Challenges & Practical Considerations for VR Performance in Real-Time Communication with an overview in Telemedicine

Anil Naik, University of California[1], Deepak Singh, MS, University of Pittsburgh[2]

Abstract: The innovative applications of virtual reality (VR) in real-time communications (RTCs) are revolutionizing how we connect and interact with others in the digital realm via social networking, teleconferencing, virtual meetings, and remote collaboration, bridging distances and enhancing communication experiences. A special focus was recently oriented on integrating such applications into the telemedicine field via collecting real-time health data for chronic disease management and post-operative monitoring. In this work, the underlying principles, benefits, and challenges associated with VR adoption in RTC applications were explored with a special focus on its potential upsides and hurdles in the telemedicine use-case. It is crucial to select the most suitable architectural model for VR-based communication systems, whether semantic, scene graph-based, generic, or distributed, based on the target application and the desired outcome. Path-loss and wideband channel models are essential tools to simulate the path-loss and wide-sense stationary behavior of propagation channel in urban or rural environments, including Okumura-Hata, Saleh & Valenzuela, and directional channel models. Deep reinforcement learning (RL) can be used to tackle the challenge of internet congestion control by utilizing RL-based frameworks (e.g., PCC Vivace, Aurora, Copa, transmission control protocol (TCP) Cubic, and BBR) to design congestion control protocols. In order to meet the requirements of lower latency and massive data transmission in VR-based RTCs, the multi-path cooperative route (MCR) schemes are significantly efficient.

Keywords: *Virtual reality (VR), real-time communication (RTM), architectural models, practical considerations, telemedicine*

1. Introduction

VIRTUAL REALITY (VR) technology has rapidly gained momentum across numerous domains, revolutionizing how we experience and interact with computer-generated environments. Thus, the adoption of VR technology burgeoned in diverse applications, spanning fields such as entertainment, education, healthcare, engineering, and others.

The entertainment industry has been at the forefront of VR adoption, as it offers unparalleled opportunities for immersive gaming, interactive storytelling, virtual theme parks, cinematic experiences, and immersive live events [1]. The current advancements in graphics, audio, and haptic technologies that enhance user engagement, lead to more captivating and memorable entertainment VR experiences. Meanwhile, VR technology has the potential to revolutionize education by providing learners with immersive and interactive environments via VR applications in education and vocational training. Adopting VR technology in education and training applications can facilitate experiential learning, improve knowledge retention, enhance practical skills, and enable virtual field trips, fostering a more engaging and effective learning process [1], [2].

VR technology has been embraced in healthcare to enhance diagnostics, treatment, rehabilitation, and training. The use of VR in several healthcare applications (e.g., medical simulations, surgical training, pain management, mental health therapy, and patient rehabilitation) highlights its benefits in providing realistic scenarios, reducing risks, improving patient outcomes, and enabling healthcare professionals to refine their skills in a safe and controlled environment. Concurrently, VR is transforming the engineering and design processes, allowing professionals to visualize and manipulate complex structures before they are physically built [3]. Therefore, utilizing VR technology in architectural design, product prototyping, virtual prototyping, and collaborative design reviews; improves the overall efficiency of the engineering and design workflows.

Meanwhile, VR technology has transcended traditional boundaries and emerged as a powerful tool for real-time communication (RTC). Nowadays, the innovative applications of VR are revolutionizing how we connect and interact with others in the digital realm via social networking, teleconferencing, virtual meetings, and remote collaboration, bridging distances and enhancing communication experiences. VR is transforming social networking by providing immersive and interactive platforms for users to connect and engage, enabling non-verbal communication cues, and creating virtual communities that transcend geographical boundaries [4]. Furthermore, teleconferencing and virtual meetings can be revolutionized by VR technology to meet and interact in a shared virtual space; enhancing engagement, improving communication effectiveness, and creating a more immersive and inclusive meeting environment. Additionally, VR offers unique opportunities for cultural exchange and language learning, allowing individuals to immerse themselves in different cultures and practice languages in realistic environments. Such VR adoption in cultural immersion experiences, language learning platforms, and virtual travel; can break down cultural barriers, promote empathy, and provide authentic language practice opportunities.

Recently, a special focus was oriented on the applications of VR-based RTC, including voice and video streaming, in the telemedicine field. Collecting real-time health data is one of the current revolutionizing trends in telemedicine, which can be enabled via VR-based RTCs [5]. In this regard, patient's data is transmitted to healthcare providers using wearable devices or sensors. Integrating such data into a virtual environment allows real-time and remote monitoring of patients, which can be crucial for chronic disease management and post-operative monitoring [6].

While virtual reality holds immense potential in RTC, it also faces challenges such as cost, accessibility, hardware limitations, and user comfort. The adoption of VR in RTC heavily relies on

Anil Naik, University of California[1],
anilgnaik9@gmail.com, ORCID: 0009-0000-8625-6952

Deepak Singh, MS, University of Pittsburgh[2]
deepaksingh.rv@gmail.com

network bandwidth and low latency to deliver a seamless experience [7]. VR-based RTC applications require high-speed internet connections to transmit large amounts of data, including audio, video, and positional tracking information. Inadequate network infrastructure or high latency can introduce delays, leading to lag and degraded user experience. As VR adoption increases, ensuring robust and reliable network connectivity becomes crucial to mitigate these issues. Vahora et al. highlighted one of the VR real-time challenges of determining the processing time to synchronize haptic and graphic updates [8]. Such an aspect can be crucial for creating a stable haptic servo loop within VR-based RTC applications. Furthermore, previously reported work focused on developing metadata communication protocols to enhance the real-time synchronization between BIM data and VR devices [9]. The use of a metadata interpretation system and cloud-based infrastructure in that work resulted in automatic and simultaneous real-time synchronization. Ge et al. proposed a software-defined networking (SDN) architecture to overcome the latency and data transmission issues of 5G networks in VR applications [10]. The proposed architecture facilitated the VR wireless transmission in 5G small cell networks via a multi-path cooperative route (MCR) scheme.

Hence, this work's main aim is to explore the underlying principles, benefits, and challenges associated with VR adoption in RTC applications. The concepts of different types of VR technologies are briefly presented, followed by a detailed insight into the various architectural models of VR communication systems. Furthermore, this work sheds light on the technical challenges and practical considerations of VR-based RTC applications, including internet congestion, and channel propagation. Most importantly, practical solutions are presented to overcome communication-related drawbacks and enhance the performance of VR systems in RTC applications, including space-time characterization techniques, path-loss prediction models, wideband channel models, deep reinforcement leaning, and multi-path cooperative transmissions.

2. Types of VR Technologies

Virtual Reality (VR) technology encompasses various modalities, each catering to specific requirements and providing unique user experiences that offer distinct experiences and capabilities. The following subsections briefly elucidate the unique characteristics and preferences of each VR modality.

2.1. Enhanced VR

Enhanced VR highlights advancements in haptic feedback, spatial mapping, eye-tracking, and augmented reality integration, providing heightened sensory experiences and advanced interactions within virtual environments.

2.2. Immersive VR

Immersive VR refers to highly immersive experiences that transport users into entirely virtual environments. The concept of Immersive VR can be characterized by the use of high-end VR headsets, motion tracking systems, and spatial audio to create a sense of presence, enabling users to interact with objects and navigate within a virtual space.

2.3. Hybrid VR

Hybrid VR combines elements of both virtual and augmented reality to create mixed reality experiences, where virtual objects coexist and interact with the real-world environment. Such a principle can be experienced via head-mounted displays (HMDs) with built-in cameras or passthrough features to enable users simultaneously see and manipulate both virtual and physical elements.

2.4. Desktop VR

Desktop VR refers to VR experiences that are primarily accessed through desktop computers or laptops, where users typically utilize lower-end VR headsets and hand controllers, often with six degrees of freedom (6DoF) tracking. Desktop VR provides advantages such as affordability, ease of setup, and access to a wide range of VR content.

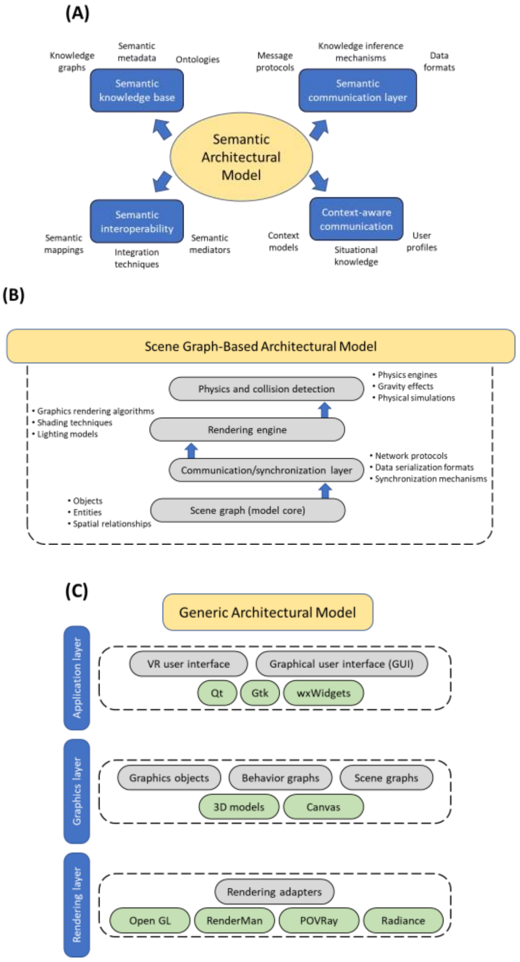
2.5. Quick-time VR

Quick-Time VR, or 360-degree VR, allows users to view and navigate static or pre-recorded 360-degree panoramic images or videos, where users can explore and interact with the captured environment with limited real-time interactivity, providing an immersive viewing experience.

3. Architectural Models of VR Systems

3.1. Semantic

Integrating semantic technologies to enhance communication, data interoperability, and knowledge representation is highly vital for effective information exchange in VR communication systems. In this regard, the semantic architectural model facilitates semantic interoperability, context-aware communication, and enhanced collaboration within VR environments. As shown in Fig. 1(A), the key semantic technologies employed in the semantic architectural model include ontologies, semantic metadata, linked data, and semantic web services.



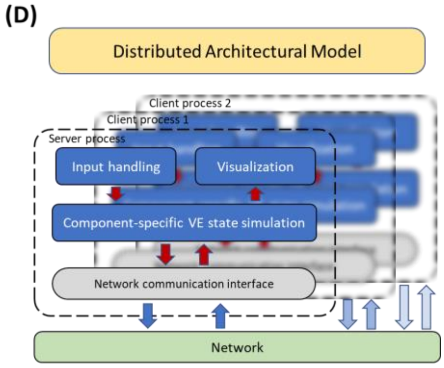


Fig. 1. Key components of architectural models in VR communication systems; semantic (A), scene graph-based (B), generic (C) (adapted from [11]), and distributed (D) (adapted from [12]).

The semantic communication layer forms the core of the semantic architectural model, which can facilitate the exchange and interpretation of meaning-rich information among VR users. The components and functionalities of the layer involve semantic messaging protocols (e.g., RDF-based protocols), data representation formats (e.g., RDF, OWL), and knowledge inference mechanisms. It emphasizes how the semantic communication layer facilitates the exchange and interpretation of meaning-rich information among VR users.

Meanwhile, the semantic knowledge base serves as a central repository of shared knowledge within VR communication systems via the creation, management, and utilization of ontologies, semantic metadata, and knowledge graphs; enabling semantic context enrichment, semantic search, and knowledge-driven interactions in VR environments. Correspondingly, semantic interoperability ensures seamless communication and data exchange among heterogeneous VR systems; using semantic mappings, semantic mediators, and semantic integration techniques. Additionally, context-awareness is crucial for personalized and adaptive communication experiences in VR systems by integrating context models, user profiles, and situational knowledge within the semantic architectural model, enhancing context-aware communication and interaction.

3.2. Scene graph-based

The scene graph-based architectural model of VR communication systems comprises several key components that facilitate the representation, synchronization, and rendering of virtual environments (Fig. 1(B)). The scene graph forms the core of the architectural model, as it represents the hierarchical structure of the virtual environment, encompassing objects, entities, spatial relationships, and properties. Such a model core organizes the visual and interactive elements of the virtual scene, enabling efficient rendering, interaction management, and synchronization across multiple users.

Another component of the scene graph-based architectural model is the communication and synchronization layer, which handles the exchange of scene graph information among VR users. The layer utilizes network protocols, data serialization formats, and synchronization mechanisms to ensure the access of all participants in the VR communication system to the latest version of the scene graph, allowing real-time collaboration, concurrent editing, and consistent rendering across distributed users. In this regard, it is crucial to avoid confusion between scene graphs with data structures utilized in visibility culling or collision queries (e.g., ABT, BSP, KdT, and octrees) [13]. Meanwhile, the rendering engine is responsible for rendering the virtual scene based on the information stored in the scene graph via graphics rendering algorithms, shading techniques, lighting models, and optimization strategies. Consequently, the visual representation of virtual environment is consistent across different VR devices, and the scene graph updates are reflected in real-time. Physics and collision detection components simulate realistic physics behaviors and handle collision detection within the VR systems. The interaction between virtual objects, gravity effects, and physical simulations is governed by physics engines and collision detection algorithms.

3.3. Generic

The generic architectural model of VR communication systems encompasses three main components fundamental to the design and implementation of such systems: application layer, graphics layer, and rendering layer (Fig. 1(C)). Steinicke et al. highlighted the primary role of each layer, where the application layer deals with creating VR user interfaces by utilizing graphical user interface (GUI) toolkits (e.g., Qt, Gtk, and wxWidgets) [11]. While the graphics and rendering layers focus on managing graphics objects via controlling scene/behavior graphs and utilizing different rendering adapters, respectively. The specific design and implementation of each component may vary based on the system's goals, platform, and targeted use cases.

3.4. Distributed

The distributed architectural model of VR communication systems comprises various components that enable the distribution of VR experiences across multiple devices and locations, allowing users to collaborate, communicate, and share virtual environments seamlessly. As shown in Fig. 1(D), the input handling unit is one of the model's three main units, which can read data from user input devices and convert them to control actions; to be transmitted to the unit of component-specific VE state simulation for simulating current VE state for the component [12]. Finally, the visualization unit can render the current VE state for each component. Then, the interaction of the current process with other processes and the initialization of new clients can be controlled via data replication management and reckoning techniques. One of the early applications of the distributed model in VR systems is Cybertennis, conducted in Geneva in 1997, where two users played tennis together remotely (40 miles away); while a virtual clone of Marilyn Monroe acted as a referee [14].

4. Technical Challenges & Limitations of VR in RTC's

4.1. Internet congestion

Internet congestion can increase the delay between sending a request and receiving a response, known as network latency, which significantly influences the performance of VR in real-time communications. In this regard, high latency caused by congestion can lead to delays between a user's actions and the corresponding feedback in the virtual environment, resulting in a noticeable lag and decreased immersion. Thus, low latency is crucial to decrease such a congestion-resulting delay.

Meanwhile, congestion may also result in packet loss, where data packets sent over the network fail to reach their destination, leading to missing or distorted visual and audio information (i.e., visual artifacts, stutters, and complete disruptions). Additionally, VR-based applications of real-time communication require a significant amount of bandwidth to transmit high-quality audio/visual content in real-time. Correspondingly, internet congestion limits the available bandwidth, resulting in degradation of the visual fidelity (e.g., reduced image quality, compression artifacts, and complete freezes in the VR environment). Thus, addressing such challenges requires a combination of network optimization, and improved congestion management techniques, to minimize latency, mitigate packet loss, increase available bandwidth, and prioritize VR traffic.

4.2. Channel propagation

Signal quality can be dramatically impacted by channel propagation and the factors affecting how the wireless signal travels through the environment from the transmitter to the receiver, including signal attenuation, reflection, diffraction, and interference. The significant latency and/or signal degradation resulting from channel propagation may lead to less responsive VR-based RTCs. Consequently, VR-based voice/video streaming can be distorted with reduced image quality and increased motion sickness. Additionally, channel propagation characteristics can cause delays in signal transmission, associated with signal echoes and interference, owing to reflections and multi-path propagation. Furthermore, the stability and consistency of the wireless connection would be greatly influenced by such a propagation issue. Channel propagation determinants, such as signal fading or interference, can be the main reasons for data streaming interruptions, glitches, or disconnections. Hence, it is crucial to consider channel propagation characteristics during system design and deployment, utilizing techniques such as proper antenna placement, signal strength optimization, and mitigation of multi-path interference. Correspondingly, the signal quality can be maintained, and wireless communication stability can be ensured.

5. Proposed Solutions & Practical Considerations

5.1. Space-time characterization

The optimization of multiple-input multiple-output (MIMO) systems requires accurate knowledge of the spatial features of channel propagation. Thus, estimating the power density at the receiver (RX) and the transmitter (TX) at each access delay can be conducted within the double-directional measurements [15]. In such measurements, a real antenna array is used at the base station (BS), while a virtual linear array is composed by successive measurement points at the mobile station (MS). As shown in Fig. 2(A), the results of double directional measurements in an urban environment resulted in different spreading behavior at BS and MS, corresponding to the power delay profile [16].

In another work, double-directional propagation channel measurements were conducted for THz band (141-148.5 GHz range), employing double-directional channel sounding using RF-over-Fiber (RFoF) extensions for measurements over 100 m distance in urban scenarios [17]. The path-loss curves maintained their shape with a slight enhancement when using longer intermediate frequency (IF) cables (i.e., higher amplification) (Fig. 2(B)) [17]. Furthermore, line-of-sight (LoS) measurements indicated a significant signal difference (around 70 dB) compared with the noise floor of the system, which was attributed to the dynamic range available for the measurement of multipaths and non-line-of-sight (NLoS) signals, using the current setup (Fig. 2(C)) [17].

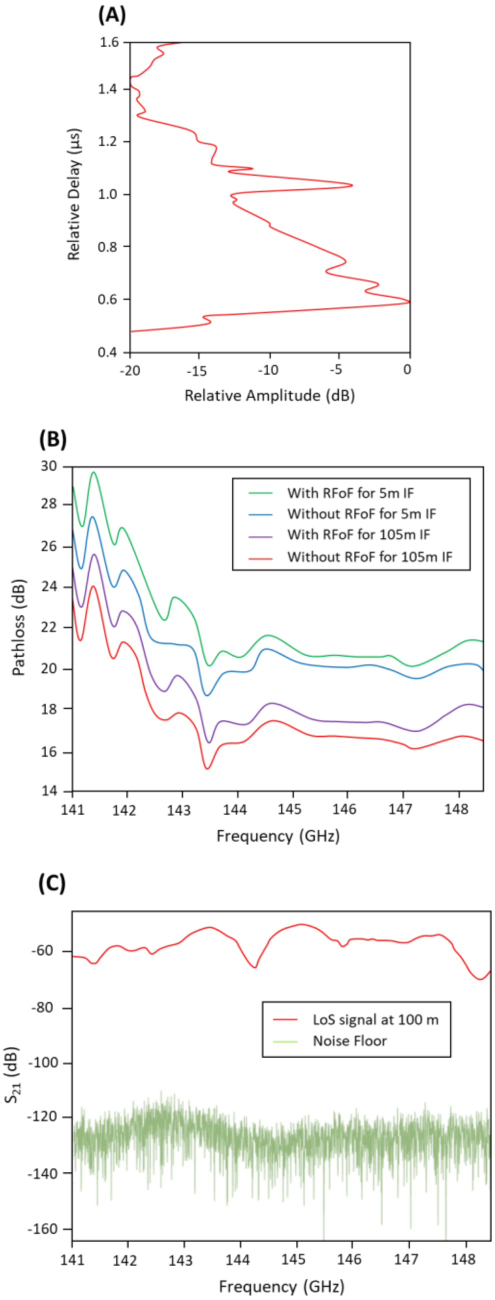


Fig. 2. Power delay profile within double directional measurement in an urban area (A), directional THz measurements of path-loss via different configurations of RFoF setup over a 2.4m indoor channel (B), and LoS in comparison with system noise floor (C) (reprinted from [15]–[17]).

Raouf et al. used Helikite flying over urban and rural areas to collect data related to spectrum measurements of various CBRS bands [18]. As displayed in Fig. 3, CBRS bands in urban environments showed much more activity compared with that in rural areas [18].

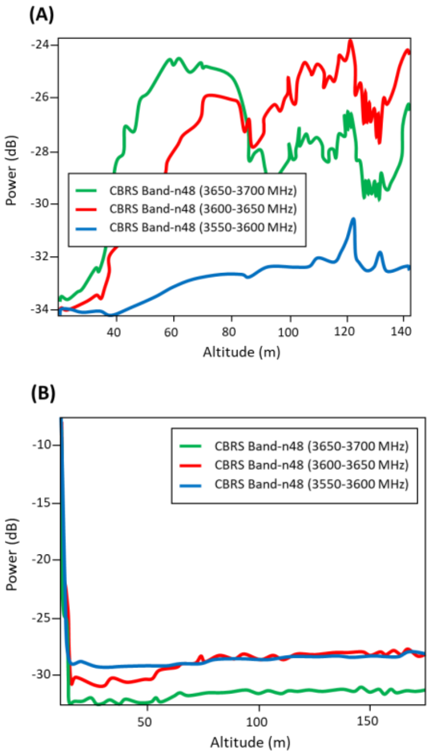


Fig. 3. Spectrum occupancy vs. altitude in different CBRS bands for urban (A), and rural (B) environments (reprinted from [18]).

5.2. Wideband models

Wideband channel models are important tools to simulate the wide sense stationary behavior of propagation channel in each environment.

5.2.1. Saleh & Valenzuela model

The model was proposed in the late 1980s, representing the concept of asset of rays with similar features (e.g., excess delay, the direction of arrival, etc.), where channel impulse response consists of different clusters (Fig. 4(A)) [19]. In this model, clusters are modeled by an exponential decay of the power in the time domain, and the arrival time of clusters and rays is defined by a Poisson random process. Such a mathematical concept is represented by the following formulas [15], [19]:

$$h(\tau) = \sum_{l=1}^{l}\sum_{k=1}^{K_l}\beta_{k,1}e^{j\theta_{k,1}}\delta(\tau - T_l - \tau_{k,l}) \qquad (1)$$

$$\beta_{k,l}^2 = \beta^2(0,0)e^{-T_l/\Gamma}e^{-\tau_{k,l}/\gamma} \qquad (2)$$

$$p(T_l - T_{l-1}) = \Lambda e^{-\Lambda(T_l - T_{l-1})} \qquad (3)$$

$$p(\tau_l - \tau_{l-1}) = \lambda e^{-\lambda(\tau_l - \tau_{l-1})} \qquad (4)$$

Where, $\beta^2(0,0)$ is the power of the first ray of the first cluster, l is the number of clusters, K_l is the number of rays of the cluster, T_l is the lth cluster arrival time, $\tau_{k,l}$ is the kth ray in the lth cluster arrival time, Λ is the clusters rate, Γ is the inter-cluster slope, λ is ray arrival time, and γ is the intra-cluster slope.

Different improvements can be added to the original model in terms of new cluster shape, new statistical laws of clusters/rays arrival time, considering the frequency response, and statistical attenuation on the amplitude of rays [15]. In this regard, the simulation of the IEEE 802.15.4a model, considering three clusters of around 670 rays, showed a contrary line of power decay to the initial Saleh and Valenzuela model (Fig. 4(B)) [20].

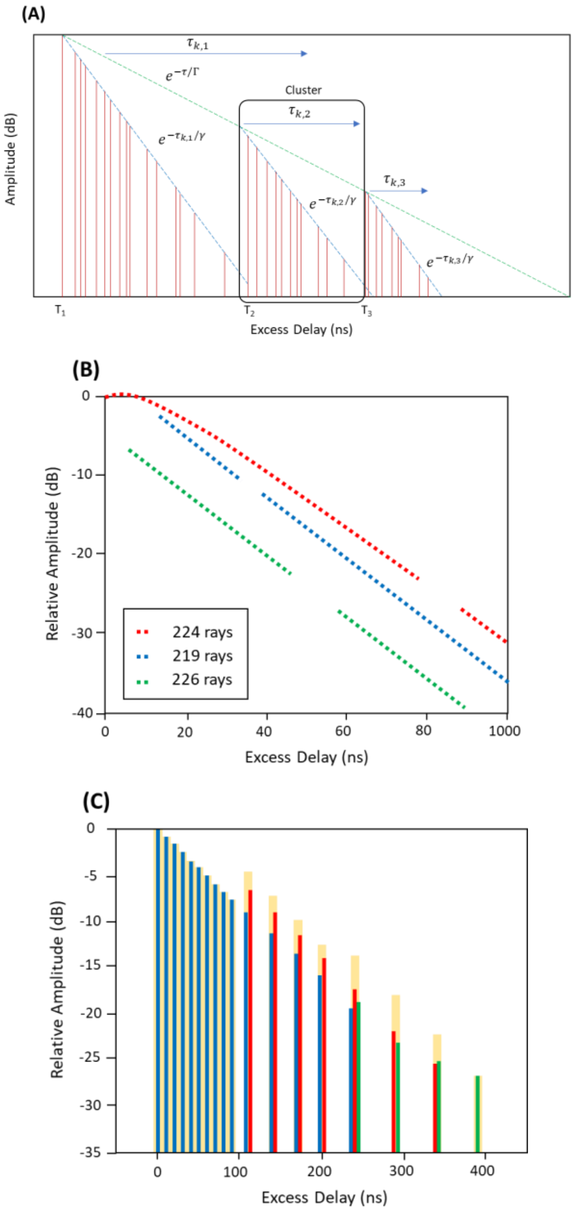


Fig. 4. Principle of Saleh and Valenzuela model (A), an example of the simulation of IEEE 802.15.4a model (Office NLOS) (B), and IEEE 802.11n channel propagation model (Model D) (C) (reprinted from [15]).

5.2.2. Directional channel model

Directional channel models are used to study systems throughput/coverage improvements via MIMO antennas [15]. In such systems, MIMO channel models are based on existing single-input and single-output (SISO) channel models, enabling backward compatibility. In this regard, one of the MIMO channel models, proposed by the 802.11n standardization group, relied on the ESTI BRAN SISO model, where the sum of three clusters was demonstrated corresponding to the initial power distribution panel (PDP) (Fig. 4(C)) [15].

5.3. Path-loss prediction models

Path-loss prediction models are mathematical tools that can be used to predict channel propagation issues and optimize network design accordingly. Such models describe the attenuation or loss of signal strength as it propagates through a wireless communication channel, taking into consideration factors, including distance, frequency, antenna characteristics, and environmental conditions.

5.3.1. Okumura-Hata model

The Okumura-Hata model was initially proposed by Okamura in the 1980s, and improved by Hata, for calculating path loss in urban and suburban environments [21], [22]. The general formula of Okumura-Hata can be expressed as follows [23], [24]:

$$P_L = A(hm) + B(hm)log_{10}(d) + C(hm) \qquad (5)$$

Where P_L (dB) is the path loss, d (km) is the distance between the transmitter (base station) and receiver (mobile station), $A(hm)$ is the mobile station antenna height correction term, and $B(hm)log_{10}(d) + C(hm)$ is the distance-dependent correction term. Such a formula theoretically estimates the path loss due to free space loss, ground reflection loss, diffraction loss, and building penetration loss based on the distance between the transmitter and receiver. The values of correction terms vary, corresponding to the frequency range and environment type (urban or suburban). Nevertheless, the Okumura-Hata model suffers from its inaccuracies in rural and mountainous environments, in addition to the sensitivity to parameter values and lack of generalization.

5.3.2. Walfisch-Ikegami model

The model was developed as a combination of Walfisch and Ikegami models to improve the path-loss estimation by considering more characteristic data of the environment, including buildings heights, road widths, buildings separating distances, and road orientation concerning direct radio path [25]. Nevertheless, the model is not yet deterministic as no topographical data is included. The general formula of the model is given as follows [24]:

For free line-of-sight (LoS) situations:

$$P_L = 42.6 + 26\ log_{10}(d) + 20\ log_{10}(f) \quad \text{for } d \geq 20 \text{ m} \quad (6)$$

For non-line-of-sight (NLoS) situations:

$$P_L = \begin{cases} P_{L_0} + P_{L_{rts}} + P_{L_{msd}} \\ P_{L_0} \end{cases} \qquad (7)$$

Where f (MHz) is the frequency band, P_{L_0} (dB) is the free space loss, $P_{L_{rts}}$ (dB) is the roof-top-to-street diffraction and scatter loss, $P_{L_{msd}}$ is multiple screen diffraction loss.

5.4. Deep reinforcement learning (RL)

Deep reinforcement learning (RL) can be one of the most promising and applicable AI tools to tackle the challenge of internet congestion control. In this regard, congestion control protocols can be developed via a deep RL approach to be employed as a local history of traffic patterns and network conditions for enhanced selection of sending rate. Formulating internet congestion control as an RL task involves solving a sequential decision-making problem within the developed RL framework.

Several RL-based frameworks can be used to design congestion control protocols, such as performance-oriented congestion control (PCC Vivace) [26], Aurora [27], Copa [28], transmission control protocol (TCP) Cubic [29], [30], BBR [31], Remy CC [32], and others. Such RL algorithms consider the agent as the traffic sender, and the actions can be translated to changes in sending rates. In this regard, Jay et al. developed Aurora an open-sourced code for evaluating RL-algorithms performance in internet congestion control. As shown in Fig. 5, the average throughput changed corresponding to the latency for each RL algorithm. Aurora and BBR achieved reasonable throughput with the superiority of Aurora in terms of low latency. In contrast, TCP Cubic showed the highest throughput but with increased latency which caused unsuitable performance within a broad range of network environments, as previously reported [29]. The most important aspect of utilizing RL-based protocols in congestion control is realizing the indication of packet loss via distinguishing congestion loss (i.e., related to exceeding network capacity) from non-congestion loss (e.g., handover between mobile base stations).

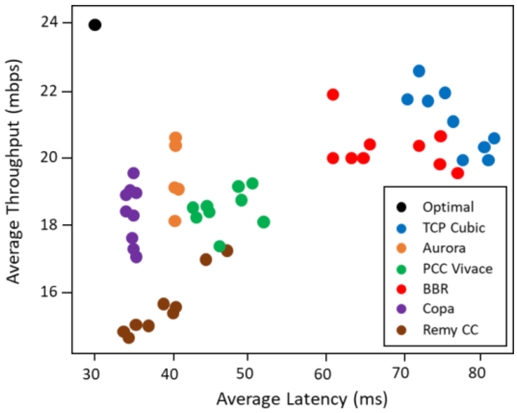
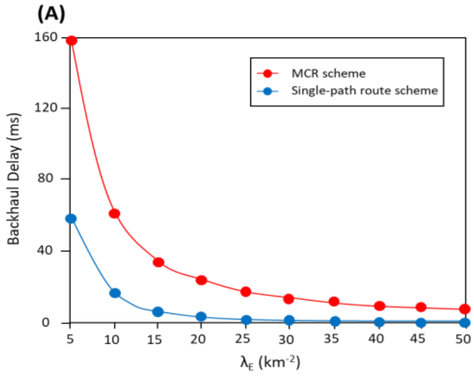


Fig. 5. Average throughput and latency (dynamic link on base latency of 32 ms, 500 packet queue, and no random loss; reprinted from [27]).

5.5. Multi-path cooperative transmissions

One of the proposed solutions to meet the requirements of lower latency and massive data transmission in VR-based RTCs is considering the multi-path cooperative route (MCR) schemes. Utilizing MCR schemes with a software-defined networking (SDN) architecture can facilitate transmissions in VR wireless networks. In this regard, Ge et al. proposed a service-effective energy optimization (SEEM) algorithm based on a delay model of MCR scheme, to enhance the wireless transmission in 5G small cell networks [10]. Simulation results indicated the superiority of MCR scheme to the conventional single-path route in terms of delay and service-effective energy (SEE). As shown in Fig. 6(A), it was clear that using MCR scheme resulted in a lower delay, concerning the density of edge data centers (EDCs), than that of the single-path route scheme, with an average backhaul delay of around 28% over 45 km^{-2} λ_E [10]. Furthermore, SEE exhibited from MCR scheme was reduced by about 11.5% (0.3206 E6 J.m^{-2}) compared with that under the single-path route scheme (Fig. 6(B)). Such results indicated that considering MCR-based algorithms (SEEM) can minimize the network energy consumption with lower delay, making it suitable for VR-based RTC applications.



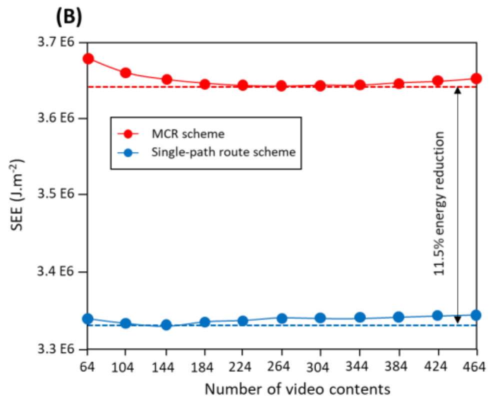


Fig. 6. Simulation results of backhaul delay (A), and SEE (B); considering MCR scheme and single=path route scheme (reprinted from [10]).

6. Implications of VR based RTC on Telehealth

6.1. Potential Upsides of VR-anchored RTC in Telehealth

Amplified Patient Involvement and Experience: The capability of virtual reality (VR) to replicate various environments can provide patients with a more captivating and interactive encounter. This can escalate patient involvement and enhance their overall healthcare journey. VR can prove particularly beneficial in areas such as mental wellness therapy, management of pain, and rehabilitation processes [33] [34].

Enhanced Diagnostic and Therapeutic Abilities: The potential of VR to facilitate the process of diagnostics and treatment is significant. For instance, VR can aid in the three-dimensional visualization of intricate anatomical structures, making surgical planning more efficient. In physical therapy and rehabilitation, VR simulations can offer an interactive and motivational environment for patients [35].

Healthcare Accessibility in Remote Locations: VR-oriented real-time communication (RTC) can help healthcare providers connect with patients in far-flung and underserved regions, diminishing the disparity in accessibility to healthcare services.

Training and Skill Enhancement for Healthcare Professionals: VR offers a secure and controlled setting for healthcare professionals to practice various medical procedures and improve their clinical expertise [36].

6.2. Potential Hurdles of VR-anchored RTC in Telehealth

Infrastructure and Technical Issues: The success of VR-oriented RTC significantly depends on robust internet connectivity and advanced hardware, which might not be readily available to all, especially in remote areas [37] [41].

Data Security and Privacy: Ensuring the security and privacy of patient data in VR-oriented RTC is paramount, given the sensitive nature of healthcare data. Any breaches can have serious repercussions including loss of patient trust, legal consequences, and financial losses [38].

Ease of Use and User Acceptability: For the successful implementation of VR-based RTC, it should be user-friendly and acceptable to both healthcare providers and patients. This includes the design of the VR system, its usability, and the comfort level of users with the technology.

7. Strategies to optimize the potential of VR-based RTC in Telehealth

Establishing Reliable and High-Speed Network Links: Investment in digital infrastructure, such as high-speed broadband, is crucial to enable uninterrupted VR-based RTC. Additionally, robust congestion control protocols can improve network performance and stability, minimizing latency and packet loss [39][41]

Instituting Robust Data Security Measures: Data encryption, secure transmission protocols, and strong authentication mechanisms can help protect sensitive patient data during VR-based RTC. Regular security audits and updates can further enhance data security.

User Acceptance and Training Programs: For successful adoption, programs focused on user acceptance and training should be implemented. This includes training healthcare professionals on the use of VR systems, as well as educating patients about the potential benefits and risks of VR-based RTC [40].

8. Conclusion

In this study, the challenges and practical considerations of virtual reality (VR)-based real-time communication (RTC) were comprehensively discussed. Internet congestion can be a critical issue by increasing the network latency, which significantly influences the performance of VR in RTCs. Furthermore, congestion may result in packet loss, leading to missing or distorted visual and audio information (i.e., visual artifacts, stutters, and complete disruptions). In this regard, using reinforcement learning (RL)-based algorithms can be very helpful in designing congestion control protocols, such as PCC Vivace, Aurora, Copa, transmission control protocol (TCP) Cubic, and Remy CC. Such RL algorithms consider the agent as the traffic sender, and the actions can be translated to changes in sending rates. The most crucial aspect of utilizing RL-based protocols in congestion control is realizing the indication of packet loss via distinguishing congestion loss (i.e., related to exceeding network capacity) from non-congestion loss (e.g., handover between mobile base stations).

Signal quality can be impacted by channel propagation through signal attenuation, reflection, diffraction, and interference, affecting the wireless signal within its travel from transmitter to receiver. Such high latency and/or signal degradation resulting from channel propagation, may lead to distortion in VR-based voice/video streaming with reduced image quality and increased motion sickness. In such an aspect, path-loss and wideband channel models are critical tools to simulate the path-loss and wide sense stationary behavior of propagation channel in urban or rural environments, such as Okumura-Hata, Saleh & Valenzuela, and directional channel models. Nevertheless, from the authors' perspective, the real challenge nowadays is developing an adaptive and robust model that can perform within any configuration (i.e., environment, distance, database, and type of link). Such a model may take advantage of all the up-to-date models (e.g., statistical, empirical, geographical, etc.) to be generalized and adopted without discontinuity and with accurate prediction.

Such integration of modeling techniques into overcoming the technical issues of VR-based RTCs shall include specific enhancements on the level of adoption in telemedicine applications. The transmission process of patient's real-time data to the healthcare provider or system for analysis shall require a reliable and low-latency data transmission mechanism. In this regard, utilizing RL-based optimizing protocols will be essential to ensure high-speed and stable network connections and employ efficient data compression techniques. Furthermore, accurate time-stamping of patient's data will be crucial so that healthcare providers can analyze the physiological responses to specific interactions within the VR session. Thus, employing synchronized clocks and time-stamping mechanisms via specific frameworks (e.g., protocols include WebRTC), shall maintain the integrity and context of the collected data. Also, future research should be focused on the storage and security of real-time patient's data, which needs to be securely stored and managed to maintain privacy and comply with relevant regulations. It is vital to consider the specific requirements and constraints of VR-based telemedicine applications and continuously update the technical solutions to adapt to evolving needs and advancements in VR technology.

Disclaimer

This disclaimer informs readers that the views, thoughts, and opinions expressed in this paper belong solely to the authors, and not necessarily to the author's employer, institution or other groups or individuals.

References

[1] V. Theoktisto and M. Fairén, "Enhancing collaboration in virtual reality applications," *Comput. Graph.*, vol. 29, no. 5, pp. 704–718, 2005, doi: https://doi.org/10.1016/j.cag.2005.08.023.

[2] V. Hohmann, R. Paluch, M. Krueger, M. Meis, and G. Grimm, "The virtual reality lab: Realization and application of virtual sound environments," *Ear Hear.*, vol. 41, no. Suppl 1, p. 31S, 2020.

[3] M. Noghabaei, A. Heydarian, V. Balali, and K. Han, "A survey study to understand industry vision for virtual and augmented reality applications in design and construction," *arXiv Prepr. arXiv2005.02795*, 2020

[4] S. Tea, K. Panuwatwanich, R. Ruthankoon, and M. Kaewmoracharoen, "Multiuser immersive virtual reality application for real-time remote collaboration to enhance design review process in the social distancing era," *J. Eng. Des. Technol.*, vol. 20, no. 1, pp. 281–298, 2022

[5] G. Riva and L. Gamberini, "Virtual reality in telemedicine," *Telemed. J. e-Health*, vol. 6, no. 3, pp. 327–340, 2000.

[6] P. Luitel *et al.*, "A novel solution for real-time wireless networks: Telehealth psychological services in rural area," *Trans. Emerg. Telecommun. Technol.*, vol. 30, no. 3, 2019, doi: 10.1002/ett.3550

[7] B. Kraut and J. Jeknić, "Improving education experience with augmented reality (AR)," in *2015 38th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*, 2015, pp. 755–760.

[8] F. Vahora, B. Temkin, T. M. Krummel, and P. Gorman, "Development of real-time virtual reality haptic application: Real-time issues," in *Proceedings 12th IEEE Symposium on Computer-Based Medical Systems (Cat. No. 99CB36365)*, 1999, pp. 290–295.

[9] J. Du, Z. Zou, Y. Shi, and D. Zhao, "Zero latency: Real-time synchronization of BIM data in virtual reality for collaborative decision-making," *Autom. Constr.*, vol. 85, pp. 51–64, 2018, doi: https://doi.org/10.1016/j.autcon.2017.10.009

[10] X. Ge, L. Pan, Q. Li, G. Mao, and S. Tu, "Multi-path cooperative communications networks for augmented and virtual reality transmission," *IEEE Trans. Multimed.*, vol. 19, no. 10, pp. 2345–2358, 2017

[11] F. Steinicke, T. Ropinski, and K. Hinrichs, "A generic virtual reality software system's architecture and application," in *Proceedings of the 2005 international conference on Augmented tele-existence*, 2005, pp. 220–227.

[12] V. Y. Kharitonov, "A software architecture for high-level development of component-based distributed virtual reality systems," in *2013 IEEE 37th Annual Computer Software and Applications Conference*, 2013, pp. 696–705

[13] M. A. A. Gutiérrez, F. Vexo, and D. Thalmann, Eds., "Architecture of Virtual Reality Systems BT - Stepping into Virtual Reality," London: Springer London, 2008, pp. 107–116. doi: 10.1007/978-1-84800-117-6_5

[14] T. Molet *et al.*, "Anyone for Tennis? Presence." MIT, 1999.

[15] P. Pajusco, "Propagation channel models for mobile communication," *Comptes Rendus Phys.*, vol. 7, no. 7, pp. 703–714, 2006.

[16] J.-M. Conrat and P. Pajusco, "Clusterization of the propagation channel in urban macrocells at 2 GHz," in *The European Conference on Wireless Technology, 2005.*, 2005, pp. 39–42

[17] N. A. Abbasi *et al.*, "Double directional channel measurements for THz communications in an urban environment," in *ICC 2020-2020 IEEE international conference on communications (ICC)*, 2020, pp. 1–6.

[18] A. H. F. Raouf, S. J. Maeng, I. Guvenc, O. Ozdemir, and M. Sichitiu, "Spectrum Monitoring and Analysis in Urban and Rural Environments at Different Altitudes," *arXiv Prepr. arXiv2301.02380*, 2023

[19] A. A. M. Saleh and R. Valenzuela, "A statistical model for indoor multi-path propagation," *IEEE J. Sel. areas Commun.*, vol. 5, no. 2, pp. 128–137, 1987

[20] A. F. Molisch *et al.*, "IEEE 802.15. 4a channel model-final report," *IEEE P802*, vol. 15, no. 04, p. 662, 2004

[21] Y. Okumura, "Field strength and its variability in VHF and UHF land-mobile radio service," *Rev. Electr. Commun. Lab.*, vol. 16, pp. 825–873, 1968

[22] M. Hata, "Empirical formula for propagation loss in land mobile radio services," *IEEE Trans. Veh. Technol.*, vol. 29, no. 3, pp. 317–325, 1980

[23] M. Farhoud, A. El-Keyi, and A. Sultan, "Empirical correction of the Okumura-Hata model for the 900 MHz band in Egypt," in *2013 Third International Conference on Communications and Information Technology (ICCIT)*, 2013, pp. 386–390.

[24] A. V., S. K.M.S., M. N., and F. T., Propagation Models and Their Applications in Digital Television Broadcast Network Design and Implementation. IntechOpen, 2010. doi: 10.5772/8471

[25] J. Walfisch and H. L. Bertoni, "A theoretical model of UHF propagation in urban environments," *IEEE Trans. Antennas Propag.*, vol. 36, no. 12, pp. 1788–1796, 1988

[26] M. Dong *et al.*, "{PCC} vivace: Online-learning congestion control," in *15th {USENIX} Symposium on Networked Systems Design and Implementation ({NSDI} 18)*, 2018, pp. 343–356

[27] N. Jay, N. Rotman, B. Godfrey, M. Schapira, and A. Tamar, "A deep reinforcement learning perspective on internet congestion control," in *International Conference on Machine Learning*, 2019, pp. 3050–3059

[28] V. Arun and H. Balakrishnan, "Copa: Practical delay-based congestion control for the internet," in *15th {USENIX} Symposium on Networked Systems Design and Implementation ({NSDI} 18)*, 2018, pp. 329–342

[29] M. Dong, Q. Li, D. Zarchy, P. B. Godfrey, and M. Schapira, "{PCC}: Re-architecting congestion control for consistent high performance," in *12th {USENIX} Symposium on Networked Systems Design and Implementation ({NSDI} 15)*, 2015, pp. 395–408

[30] S. Ha, I. Rhee, and L. Xu, "CUBIC: a new TCP-friendly high-speed TCP variant," *ACM SIGOPS Oper. Syst. Rev.*, vol. 42, no. 5, pp. 64–74, 2008.

[31] N. Cardwell, Y. Cheng, C. S. Gunn, S. H. Yeganeh, and V. Jacobson, "Bbr: Congestion-based congestion control: Measuring bottleneck bandwidth and round-trip propagation time," *Queue*, vol. 14, no. 5, pp. 20–53, 2016

[32] K. Winstein and H. Balakrishnan, "Tcp ex machina: Computer-generated congestion control," *ACM SIGCOMM Comput. Commun. Rev.*, vol. 43, no. 4, pp. 123–134, 2013

[33] Hoffman HG, Chambers GT, Meyer WJ, et al. Virtual reality as an adjunctive non-pharmacologic analgesic for acute burn pain during medical procedures. Ann Behav Med 2011;41(2):183–191

[34] Garrett B, Taverner T, McDade P. Virtual Reality as an Adjunct Home Therapy in Chronic Pain Management: An Exploratory Study. JMIR Med Inform 2017;5(2):e11

[35] Ma M, Fallavollita P, Seelbach I, et al. Personalized augmented reality for anatomy education. Clin Anat 2016;29(4):446–453.

[36] Kyaw BM, Saxena N, Posadzki P, et al. Virtual Reality for Health Professions Education: Systematic Review and Meta-Analysis by the Digital Health Education Collaboration. J Med Internet Res 2019;21(1):e12959

[37] AlDossary S, Martin-Khan MG, Bradford NK, et al. A systematic review of the methodologies used to evaluate telemedicine service initiatives in hospital facilities. Int J Med Inform 2017; 97:171–194

[38] Wirth F, Tabone F, Azzopardi L, et al. Privacy and Security in Mobile Health: A Research Agenda. IEEE Comput 2018;51(6):22–30.

[39] Aucinas A, Evans J, Pezaros D. TCP Congestion Control for Virtual Reality Streaming. In: Proceedings of the 1st Workshop on Virtual Reality and Augmented Reality Network. VR/AR Network '17. New York, NY, USA: Association for Computing Machinery; 2017. p. 7–12.

[40] Van Velsen L, Wildevuur S, Flierman I, et al. Trust in telemedicine portals for rehabilitation care: an exploratory focus group study with patients and healthcare professionals. BMC Med Inform Decis Mak 2016; 16:11

[41] Deepak Singh, Anil Naik. Redefining Boundaries in Healthcare: An In-Depth Exploration of Telemedicine's and Wireless Technology’s Role in Emergency Care and Disaster Management. https://doi.org/10.14445/22312803/IJCTT-V71I6P102