically, the proposed grid of CSI-RS beams has an approximate angular granularity of 30° and 4° for the BS antennas depicted in Figure 5. The assumed granularity is based on a first null beamwidth analysis, ensuring minimal inter-beam interference. Given the angular span and the applied granularity, the GoB comprises 5 × 5 = 25 and 31 × 31 = 961 beams for the 3.5 GHz and 26 GHz scenarios, respectively. Figure 7 illustrates the proposed GoB for the azimuth plane (at an elevation of 0°) for both carriers.

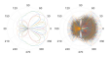

*Figure 7: Azimuth plane view of the Grid of Beams for the 3.5 GHz (left) and 26 GHz (right) carriers.*

Each UE selects the serving BS antenna towards which it experiences the strongest instantaneous channel gain as its serving cell, while its serving cell cluster may be extended beyond this 'best server' by including additional BS antennas towards which the UE experiences an average channel gain no more than $\zeta$ dB lower than towards the best server. With configuration parameter $\zeta$ set to a value greater than 0 dB this implies the use of D-MIMO where a UE may be concurrently served by the aggregation of multiple BS antenna arrays.

At the serving cell, the UE selects the strongest CSI-RS beam, indicated via the CSI-RS Resource Indicator (CRI) feedback, with a correspondingly precoded Physical Downlink Shared CHannel (PDSCH) traffic beam used for downlink transmission. Receive beamforming is used at the UE side in the form of Maximum Ratio Combining (MRC). The combination of a selected CSI-RS beam and a selected MRC configuration at the UE receiver, constitutes a so-called downlink 'beam pair'. Given the use of a TDD carrier, 'beam correspondence' is assumed and the same beam pair is applied for uplink communications. This entails that the precoder of the selected CSI-RS is effectively used as a combiner at the BS for uplink reception, while the combining parameters optimised by the UE for downlink reception are also used for uplink transmission beamforming.

A QoS-oriented packet scheduler governs up- and (sub- or wideband-based) downlink transmissions, given the resources configured by an applied TDD duplexing scheme, configured by a given frame size and an UL/DL resource split that is optimised in line with the asymmetry in traffic load and spectral efficiency. Downlink scheduling is based on periodic CSI feedback from the UE with a configurable periodicity. Head of line packet dropping is applied in the transmit buffers in case the packet's corresponding VR frame is determined not to meet the aforementioned latency budget $\Delta_{RAN}$. An effective downlink-oriented scheduler may be the EXP/PF or M-LWDF schedulers analysed in e.g. [24][1]. In the uplink we propose the use of an open-loop power control scheme with a

target received power density of e.g. -80 dBm/PRB (Physical Resource Block) combined with a fractional path loss compensation factor of 0.8, complemented by an Automatic Transmission Bandwidth (ATB) scheme providing each UE with the number of consequently supportable PRBs. In both the up- (only when the PRB resources are scarce) and the downlink, MU-MIMO-based co-scheduling of multiple UEs in the same time-frequency resources is applied. The applied criterion thereby is that two UEs can be co-scheduled if their selected CSI-RS beams are at least $\kappa$ beams apart in the GoB, with $\kappa \in \{0, 1, 2, \ldots\}$ a configurable scheduling parameter. The smaller $\kappa$, the higher the applied degree of MU-MIMO, while for very large $\kappa$ the MU-MIMO is effectively turned off and only single-user scheduling is used. Note that in the extreme case of $\kappa = 0$ distinct polarisations must be applied to avoid excessive interference.

The adaptive modulation and coding scheme estimates the received Signal-to-Interference-plus-Noise-Ratio (SINR) and maps this to a selected and fed back MCS (Modulation and Coding Scheme) based on a 10% BLER (BLock Error Rate) target and link-level results available from [26]. To adequately compensate for feedback/measurement delay-induced errors in deriving the optimal transmission parameters, a dynamically tuned UE-specific outer-loop link adaptation scheme is applied.

## VII. Concluding Remarks

We have established an integral modelling framework for the feasibility assessment and performance optimisation of the radio access network for Social VR applications in indoor office scenarios. The framework covers in detail the relevant aspects regarding meeting room layout, human behaviour, application-layer specifics, network deployment and configuration, signal propagation and resource management. When integrated in a simulation tool, the involved models can be used for performance optimisation and a derivation of guidelines for indoor network deployment and configuration. In this paper we focused on an indoor business meeting scenario, but our modelling framework provides a good basis for extension towards other Social VR use cases.

A topic for future research exploiting the proposed modelling approach is the development of (potentially AI/ML-based) methods for dynamic configuration of radio access parameters to support Social VR applications. This involves cross-layer adaptation, e.g. the adaptation of application-layer data rates (and hence QoE) to actual radio network capabilities, optimisation of the connect-compute trade-off in the radio access (incl. edge cloud), and pro-active beam management based on human movement predictions. Ultimately, we envision the design of a fully automated 'Social VR network slice', i.e. a self-configuring and autonomously managed 5G (radio) network slice to support Social VR applications.

## References


[1] H. Holma, A. Toskala and T. Nakamura, *'5G technology - 3GPP New Radio'*, John Wiley & Sons, Hoboken, USA, 2020.

[2] 5G PPP, *'5G innovations for new business opportunities'*, white paper, see https://5g-ppp.eu/wp-content/uploads/2017/01/5GPPP-brochure-MWC17.pdf, 2017.

[3] M. Latva-aho and K. Leppänen (eds.), *'Key drivers and research challenges for 6G ubiquitous wireless intelligence'*, white paper, University of Oulu, Finland, 2019.

[4] 3GPP, *'Extended Reality (XR) in 5G'*, TR26.928, v16.0.0, 2020 .

[5] F. Hu, Y. Deng, W. Saad, M. Bennis and A. H. Aghvami, 'Cellular-connected wireless virtual reality: requirements, challenges, and solutions', *IEEE Communications Magazine*, vol. 58, no. 5, 2020.

[6] E. Bastug, M. Bennis, M. Medard and M. Debbah, 'Toward interconnected virtual reality: opportunities, challenges and enablers', *IEEE Communications Magazine*, vol. 55, no. 6, 2017.

[7] S. Dijkstra-Soudarissanane, K. El Assal, S.N.B. Gunkel, F. Ter Haar, R. Hindriks, J.W. Kleinrouweler and O.A. Niamut 'Multi-sensor capture and network processing for virtual reality conferencing', *Proceedings of MMSys '19*, Amherst, USA, 2019.

[8] T.V. Doan, D. You, H. Salah, G.T. Nguyen and F. H.P. Fitzek, 'MEC-assisted immersive services: orchestration framework and protocol', *Proceedings of BMSB '19*, Jeju, South-Korea, 2019.

[9] M.S. Elbamby, C. Perfecto, M. Bennis and K. Doppler, 'Edge computing meets millimeter-wave enabled VR: paving the way to cutting the cord', *Proceedings of WCNC '18*, Barcelona, Spain, 2018.

[10] M. S. Elbamby, C. Perfecto, M. Bennis and K. Doppler, 'Toward low-latency and ultra-reliable virtual reality', *IEEE Network*, vol. 32, no. 2, 2018.

[11] S. Barbarossa, E. Ceci and M. Merluzzi, 'Overbooking radio and computation resources in mmW-mobile edge computing to reduce vulnerability to channel intermittency', *Proceedings of EuCNC '17*, Oulu, Finland, 2017.

[12] M. Chen, W. Saad and C. Yin, 'Virtual reality over wireless networks: quality-of-service model and learning-based resource management', *IEEE Transactions on Communications*, vol. 66, no. 11, 2018.

[13] O. Abari, D. Bharadia, A. Duffield and D. Katabi, 'Cutting the cord in virtual reality', *Proceedings of HotNets '16*, Atlanta, USA, 2016.

[14] S. Fremerey, A. Singla, K. Meseberg and A. Raake, 'AVtrack360: an open dataset and software recording people's head rotations watching 360° videos on an HMD', *Proceedings of MMSys '18*, Amsterdam, The Netherlands, 2018.

[15] E.J. David, J. Gutiérrez, A. Coutrot, M. Perreira Da Silva and P. Le Callet, 'A dataset of head and eye movements for 360° videos', *Proceedings of MMSys '18*, Amsterdam, The Netherlands, 2018.

[16] S. N. B. Gunkel, H. M. Stokking, M. J. Prins, N. van der Stap, F.B.T. Haar and O.A. Niamut, 'Virtual reality conferencing: multi-user immersive VR experiences on the web', *Proceedings of MMSys '18*, Amsterdam, The Netherlands, 2018.

[17] S. Mangiante, G. Klas, A. Navon, Z. GuanHua, J. Ran and M. Dias Silva, 'VR is on the edge: how to deliver 360° videos in mobile networks', *Proceedings of VR/AR Network '17*, Los Angelas, USA, 2017.

[18] E. Björnson, L. Sanguinetti, H. Wymeersch, J. Hoydis and T.L. Marzetta, *'Massive MIMO is a reality – What is next? Five promising research directions for antenna arrays'*, arXiv:1902.07678, 2019

[19] 3GPP, *'Study on channel model for frequencies from 0.5 to 100 GHz'*, TR38.901, v16.1.0, 2019.

[20] Fraunhofer, *'Quadriga – The next generation radio channel model'*, www.quadriga-channel-model.de, 2020.

[21] S. Ju, O. Kanhere, Y. Xing and T. S. Rappaport, 'A millimeter-wave channel simulator NYUSIM with spatial consistency and human blockage', *Proceedings of Globecom '19*, Waikoloa, USA, 2019.

[22] J. S. Lu, D. Steinbach, P. Cabrol and P. Pietraski, 'Modeling human blockers in millimeter wave radio links', *ZTE Communications*, vol. 10, no. 4, 2012.

[23] J. S. Lu, D. Steinbach, P. Cabrol and P. Pietraski, 'Modeling human blockers in millimeter wave radio links', *ZTE Communications*, vol. 10, no. 4, 2012.

[24] F. Afroz, K. Sandrasegaran and P. Ghosal, 'Performance analysis of PF, M-LWDF and EXP/PF packet scheduling algorithms in 3GPP LTE downlink', *Proceedings of ATNAC '14*, Melbourne, Australia, 2014.

[25] J.-H. Rhee, J.M. Holtzman and D.-K. Kim, 'Scheduling of real/non-real time services: adaptive EXP/PF algorithm', *Proceedings of VTC '03 (Spring)*, Jeju, South-Korea, 2003.

[26] S. Pratschner, B. Tahir, L. Marijanovic, M. Mussbah, K. Kirev, R. Nissel, S. Schwarz and M. Rupp, 'Versatile mobile communications simulation: the Vienna 5G link level simulator', *EURASIP Journal on Wireless Communications and Networking*, 2018.