Networking for the Metaverse

Benjamin Finley, Nitinder Mohan, and Peng Yuan Zhou

1 User Experienced Delay

In order to deliver the same level of experience as reality, it is vital in the metaverse to ensure the user's immersion by minimizing the motion to photon (MTP) latency.[1] Researchers have determined that MTP latency must be within the human perceptible limit for users to interact fluidly and immediately with holographic enhancements. During the AR registration process, for instance, significant latency frequently causes virtual objects to lag behind their intended position [2], which can create nausea and vertigo. As a result, reducing end-to-end latency is essential for the metaverse, particularly in situations where real-time data processing is required, such as real-time AR interaction with the physical world like AR surgeries [3–5], or real-time user interactions in the metaverse, such as multiplayer interactive exhibits in VR [6] or multiple players battling in Fortnite.

As previously stated, the metaverse frequently necessitates intense computing for mobile devices, which increases latency. To compensate for the restricted capacity of graphics and chipsets in mobile interfaces (AR glasses and VR headsets, etc.), offloading is frequently employed to alleviate the compute and memory strain at the

[1] MTP latency is the amount of time between the user's action and its corresponding consequence being reflected on the display screen, and is one of the most influential elements in the immersive experience [1].

B. Finley (✉)
University of Helsinki, Helsinki, Finland
e-mail: benjamin.finley@tutanota.com

N. Mohan
Delft University of Technology, Delft, The Netherlands

P. Y. Zhou
Aarhus University, Aarhus, Denmark

P. Hui et al. (eds.), *Handbook of the Metaverse*,
https://doi.org/10.1007/978-3-032-03296-6_6

expense of increased networking latency [7, 8]. A balanced trade-off is essential to make the offloading process transparent to the user experience in virtual environments. But it is difficult. For instance, creating a locally navigable viewport larger than the headset's field of view is required to compensate for network latency during offloading [9]. However, there is a conflict between the required viewport size and the networking latency: a longer delay necessitates a wider viewport and the streaming of more material, which results in even longer latency. Consequently, a solution that improves physical deployment may be more practical than pure resource orchestration.

Due to the fluctuating and unpredictable high latency [10–13], cloud offloading cannot always achieve the optimal balance and results in long-tail latency performance, which harms the user experience [14]. Recent cloud reachability measurements indicate that the current cloud distribution can provide network latency of less than 100 milliseconds [15–17]. However, only a handful of nations (24 out of 184) reliably reach the MTP level via wired networks, while only China (out of 184) satisfies the MTP requirement using wireless networks [1]. To ensure a seamless and immersive user experience within the metaverse, a complementary solution is required.

Edge computing, which computes, stores, and transmits data closer to end-users and their devices, can reduce user-experienced latency relative to cloud offloading [18]. Satyanarayanan et al. [11] identified in 2009 that deploying strong cloud-like infrastructure just one wireless hop away from mobile devices, i.e., cloudlet, may revolutionize the computing paradigm, as demonstrated by numerous subsequent works. Specifically, Chen et al. [19] examined the latency performance of edge computing by conducting practical tests on various applications. They demonstrated that LTE cloudlets could deliver significant benefits (60% less latency) over cloud offloading. Similarly, Ha et al. [20] discovered through measurements that edge computing can cut service latency by at least 80 ms on average compared to the cloud.

Researchers have proposed strategies to enhance the performance of metaverse applications by leveraging the latency advantage of edge computing. For example, EdgeXAR [21] is a mobile augmented reality framework that takes advantage of edge offloading to provide lightweight tracking with six degrees of freedom while hiding the offloading latency from the user. Jaguar challenges the limits of end-to-end latency in mobile augmented reality by exploiting hardware acceleration on edge clouds with GPUs [22]. EAVVE [12] provides a cooperative AR vehicular vision system assisted by edge servers to lower the total offloading latency and compensate for insufficient in-vehicle computational capability. Bartolomeo et al. [8] proposed scAtteR, a distributed stream processing-based AR architecture that decouples the different compute blocks of AR pipelines as different microservices for improved scalability and end-to-end performance. Likewise, similar strategies have been proposed for VR services. Lin et al. [23] turned the problem of energy-aware VR experience into a Markov decision process and utilized pervasive edge computing to realize an immersive wireless VR experience. Gupta et al. [24] combined scalable

360-degree content, VR user viewport modeling, mmWave connectivity, and edge computing to realize a low-latency 8 K 360-degree video mobile VR arcade streaming system. Elbamby et al. [25] suggested a proactive edge computing and mmWave communication system to enhance the performance of an interactive VR network game arcade that demands real-time rendering of HD video frames. As the resolution increases, edge computing will play an increasingly crucial role in reducing the latency of metaverse streaming at 16, 24 K, and even higher resolutions.

2 Multi-access Edge Computing

In the eyes of many industry insiders, edge computing's exceptional performance in lowering latency in virtual worlds has made it a crucial pillar in the building of the metaverse. Apple, for instance, employs a Mac with a VR headset attached to support 360-degree VR rendering [26]. Thanks to its powerful Qualcomm Snapdragon XR2 CPU, the Facebook Oculus Quest 2 can give VR experiences without the need for a PC. In contrast to a strong PC, the standalone VR experience suffers from lower frame rates and thus less detailed VR scenes. By offloading to an edge server (such as a personal computer), consumers can enjoy a more dynamic and immersive experience at higher frame rates without compromising detail. The Facebook-announced Oculus Air Link [27] in April 2021 enables Quest 2 to offload to the edge at up to 1200 Mbps via the home Wi-Fi network, offering a lag-free VR experience with enhanced mobility. However, these products are limited to interior areas with restricted user movement.

For consumers to experience a truly omnipresent metaverse, cellular networks must offer seamless performance for AR/VR users during their outdoor movements. So far, last-mile access has been the bottleneck for latency in LTE and 5G networks [28]. Multi-access edge computing (MEC) is anticipated to enhance the metaverse user experience by delivering standard and universal edge offloading services one hop away from cellular-connected user devices, such as AR glasses, as 5G and 6G evolve. It not only decreases the round-trip time (RTT) of packet delivery [29], but also enables near real-time orchestration for multi-user interactions [30]. MEC is required for outdoor metaverse services to comprehend the specific local context and coordinate close cooperation between adjacent users or devices. 5G MEC servers, for instance, may manage the AR content of adjacent users with a single hop of packet transfer and enable real-time user interaction for social AR applications such as Pokémon GO [31]. Figure 1 depicts an illustration of an MEC solution offered by ETSI [32]. Utilizing MEC to enhance the metaverse experience has garnered scholarly interest. To increase the QoE of wireless VR applications, Dai et al. [33] built a view synthesis-based 360-degree VR caching system over MEC-Cache servers in Cloud Radio Access Network (CRAN). Gu et al. [34] and Liu et al. [35] both employed sub-6 GHz links and mmWave links in conjunction with MEC resources to address the restricted resources on VR HMDs and the transmission rate bottle-

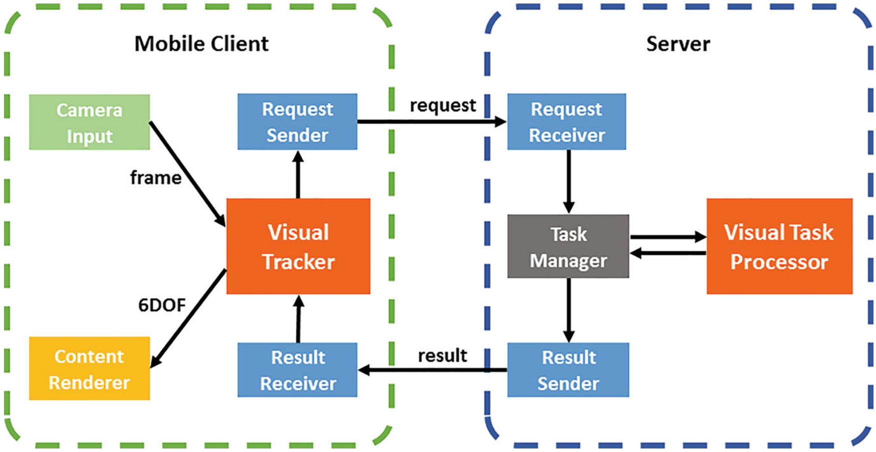

Fig. 1 The overview of EdgeXAR framework introduced by [21]

neck for standard VR and panoramic VR video (PVRV) delivery, respectively. Zhou et al. [29] explored a practical issue in scalable outdoor mobile AR scenarios, focusing on the handoff problem considering both computation and signal when users move across MECs.

In fact, metaverse companies have begun using MEC to enhance the user experience. DoubleMe, a leading volumetric capture company, announced a proof of concept project, Holoverse, in collaboration with Telefónica, Deutsche Telekom, TIM, and MobiledgeX, to test the optimal 5G Telco Edge Cloud network infrastructure for the seamless deployment of various services using the metaverse in August 2021 [36]. The renowned developer of "Ingress," "Pokémon GO," and "Harry Potter: Wizards Unite," Niantic anticipates creating a "Planet-Scale AR." It has partnered with global telecommunications operators, such as Deutsche Telekom, EE, Globe Telecom, Orange, SK Telecom, SoftBank Corp., TELUS, Verizon, and Telstra, to improve the performance of their AR services using MEC [37]. With the advancement of 5G and 6G technologies, last-mile latency will continue to decrease. Therefore, MEC promises to enhance its contribution to the universal metaverse experience.

Because the metaverse will capture more user data than ever, the privacy risk is among the major concerns [38, 39]. For example, Amazon, Apple, Google (Alphabet), Facebook, and Microsoft have long supported password-less authentication [40, 41], which validates identification with a fingerprint, face recognition, or PIN. The metaverse will certainly continue in this manner, maybe incorporating more biometrics such as audio and iris recognition [42, 43]. Previously, if a user forgot their password, the worst-case scenario was that they lost some data and had to create a new password to protect other data. Since biometrics are permanently connected with a user, if they are compromised (taken by an imposter), they cannot be canceled, and the user is in serious trouble [44, 45].

3 Multimodal Networking

A given metaverse can support multiple human sensing modalities (such as sight, sound, and touch) and thus can provide varying levels of immersion [46]. These multimodal metaverses henceforth also present different networking requirements for the different sensing modalities. For example, high-resolution visual information might require high throughput and medium latency, whereas haptic information might require lower throughput and low latency (see Fig. 2 and Table 5.3.3-1 in [47]). Thus networks should be able to satisfy these differing requirements simultaneously so that users can successfully integrate these modalities (within a temporal window of integration [46] often ranging from 10 to 30 ms [47]) and therefore experience coherent events (such as picking up an object while also looking at that object).

In the context of current 3GPP mobile networks, a simple solution is to simply leverage several different downlink network slices with different networking applications such as enhanced mobile broadband (eMBB) for visual information and ultra-reliable low latency (URLLC) for haptic information. However, multiplexing different applications on a single RAN can cause self-interference between the two slices [48]. In general, 3GPP prescribes two methods for slice multiplexing: puncture scheduling and orthogonal scheduling. In puncture scheduling, URLLC traffic is placed in short time gaps within the eMBB traffic (thus preempting such traffic) so as to ensure URLLC reliability. The main downside is that the eMBB traffic reliability can suffer. Alternatively, in orthogonal scheduling resources are reserved in advance for URLLC traffic, however, given bursty URLLC traffic many of these reservations are empty thus resulting in inefficiencies. Therefore, research into alternative multiplexing schemes (including some that take multimodal VR into account) is ongoing. We discuss a few examples of such research, refer to [49] for a comprehensive survey (in Sect. 5) of different approaches.

As an early example, [48] study such a multiplexing scenario and compare orthogonal multiple access (OMA) to non-orthogonal multiple access (NOMA) with suc-

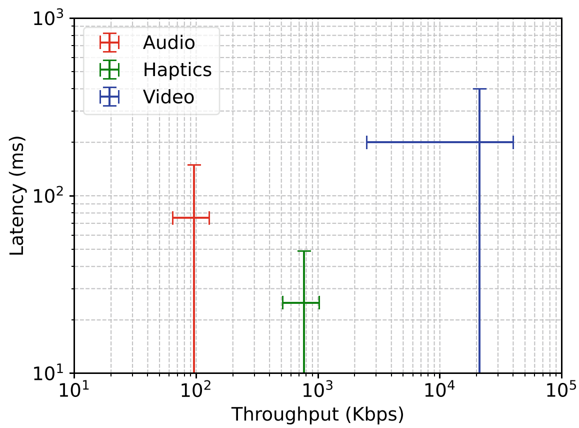

Fig. 2 Typical network requirements in terms of throughput and latency for different modalities to provide satisfactory quality of experience [47]. The throughputs represent typical ranges for the modalities while the latencies represent the maximum latencies and extend from these maximums down to zero

cessive interference cancellation (SIC) in terms of multiplexing the radio resource in multimodal VR. They use a stochastic geometric network context to perform both analytical analysis and a validating simulation. The results show that NOMA (with the proposed SIC) outperforms OMA for higher resolutions (both visual and haptic) in terms of a well-known metric of human perception (the just-noticeable difference metric). The results are relevant because current networks primarily use OMA (such as the orthogonal frequency-division multiplexing of downlink LTE and 5G NR).

More recently, [50] proposed a deep reinforcement learning (DRL) method to improve upon puncture scheduling by also taking into account the eMBB reliability requirements rather than just eMBB throughput maximization. eMBB reliability (in the study defined in terms of throughput variance) is important in the metaverse context given that visual information still requires moderate reliability. The proposed method leverages traditional optimization methods for a majority of the scheduling yet also uses a DRL method to help ensure the URLLC requirements are met. The DRL method is needed because traditional optimization methods often require significant simplification of complex problems to allow efficient solving (for example, relaxing of constraints). Therefore, this combination method merges the efficiency benefits of traditional methods with the complexity handling benefits of DRL methods.

While [51] suggests a genetic algorithm method that partitions the spectrum between eMBB and URLLC. They argue that a genetic algorithm method is more computationally efficient than a DRL method yet can still handle complex (often non-convex) optimization problems. They show their approach is superior to a basic Q-Learning-based DRL approach in terms of managing the trade-off between performance and computational efficiency.

Alternatively, future mobile networks look to provide high-rate and high-reliability low-latency communications (HRLLC) as a new application over 6G networks and potentially new bands such as those with terahertz wavelengths. Thus a single application could support all metaverse modalities.

Specifically, [52] study the potential for HRLLC over THz network through an analytical model and simulation. Quantitatively, the analysis shows that, such a network can provide the high reliability (>99.999% with latency <20 ms) and throughput (>18 Gbps) needed for future VR and metaverse applications. However, such a network requires very high BS density and potentially new networking techniques such as ultra-massive MIMO and intelligent reflecting environmental surfaces (IRSs) [53] to ensure enough LoS or quasi-LoS for such reliability and throughput. IRSs are low-power software-tunable RF scattering surfaces that can be placed in the environment to, for example, constructively reflect (through a phase shift) a BS signal thus reducing loss in an NLoS situation.

In addition to 3GPP-based mobile networking solutions, research on multimodal transport layer solutions and protocols is ongoing. These solutions are typically built on existing transport protocols such as TCP and QUIC.

Reference [54] proposes an addition to the QUIC protocol that adjusts each stream's priority based on the stream's throughput requirement. Specifically, priority is based on weighted fairness criteria that is inversely proportional to the throughput requirement. Thus streams with the lowest throughput requirements (such as audio and haptic streams in the metaverse) are prioritized at a relatively small cost to the high-throughput streams (such as visual streams). This prioritization is roughly similar to that which would be implied by the QoE requirements discussed at the beginning of the section. The major benefits of such a strategy are the general nature (for example, the strategy is also applicable in non-metaverse scenarios such as video calls) and simplicity. The strategy can be easily applied to QUIC as a code patch with minimal changes.

So far the solutions discussed either deal with only a single domain or autonomous system (like a single mobile network) or in the case of transport protocols automatically assume the underlying path (often across multiple domains) actually has the baseline capability to support the requirements of metaverse applications. This might not be the case for metaverse applications that need to traverse parts of the public Internet where there may be bottleneck domains and the QoS model is generally the best effort. Therefore, to address this issue, [55], for example, proposes a cross-domain QoS concept that introduces and leverages a metaverse network broker (similar to a central controller in a software-defined network) to coordinate with the network providers (domains) along a path, provision and monitor the connections, and reward network providers for achieving the agreed upon QoS levels. The concept also adds domain-specific information in the IPv6 routing header to allow tracking performance for each domain and uses an extended version of DNS to allow clients to specify their metaverse requirements and get back a suitable domain path (i.e., a sequence of domains).

Finally, as mentioned, these modalities must be kept in sync within small thresholds and potentially between multiple users. Such synchronization between many modalities and between multiple users does not yet exist in the 5G standard and remains an open issue [47].

4 Semantic and Goal/Deadline-Aware Networking

Semantic- or goal-aware networking is an extension of traditional networking that expands the task of networking from simply and accurately transferring arbitrary bits or symbols to transferring only the information relevant for the receiver task (and thus leveraging semantic information) [56]. For example, the goal of a specific networking activity might be to share information between devices about nearby objects that each device detects (via a camera), the receiver devices would then leverage the information to build area models. In this context, a traditional networking task might be to transfer the bits of high-resolution images. In contrast, a semantic- or goal-aware task might be to process (through an AI model) such images before transferring and extracting the object identification information and transferring just object outlines, location information, and labels. This significantly reduces the volume of data to transfer (compared to traditional general image compression methods

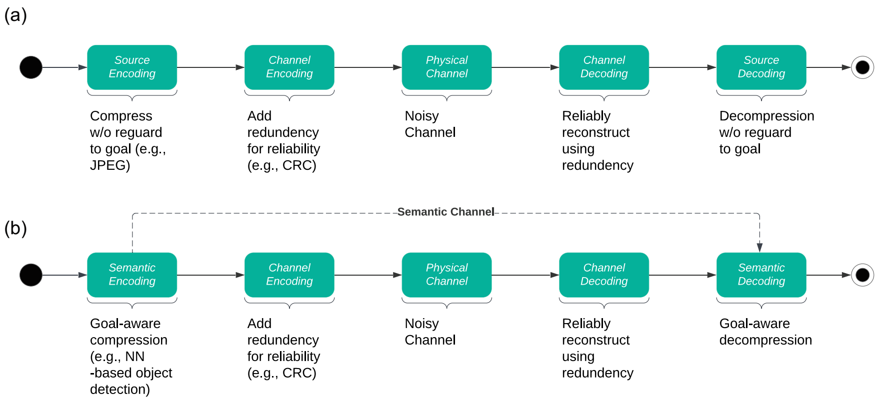

Fig. 3 Comparison of traditional networking steps **a** versus semantic networking steps **b**

such as JPG2000) and can be seen as a form of task-specific compression from the signal processing area (Fig. 3).

In terms of the metaverse, semantic networking has significant potential since the metaverse includes a wide variety of data types (such as rendered frames, video, and text) and tasks that are well established in semantic communication research and frameworks. However, given the diversity of data types and tasks within even a single given metaverse, having separate semantic communication frameworks (including semantic encoding and decoding steps) for each task is problematic and adds significant complexity. Thus general and adaptable frameworks might be especially important for the metaverse. We briefly describe a few frameworks from recent research.

Reference [56] proposes a semantic networking framework that uses directed bipartite graphs with nodes from one side representing entities (for example, different objects from an image) and the other side representing predicates or relationships (such as overlapping or similarity). The edges thus indicate relationships between specific entities. The framework is general and flexible enough that, as the authors illustrate, many different types of tasks (object detection, video captioning, scene graph generation, and speech recognition) can be straightforwardly mapped to the framework. Additionally, the use of a bipartite graph structure allows more efficient processing than other structures such as normal graphs.

While [57] introduces a semantic framework for image transmission that leverages deep learning to allow rapid adaptation to varying receiver tasks while the sender can be essentially task-unaware and thus more general and reusable. Specifically, the framework leverages DNNs for semantic encoding on the sender and decoding on the receiver. The receiver uses a loss function that takes into account both the observable (traditional networking) performance and task performance to optimize both the encoding and decoding networks. Additionally, a transfer learning approach (using a GAN) helps with adapting the encoder and decoder networks to varying image-based tasks thus lowering retraining costs.

5 Multipath Networking

Multipath networking refers to aggregating the available capacity of several different physical networks (for example, with multipath TCP (MPTCP) over Wi-Fi and 5G) or several different base stations of the same network (for example, with dual-connectivity over a 3GPP network). Moreover, Multipath TCP (MPTCP) is widely deployed in the Internet as it was upstreamed in the default Linux kernel since May 2020 [58]. This allows end-hosts to switch their existing TCP connections to remote servers to MPTCP, which can help improve the capacity available for bandwidth-intensive metaverse applications.

However, for cases with several different physical networks, the often varying network conditions (which can be a natural result of the different underlying technologies) can result in issues given the strict requirements of many metaverse applications. Specifically, the congestion control algorithms of the network protocols (for example, MPTCP and MP-QUIC) can struggle to adapt to rapid order of magnitude changes in network performance (from fading and congestion conditions) and with deciding how much traffic to send over each network (while remaining fair to other flows) [59]. This can result in sub-optimal aggregate throughput and head-of-line (HoL) blocking, impacting latency [60].

To help with this challenge, researchers are looking at novel heuristic, AI, and hybrid (heuristic and AI fusion) algorithms [61], with some work focusing on the metaverse as an application area [62]. In terms of AI algorithms, most of these leverage deep reinforcement learning which uses a deep neural network model to learn the best responses to the changing network conditions [61]. Such algorithms show improvements over heuristic approaches in many studies; however, the downsides include the need for periodic or continuous training to ensure that the model can perform well even in conditions not in the original training dataset and high computational requirements from both training and inference than heuristic algorithms.

In terms of metaverse applications, [62] propose a novel multipath transport layer protocol that optimally aggregates the available capacity of the networks even under varying network conditions and prevents HoL blocking. Specifically, the protocol uses two single-path TCP connections (one over each network) and for a given block of data it starts sending that data from front to back over one connection and from back to front over the other connection. Figure 4 illustrates this process. Thus even if one connection becomes rapidly unusable the other connection does not need to wait and can send all the remaining data in that block. This prevents many cases of HoL blocking. They also propose an algorithm for minimizing viewport distortion in streaming 360° video (similar to a metaverse) over multiple paths. The algorithm decides the compression rates of different viewport tiles by leveraging the probabilities that the user will view each tile and while also respecting the throughput constraints of each network connection. They test the protocol empirically using real-world mobile 360° video scenarios.

Similarly, [63] also proposes a similar algorithm for deciding compression rates of viewport tiles for transmission over multiple network paths. However, they use the

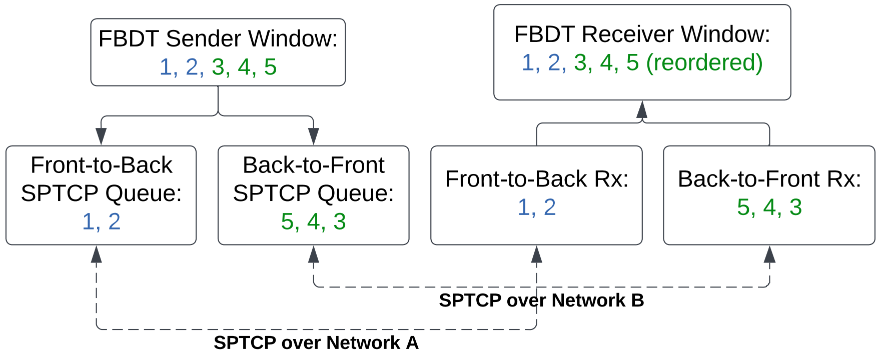

Fig. 4 Diagram of the Forward and Backward Data Transmission protocol (FBDT) sending five packets by leveraging single-path TCP (SPTCP) connections over networks A and B. In this case, the Back-to-Front connection reaches packet three first and thus sends the packet rather than the Front-to-Back connection. Also note that extra reordering work is required for the Back-to-Front packets on the receiver side

established MPTCP protocol rather than proposing a novel transport protocol. Note that multipath solutions are often similar in spirit to multimodal solutions though typically a difference remains in terms of assuming several independently managed networks versus a single network.

6 Ubiquitous Network Coverage

As previously mentioned, metaverse applications often require high-throughput networking, such as with 5/6G mobile networks that leverage the significant bandwidth available in higher frequency bands. However, these frequencies often require dense networks with many base stations due to poor propagation characteristics. Therefore, supporting metaverse applications in some suburban and rural areas where dense networks are not economically feasible is difficult. Additionally, metaverses that leverage real-world data (such as those with digital twins) need to collect significant data from these areas, which is difficult without such mobile networks. To help with these issues, researchers suggest ubiquitous network coverage and data collection through the use of Unmanned Aerial Vehicles (UAVs) as a possible solution.

For example, [64] proposes a dual-agent reinforcement learning method for selecting both the communication channel and flight trajectory of the UAV moving past a set of clients that have data for collection for metaverse digital twins. The reward function takes into account the total time taken to complete the full flight along with penalties for not collecting the full data from clients. Such work builds on significant prior studies on the optimization of UAVs in the communication and networking context [65] often work with non-metaverse applications. Naturally, metaverse applications can pose greater challenges than other applications due, for example, to the potential for diversity in terms of requirements. In a digital twin context, certain properties are much more tolerant of delays in updating.

7 Networking for Localizing and Positioning

Metaverse applications including those leveraging digital twins often rely on highly accurate real-world positioning information; such applications and services are known, for example, as localized mobile metaverse services by 3GPP [66]. GPS information from clients with GPS sensors can help in certain situations such as outdoors; however, in indoor or more challenging situations, network-assisted localization can be a potential solution.

State-of-the-art network localization algorithms can be split into conventional methods and learning-based methods (including machine learning and neural network solutions) [67]. Conventional methods include, for example, traditional triangulation, fingerprinting, Kalman filtering, and compressive sensing. While learning methods span the range of ML and AI approaches, for example, k-nearest neighbors, support vector machines, convolutional NN, federated learning, and transfer learning. These methods typically offer different tradeoffs in terms of key indicators such as location accuracy, precision, latency, coverage, and stability (variation in accuracy over time).

Beyond basic network-assisted approaches, in a metaverse context, cooperative localization (CL) is potentially an important topic given the social nature of many metaverses. In CL users rely on signals to both fixed base stations and to other users to achieve localization. Naturally, privacy is a concern in CL because location data is sensitive and other users are generally untrustworthy. However, in the metaverse, user location is potentially already being shared with nearby users (for interaction purposes), thus the privacy issues in this context are of less concern. Overall CL improves localization precision and coverage at the expense of more computation and energy usage at the user device [68].

State-of-the-art CL work studies situations where CL localization can occur even in cases with a single or no base stations but with one or more of the aforementioned passive IRSs [69]. Given the proposed ubiquity of the metaverse, low-cost solutions such as IRSs are attractive to keep deployment costs down. In contrast, normal 5G requires at least two antenna-array base stations for localization. As an example of such a low-cost IRS, [70] presents an IRS design consisting of an interconnected 2D array of cells with each cell containing an antenna and RF switch to allow for passive 3D beamforming. This beamforming essentially redirects the incoming RF signal to a specified direction (through phase shifts generating constructive interference) without the need for active RF components. An attached microcontroller allows rapid reprogramming of the array to change the phase shifting and thus the beam direction. The work estimates a cost per cell of two US dollars or less when manufacturing at large scales, whereas a 4G or 5G small cell can cost between 4000 and 20,000 dollars [71].

Additionally, the sensors common to many metaverse headsets such as high-resolution RGB cameras and LiDAR can help in localization through the use of spatial mapping. In spatial mapping such sensor data is compared to data from a known map of spatial information (such as room contents, landmarks, and dimensions) previously collected by other users, a mapping service like Google Maps, or even city or property owners. In the network context, future 3GPP standards suggest that this spatial mapping could be a 5G network service available to metaverse applications [47]. Storing and updating such spatial maps in network edge servers close to the physical locations would naturally provide several advantages such as reducing latency and backhaul load (compared to cloud storage by a third-party metaverse provider).

8 Content Delivery Networks

Besides the commonly addressed stricter requirement on low latency, metaverse content delivery also faces unique challenges such as context-dependent multimedia retrieval and trust issues across multiple metaverse platforms.

Caching stands at the core of CDN and may require customized changes for metaverse content delivery. First, it might be more meaningful to cache background scenes for XR instead of caching bulk videos which is currently employed, to improve the cache hit ratio. Because the clients may have different FoVs even when watching the same scene, but with a largely overlapped background. Second, varied forms of cached contents, e.g., chunks or tiles, have also been addressed by many works [72–74], to allow finer granularity of transmission tailored to the user's FoV movement. Third, edge servers, especially MECs have been considered by some works as a key part of pushing popular metaverse content near the users [75] to meet low-latency requirements. A more progressive approach to edge is opportunistic content delivery that uses a nomadic user's device as the data source to distribute demanded XR content among nearby users [76].

Request routing and content retrieval methods are also critical. A few studies have proposed requesting routing methods to optimize XR delivery. Reference [77] proposed a scalable request routing approach to optimize VR video delivery. They proposed to first direct the streaming requests to the best proximal cluster using DNS mappings, then process the requests in batch to optimize their allocation in an online fashion using a fast linear programming-based heuristic. Reference [78] proposed a viewport request routing problem customized for the 5G network, jointly considering the rendering-aware tile cache placement. Reference [79] applied segment routing to enable fast routing through a service chain, incorporating service function chaining and microservices, deployed at different infrastructure levels (i.e., edge, fog, cloud) for VR content delivery.

A few works have been proposed surrounding Experience Delivery Network (XDN),[2] a newer version of CDN specifically targeting enabling immersive experience and XR services via 5G and Edge Computing. For example, [80] identified the key cloud-native building blocks for edge cloud XDN architecture and presented an end-to-end 360 immersive media streaming solution leveraging cloud-native modular microservices. Reference [81] proposed a technique based on statistical models to generate and deploy virtual users to simulate a real-world environment for designing an XDN that can provide XR services to a large number of people. Such simulation can help design the XDN via network design and dynamic server/data allocation to meet QoS requirements.

Next-gen networking has also been considered. Information-Centric Network (ICN) was proposed as an enhancement of the current host-centric IP network, using identified information or content as the focal point for packet routing. Some researchers believe ICN is a better fit for metaverse content delivery than the current IP network. For example, Burke [82] proposed to use Named Data Networking (NDN), a branch of ICN that uses application-defined names for data forwarding, to deliver AR content via semantic-enabling web interfaces. The author claims that NDN provides numerous benefits over IP networks in terms of AR content delivery, including better trust management with heterogeneous content providers powered by schematized trust [83], and allowing multiple content providers to efficiently access the user's context through its intrinsic multicast and caching support.

9 Standards and Interoperability

Several diverse groups across the ICT domain are pursuing standardization of metaverse topics. Overarchingly, the Metaverse Standards Forum is an umbrella organization founded in 2023 that aims to coordinate different standards organizations, companies, and institutions to help create an inclusive and open metaverse. The forum does not create standards itself but instead looks to generate outputs such as use cases, pilots, testbeds, tooling, best practices, and guidelines to help speed up standardization.

As an example, the Real/Virtual World Integration domain group of the Metaverse Standards Forum is developing a use case in which both a ride requester and driver (when within visual distance) have (spatially anchored) real-time AR signs that help coordinate a mutual safe pick-up location [84]. Figure 5 illustrates a potential example of such a sign from the ride requester viewpoint. Such use cases help drive potential requirements that standard organizations can leverage when actually developing standards.

In terms of standards organizations, at least the following organizations have established working groups (or equivalents) on the metaverse: The Institute of Electrical and Electronics Engineers (IEEE), International Organization for Standardiza-

[2] https://www.red5.net/blog/introducing-xdn-experience-delivery-network/.

Fig. 5 A potential example of a (spatially anchored) real-time AR sign for a ride pick-up location (from ride requester viewpoint). The Real/Virtual World Integration domain group of the Metaverse Standards Forum is developing a use case that leverages real-time AR signs for ride hailing [84]

tion (ISO), World Wide Web Consortium (W3C), International Telecommunication Union (ITU), and 3rd Generation Partnership Project (3GPP).

In the network context, there are several standardization efforts focusing at different network layers and network locations.

For mobile networks, 3GPP has released a technical report that surveys the network requirements for different metaverse use cases over 5G and briefly describes new functionality that 5G networks would need to fulfill these requirements [47]. 3GPP is now leveraging this report to develop a technical specification for mobile metaverse services [66] that will eventually become a published standard to extend 5G.

The technical report use cases include a mixture more traditional communication cases in the metaverse context and metaverse specific cases [47]. For context a few of these are listed below:

- 5G-enabled traffic flow simulation and situational awareness.
- Collaborative and concurrent engineering.
- Metaverse-based tele-operated driving.
- Movie streaming from metaverse server to the rendering device.
- Avatar information streaming between remote UEs.
- Interactive data exchange: voice and text between remote UEs.

The new functionality to support such use cases is categorized as in the list below with each category having roughly 5–10 functionalities. Interestingly, a majority of the new functionality is not related to headline network performance characteristics (such as low latency and high throughput) but instead more mundane issues such as exposure of APIs to help with digital representation of users and charging.

- Localized mobile metaverse service functionality.
- Digital representation of users and avatar functionality.
- Efficiency, exposure, and coordination of mobile metaverse services.
- Security and privacy aspects of mobile metaverse services.
- Digital asset management.
- Charging requirements for mobile metaverse services.

While ITU has also released several in-progress technical reports and technical specifications that cover topics such as metaverse applications, architecture, interoperability, security, economics, and accessibility.[3]

In terms of the web, W3C has released a draft recommendation of standard APIs (known as WebXR) for AR and VR devices to interface with web applications [85]. Specifically, web browsers can implement these APIs and then negotiate between metaverse web applications and user's metaverse devices (like VR goggles). The main conceptual object of WebXR APIs is the XRSession which represents an in-progress XR usage session and allows configuration and control of the session. For example, the application uses the session object to query the device position and orientation.

Overall, metaverse standardization is still in an early stage of development with many organizations only releasing initial artifacts in mid-2023. For reference, historically 3GPP has taken on average 692 days to develop a standard [86]. However, the standard deviation of development time is 649 days, thus illustrating high variability. Luckily many efforts (like W3C) are building on existing XR and VR results and thus may not take as long as other standardization processes. For example, Blink (the browser engine of both Google Chrome and Microsoft Edge) already passes roughly half of the 905 WebXR implementation tests [87] and supports the entire WebXR core.

References

1. N. Mohan, L. Corneo, A. Zavodovski, S. Bayhan, W. Wong, and J. Kangasharju, "Pruning edge research with latency shears," in *Proceedings of the 19th ACM Workshop on Hot Topics in Networks*, ser. HotNets '20. New York, NY, USA: Association for Computing Machinery, 2020, p. 182–189. [Online]. Available: https://doi.org/10.1145/3422604.3425943
2. R. L. Holloway, "Registration error analysis for augmented reality," *Presence: Teleoperators & Virtual Environments*, vol. 6, no. 4, pp. 413–432, 1997.
3. H. Fuchs, M. A. Livingston, R. Raskar, K. Keller, J. R. Crawford, P. Rademacher, S. H. Drake, A. A. Meyer *et al.*, "Augmented reality visualization for laparoscopic surgery," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 1998, pp. 934–943.
4. L. Soler, S. Nicolau, J. Schmid, C. Koehl, J. Marescaux, X. Pennec, and N. Ayache, "Virtual reality and augmented reality in digestive surgery," in *Third IEEE and ACM International Symposium on Mixed and Augmented Reality*. IEEE, 2004, pp. 278–279.

[3] https://www.itu.int/en/ITU-T/focusgroups/mv/Pages/deliverables.aspx.

5. P. Rhienmora, K. Gajananan, P. Haddawy, M. N. Dailey, and S. Suebnukarn, "Augmented reality haptics system for dental surgical skills training," in *Proceedings of the 17th ACM Symposium on Virtual Reality Software and Technology*, 2010, pp. 97–98.
6. L. Li and J. Zhou, "Virtual reality technology based developmental designs of multiplayer-interaction-supporting exhibits of science museums: taking the exhibit of "virtual experience on an aircraft carrier" in china science and technology museum as an example," in *Proceedings of the 15th ACM SIGGRAPH Conference on Virtual-Reality Continuum and Its Applications in Industry-Volume 1*, 2016, pp. 409–412.
7. T. Braud, Z. Pengyuan, J. Kangasharju, and H. Pan, "Multipath computation offloading for mobile augmented reality," in *2020 IEEE International Conference on Pervasive Computing and Communications (PerCom)*. IEEE, 2020, pp. 1–10.
8. G. Bartolomeo, J. Cao, X. Su, and N. Mohan, "Characterizing distributed mobile augmented reality applications at the edge," in *Companion of the 19th International Conference on Emerging Networking EXperiments and Technologies*, ser. CoNEXT 2023. New York, NY, USA: Association for Computing Machinery, 2023, p. 9–18. [Online]. Available: https://doi.org/10.1145/3624354.3630584
9. A. Yaqoob and G.-M. Muntean, "A combined field-of-view prediction-assisted viewport adaptive delivery scheme for 360° videos," *IEEE Transactions on Broadcasting*, vol. 67, no. 3, pp. 746–760, 2021.
10. A. Li, X. Yang, S. Kandula, and M. Zhang, "Cloudcmp: Comparing public cloud providers," in *Proceedings of the 10th ACM SIGCOMM Conference on Internet Measurement*, ser. IMC '10. New York, NY, USA: Association for Computing Machinery, 2010, p. 1–14.
11. M. Satyanarayanan, P. Bahl, R. Caceres, and N. Davies, "The case for vm-based cloudlets in mobile computing," *IEEE Pervasive Computing*, vol. 8, no. 4, pp. 14–23, 2009.
12. P. Zhou, T. Braud, A. Zavodovski, Z. Liu, X. Chen, P. Hui, and J. Kangasharju, "Edge-facilitated augmented vision in vehicle-to-everything networks," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 10, pp. 12 187–12 201, 2020.
13. T. Braud, F. H. Bijarbooneh, D. Chatzopoulos, and P. Hui, "Future networking challenges: The case of mobile augmented reality," in *2017 IEEE 37th International Conference on Distributed Computing Systems (ICDCS)*. pages 1796-1807, June 2017.
14. J. Dean and L. A. Barroso, "The tail at scale," *Communications of the ACM*, vol. 56, pp. 74–80, 2013. [Online]. Available: http://cacm.acm.org/magazines/2013/2/160173-the-tail-at-scale/fulltext
15. L. Corneo, M. Eder, N. Mohan, A. Zavodovski, S. Bayhan, W. Wong, P. Gunningberg, J. Kangasharju, and J. Ott, "Surrounded by the clouds: A comprehensive cloud reachability study," in *Proceedings of the Web Conference 2021*, ser. WWW '21. New York, NY, USA: Association for Computing Machinery, 2021, p. 295–304. [Online]. Available: https://doi.org/10.1145/3442381.3449854
16. T. K. Dang, N. Mohan, L. Corneo, A. Zavodovski, J. Ott, and J. Kangasharju, "Cloudy with a chance of short rtts: Analyzing cloud connectivity in the internet," in *Proceedings of the 21st ACM Internet Measurement Conference*, ser. IMC '21. New York, NY, USA: Association for Computing Machinery, 2021, p. 62–79. [Online]. Available: https://doi.org/10.1145/3487552.3487854
17. L. Corneo, N. Mohan, A. Zavodovski, W. Wong, C. Rohner, P. Gunningberg, and J. Kangasharju, "(how much) can edge computing change network latency?" in *2021 IFIP Networking Conference (IFIP Networking)*, 2021, pp. 1–9.
18. S. Shi, V. Gupta, M. Hwang, and R. Jana, "Mobile vr on edge cloud: a latency-driven design," in *Proceedings of the 10th ACM Multimedia Systems Conference*, 2019, pp. 222–231.
19. Z. Chen, W. Hu, J. Wang, S. Zhao, B. Amos, G. Wu, K. Ha, K. Elgazzar, P. Pillai, R. Klatzky, D. Siewiorek, and M. Satyanarayanan, "An empirical study of latency in an emerging class of edge computing applications for wearable cognitive assistance," in *Proceedings of the Second ACM/IEEE Symposium on Edge Computing*, ser. SEC '17. New York, NY, USA: Association for Computing Machinery, 2017. [Online]. Available: https://doi.org/10.1145/3132211.3134458

20. K. Ha, Z. Chen, W. Hu, W. Richter, P. Pillai, and M. Satyanarayanan, "Towards wearable cognitive assistance," in *Proceedings of the 12th annual international conference on Mobile systems, applications, and services*, 2014, pp. 68–81.
21. W. Zhang, S. Lin, F. Bijarbooneh, H.-F. Cheng, T. Braud, P. Zhou, L.-H. Lee, and P. Hui, "Edgexar: A 6-dof camera multi-target interactionframework for mar with user-friendly latencycompensation using edge computing," in *Proceedings of the ACM on HCI (Engineering Interactive Computing Systems)*, 2022.
22. W. Zhang, B. Han, and P. Hui, "Jaguar: Low latency mobile augmented reality with flexible tracking," in *Proceedings of the 26th ACM international conference on Multimedia*, 2018, pp. 355–363.
23. P. Lin, Q. Song, D. Wang, R. Yu, L. Guo, and V. Leung, "Resource management for pervasive edge computing-assisted wireless vr streaming in industrial internet of things," *IEEE Transactions on Industrial Informatics*, 2021.
24. S. Gupta, J. Chakareski, and P. Popovski, "Millimeter wave meets edge computing for mobile vr with high-fidelity 8k scalable 360° video," in *2019 IEEE 21st International Workshop on Multimedia Signal Processing (MMSP)*. IEEE, 2019, pp. 1–6.
25. M. S. Elbamby, C. Perfecto, M. Bennis, and K. Doppler, "Edge computing meets millimeter-wave enabled vr: Paving the way to cutting the cord," in *2018 IEEE Wireless Communications and Networking Conference (WCNC)*. IEEE, 2018, pp. 1–6.
26. Apple, "View 360° video in a vr headset in motion," June 2021. [Online]. Available: https://support.apple.com/guide/motion/motncec93c28/mac
27. Facebook, "Introducing oculus air link, a wireless way to play pc vr games on oculus quest 2, plus infinite office updates, support for 120 hz on quest 2, and more." April 2021. [Online]. Available: https://www.oculus.com/blog/introducing-oculus-air-link-a-wireless-way-to-play-pc-vr-games-on-oculus-quest-2-plus-infinite-office-updates-support-for-120-hz-on-quest-2-and-more/
28. I. Hadžić, Y. Abe, and H. C. Woithe, "Edge computing in the epc: A reality check," in *Proceedings of the Second ACM/IEEE Symposium on Edge Computing*, 2017, pp. 1–10.
29. P. Zhou, S. Fu, B. Finley, X. Li, S. Tarkoma, J. Kangasharju, M. Ammar, and P. Hui, "5g mec computation handoff for mobile augmented reality," *arXiv preprint* arXiv:2101.00256, 2021.
30. P. Ren, X. Qiao, Y. Huang, L. Liu, C. Pu, S. Dustdar, and J.-L. Chen, "Edge ar x5: An edge-assisted multi-user collaborative framework for mobile web augmented reality in 5g and beyond," *IEEE Transactions on Cloud Computing*, 2020.
31. H.-Y. Chen, R.-T. Hsu, Y.-C. Chen, W.-C. Hsu, and P. Huang, "Ar game traffic characterization: a case of pokémon go in a flash crowd event," in *Proceedings of the 19th Annual International Conference on Mobile Systems, Applications, and Services*, 2021, pp. 493–494.
32. Y. C. Hu, M. Patel, D. Sabella, N. Sprecher, and V. Young, "Mobile edge computing–a key technology towards 5g," *ETSI white paper*, vol. 11, no. 11, pp. 1–16, 2015.
33. J. Dai, Z. Zhang, S. Mao, and D. Liu, "A view synthesis-based 360° vr caching system over mec-enabled c-ran," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 30, no. 10, pp. 3843–3855, 2019.
34. Z. Gu, H. Lu, P. Hong, and Y. Zhang, "Reliability enhancement for vr delivery in mobile-edge empowered dual-connectivity sub-6 ghz and mmwave hetnets," *IEEE Transactions on Wireless Communications*, 2021.
35. Y. Liu, J. Liu, A. Argyriou, and S. Ci, "Mec-assisted panoramic vr video streaming over millimeter wave mobile networks," *IEEE Transactions on Multimedia*, vol. 21, no. 5, pp. 1302–1316, 2018.
36. Yahoo, "Real-world metaverse 'twinworld' selected as 5g telco edge cloud testbed for 3 global mobile carriers," August 2021. [Online]. Available: https://currently.att.yahoo.com/att/real-world-metaverse-twinworld-selected-161500696.html
37. Niantic, "Niantic planet-scale ar alliance accelerates social ar future in codename: Urban legends," March 2021. [Online]. Available: https://nianticlabs.com/blog/niantic-planet-scale-ar-5g-urban-legends/

38. R. Leenes, "Privacy in the metaverse," in *IFIP International Summer School on the Future of Identity in the Information Society*. Springer, 2007, pp. 95–112.
39. B. Falchuk, S. Loeb, and R. Neff, "The social metaverse: Battle for privacy," *IEEE Technology and Society Magazine*, vol. 37, no. 2, pp. 52–61, 2018.
40. F. Alqubaisi, A. S. Wazan, L. Ahmad, and D. W. Chadwick, "Should we rush to implement password-less single factor fido2 based authentication?" in *2020 12th Annual Undergraduate Research Conference on Applied Computing (URC)*. IEEE, 2020, pp. 1–6.
41. M. J. Haber, "Passwordless authentication," in *Privileged Attack Vectors*. Springer, 2020, pp. 87–98.
42. J. Lodge, "Nameless and faceless: The role of biometrics in realising quantum (in) security and (un) accountability," in *Security and Privacy in Biometrics*. Springer, 2013, pp. 311–337.
43. G. Boddington, "The internet of bodies–alive, connected and collective: the virtual physical future of our bodies and our senses," *Ai & Society*, pp. 1–17, 2021.
44. N. Ratha, J. Connell, R. M. Bolle, and S. Chikkerur, "Cancelable biometrics: A case study in fingerprints," in *18th International Conference on Pattern Recognition (ICPR'06)*, vol. 4. IEEE, 2006, pp. 370–373.
45. O. Ouda, N. Tsumura, and T. Nakaguchi, "Bioencoding: A reliable tokenless cancelable biometrics scheme for protecting iriscodes," *IEICE TRANSACTIONS on Information and Systems*, vol. 93, no. 7, pp. 1878–1888, 2010.
46. D. Martin, S. Malpica, D. Gutierrez, B. Masia, and A. Serrano, "Multimodality in vr: A survey," *ACM Comput. Surv.*, vol. 54, no. 10s, sep 2022. [Online]. Available: https://doi.org/10.1145/3508361
47. 3GPP, "Feasibility Study on Localized Mobile Metaverse Services," 3rd Generation Partnership Project (3GPP), Technical Specification (TS) 22.856, 09 2023, version 19.1.0. [Online]. Available: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=4046
48. J. Park and M. Bennis, "Urllc-embb slicing to support vr multimodal perceptions over wireless cellular systems," in *2018 IEEE Global Communications Conference (GLOBECOM)*, 2018, pp. 1–7.
49. M. E. Haque, F. Tariq, M. R. Khandaker, K.-K. Wong, and Y. Zhang, "A survey of scheduling in 5g urllc and outlook for emerging 6g systems," *IEEE Access*, 2023.
50. M. Alsenwi, N. H. Tran, M. Bennis, S. R. Pandey, A. K. Bairagi, and C. S. Hong, "Intelligent resource slicing for embb and urllc coexistence in 5g and beyond: A deep reinforcement learning based approach," *IEEE Transactions on Wireless Communications*, vol. 20, no. 7, pp. 4585–4600, 2021.
51. H. Peng, L.-C. Wang, and Z. Jian, "Data-driven spectrum partition for multiplexing urllc and embb," *IEEE Transactions on Cognitive Communications and Networking*, vol. 9, no. 2, pp. 386–397, 2022.
52. C. Chaccour, M. N. Soorki, W. Saad, M. Bennis, and P. Popovski, "Can terahertz provide high-rate reliable low-latency communications for wireless vr?" *IEEE Internet of Things Journal*, vol. 9, no. 12, pp. 9712–9729, 2022.
53. S. Gong, X. Lu, D. T. Hoang, D. Niyato, L. Shu, D. I. Kim, and Y.-C. Liang, "Toward smart wireless communications via intelligent reflecting surfaces: A contemporary survey," *IEEE Communications Surveys & Tutorials*, vol. 22, no. 4, pp. 2283–2314, 2020.
54. N. Rozen-Schiff, A. Navon, I. Pechtalt, L. Bruckman, and Y. Boyuan, "Aqua: Adding bandwidth allocation to quic for metaverse multi-stream applications," in *IEEE International Conference on Metaverse Computing, Networking and Applications (MetaCom)*. IEEE, 2023.
55. H. Song, "Enable cross-domain qos for internet-scale metaverse," in *IEEE International Conference on Metaverse Computing, Networking and Applications (MetaCom)*. IEEE, 2023.
56. M. Kalfa, M. Gok, A. Atalik, B. Tegin, T. M. Duman, and O. Arikan, "Towards goal-oriented semantic signal processing: Applications and future challenges," *Digital Signal Processing*, vol. 119, p. 103134, 2021.
57. H. Zhang, S. Shao, M. Tao, X. Bi, and K. B. Letaief, "Deep learning-enabled semantic communication systems with task-unaware transmitter and dynamic data," *IEEE Journal on Selected Areas in Communications*, vol. 41, no. 1, pp. 170–185, 2022.

58. F. Aschenbrenner, T. Shreedhar, O. Gasser, N. Mohan, and J. Ott, "From single lane to highways: Analyzing the adoption of multipath tcp in the internet," in *2021 IFIP Networking Conference (IFIP Networking)*, 2021, pp. 1–9.
59. T. Shreedhar, N. Mohan, S. K. Kaul, and J. Kangasharju, "Qaware: A cross-layer approach to mptcp scheduling," in *2018 IFIP Networking Conference (IFIP Networking) and Workshops*, 2018, pp. 1–9.
60. N. Mohan, T. Shreedhar, A. Zavodovski, J. Kangasharju, and S. K. Kaul, "Is two greater than one?: Analyzing multipath tcp over dual-lte in the wild," 2019.
61. S. J. Siddiqi, F. Naeem, S. Khan, K. S. Khan, and M. Tariq, "Towards ai-enabled traffic management in multipath tcp: A survey," *Computer Communications*, vol. 181, pp. 412–427, 2022.
62. S. Srinivasan, S. Shippey, E. Aryafar, and J. Chakareski, "Fbdt: Forward and backward data transmission across rats for high quality mobile 360-degree video vr streaming," in *Proceedings of the 14th Conference on ACM Multimedia Systems*, 2023, pp. 130–141.
63. W. Wei, J. Han, Y. Xing, K. Xue, J. Liu, and R. Zhuang, "Mp-vr: An mptcp-based adaptive streaming framework for 360-degree virtual reality videos," in *ICC 2021-IEEE International Conference on Communications*. IEEE, 2021, pp. 1–6.
64. P. Si, W. Yu, J. Zhao, K.-Y. Lam, and Q. Yang, "Uav aided metaverse over wireless communications: A reinforcement learning approach," *arXiv preprint* arXiv:2301.01474, 2023.
65. R. Shahzadi, M. Ali, H. Z. Khan, and M. Naeem, "Uav assisted 5g and beyond wireless networks: A survey," *Journal of network and computer applications*, vol. 189, p. 103114, 2021.
66. 3GPP, "Mobile Metaverse Services," 3rd Generation Partnership Project (3GPP), Technical Specification (TS) 22.156, 09 2023, version 1.0.0. [Online]. Available: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=4199
67. S. E. Trevlakis, A.-A. A. Boulogeorgos, D. Pliatsios, K. Ntontin, P. Sarigiannidis, S. Chatzinotas, and M. Di Renzo, "Localization as a key enabler of 6g wireless systems: A comprehensive survey and an outlook," *arXiv preprint* arXiv:2302.02107, 2023.
68. N. Chukhno, S. Trilles, J. Torres-Sospedra, A. Iera, and G. Araniti, "D2d-based cooperative positioning paradigm for future wireless systems: A survey," *IEEE Sensors Journal*, vol. 22, no. 6, pp. 5101–5112, 2021.
69. M. Ammous, H. Chen, H. Wymeersch, and S. Valaee, "Zero access points 3d cooperative positioning via ris and sidelink communications," *arXiv preprint* arXiv:2305.08287, 2023.
70. M. Rossanese, P. Mursia, A. Garcia-Saavedra, V. Sciancalepore, A. Asadi, and X. Costa-Perez, "Designing, building, and characterizing rf switch-based reconfigurable intelligent surfaces," in *Proceedings of the 16th ACM Workshop on Wireless Network Testbeds, Experimental evaluation & CHaracterization*, 2022, pp. 69–76.
71. Widelity, "Secure and trusted communications networks reimbursement program report," FCC, Tech. Rep., 2021, https://docs.fcc.gov/public/attachments/DA-21-355A1.pdf
72. A. Zare, A. Aminlou, M. M. Hannuksela, and M. Gabbouj, "Hevc-compliant tile-based streaming of panoramic video for virtual reality applications," in *Proceedings of the 24th ACM international conference on Multimedia*, 2016, pp. 601–605.
73. M. Hosseini and V. Swaminathan, "Adaptive 360 vr video streaming: Divide and conquer," in *2016 IEEE International Symposium on Multimedia (ISM)*. IEEE, 2016, pp. 107–110.
74. K. K. Sreedhar, A. Aminlou, M. M. Hannuksela, and M. Gabbouj, "Viewport-adaptive encoding and streaming of 360-degree video for virtual reality applications," in *2016 IEEE International Symposium on Multimedia (ISM)*. IEEE, 2016, pp. 583–586.
75. C. Westphal, "Challenges in networking to support augmented reality and virtual reality," *IEEE ICNC*, 2017.
76. S. Sukhmani, M. Sadeghi, M. Erol-Kantarci, and A. El Saddik, "Edge caching and computing in 5g for mobile ar/vr and tactile internet," *IEEE MultiMedia*, vol. 26, no. 1, pp. 21–30, 2018.
77. P.-L. Poirion, J. Leguay, and L. Ruosi, "Scalable request routing for vr-ready cdns," in *2018 21st Conference on Innovation in Clouds, Internet and Networks and Workshops (ICIN)*. IEEE, 2018, pp. 1–6.

78. Y. Liu, J. Liu, A. Argyriou, L. Wang, and Z. Xu, "Rendering-aware vr video caching over multi-cell mec networks," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 3, pp. 2728–2742, 2021.
79. J. Santos, J. van der Hooft, M. T. Vega, T. Wauters, B. Volckaert, and F. De Turck, "Srfog: A flexible architecture for virtual reality content delivery through fog computing and segment routing," in *2021 IFIP/IEEE International Symposium on Integrated Network Management (IM)*. IEEE, 2021, pp. 1038–1043.
80. G. Shen, J. Dai, H. Moustafa, and L. Zhai, "5g and edge computing enabling experience delivery network (xdn) for immersive media," in *2021 IEEE 22nd International Conference on High Performance Switching and Routing (HPSR)*. IEEE, 2021, pp. 1–7.
81. S. Oh, I.-K. Jeong, I. Aliyu, C.-R. Yu, T.-W. Um, and J. Kim, "Simulation of user arrangement in virtual reality for design metaverse network: Focus on statistical distribution," in *2023 14th International Conference on Information and Communication Technology Convergence (ICTC)*. IEEE, 2023, pp. 1521–1523.
82. J. Burke, "Browsing an augmented reality with named data networking," in *2017 26th International Conference on Computer Communication and Networks (ICCCN)*. IEEE, 2017, pp. 1–9.
83. Y. Yu, A. Afanasyev, D. Clark, K. Claffy, V. Jacobson, and L. Zhang, "Schematizing trust in named data networking," in *proceedings of the 2nd ACM Conference on Information-Centric Networking*, 2015, pp. 177–186.
84. Real/Virtual World Integration Domain Working Group, "Real/Virtual World Integration Domain Working Group Charter," Metaverse Standards Forum, Group Charter, 02 2023, version 1.0.0. [Online]. Available: https://portal.metaverse-standards.org/document/dl/5048
85. B. Jones, M. Goregaokar, and R. Cabanier, "WebXR device API," W3C, Candidate Recommendation, oct 2023, https://www.w3.org/TR/2023/CRD-webxr-20231005/
86. T. Rosa, "Cooperation, competition and patents: Understanding innovation in the telecommunication sector," *working paper*, 2019. [Online]. Available: https://wwws.law.northwestern.edu/research-faculty/clbe/events/innovation/documents/rosa_jmp.pdf
87. "WebXR implementation report," W3C, Tech. Rep., oct 2023, https://wpt.fyi/results/webxr