frontiers in Communications and Networks

TYPE Systematic Review
PUBLISHED 14 April 2026
DOI 10.3389/frcmn.2026.1761635

EDITED BY
Mauro Femminella,
University of Perugia, Italy

REVIEWED BY
Olga Chukhno,
Mediterranea University of Reggio Calabria, Italy
Xin Li,
Beijing University of Posts and Telecommunications (BUPT), China

*CORRESPONDENCE
Bowen Qin,
bowen.qin@bth.se

RECEIVED 05 December 2025
REVISED 13 March 2026
ACCEPTED 16 March 2026
PUBLISHED 14 April 2026

CITATION
Qin B, Tutschku K and Hu Y (2026) Requirements for network resource management in multi-user extended reality systems: a systematic review. *Front. Commun. Netw.* 7:1761635. doi: 10.3389/frcmn.2026.1761635

Requirements for network resource management in multi-user extended reality systems: a systematic review

Bowen Qin*, Kurt Tutschku and Yan Hu

Department of Computer Science, Blekinge Institute of Technology, Karlskrona, Sweden

Introduction: Extended Reality (XR) applications have been widely adopted in gaming, healthcare, professional training, and social interaction. Among them, networked multi-user XR (MU-XR) systems are increasingly important. They allow users to communicate and collaborate remotely by streaming real-time data. Moreover, MU-XR systems are becoming increasingly interactive, offering a variety of modes and sensors for user interaction and even allowing users to share their social engagements. Hence, global network conditions increasingly shape user experience (UX). However, current performance engineering remains largely network-centric and predominantly evaluates network capabilities from an individual-user perspective, thereby overlooking how global network resource management (NRM) affects multi-user interaction, multi-user UX, and immersion.
Methods: This has led to a lack of a unified framework that can connect multi-user experiences from an application perspective with underlying NRM. To address this gap, this paper conducts a systematic literature review (SLR) of 56 studies published between 2015 and 2025, analyzing UX evaluation methods, multi-user network performance metrics, and NRM-related network functions.
Results: The findings reveal that most studies focus on latency minimization to ensure a real-time, immersive collaborative experience, while only a few address network-level support for resource allocation and synchronization for user groups that require immersion. Among the UX evaluation methods, only 6 of the 56 studies conducted user studies, indicating a clear lack of empirical support.
Discussion: Furthermore, this SLR proposes a new model for structuring NRM requirements across various MU-XR scenarios, starting at the XR application level, identifying immersion techniques at the XR level, and connecting them to global network performance, aiming to bridge MU-XR applications and their NRM requirements.

KEYWORDS
artificial intelligence, augmented reality, extended reality, mixed reality, multi-user, network resource management, virtual reality

1 Introduction

Extended Reality (XR) has been identified by IMT-2030 as a pivotal application for delivering immersive experiences and seamless human-machine interactions, serving as a cornerstone of the next-generation of digital infrastructure (ITU-R, 2023). In multi-user XR (MU-XR), synchronous interaction in shared virtual environments can further enhance immersiveness (Guertin-Lahoud et al., 2023). Meanwhile, driven by practical demands such as cost reduction (Bayro et al., 2025), remote collaboration (Chang et al., 2023), and environmental sustainability (e.g., remote meetings and training to reduce travel-related carbon emissions (Krodel et al., 2024)), distributed MU-XR has attracted significant attention in recent years (Alsader et al., 2025). However, as the underlying data-transmission infrastructure for distributed MU-XR, the network must support ultra-high bandwidth, real-time responsiveness, high reliability, and scalable computing resources (Huang et al., 2023). Meeting these requirements simultaneously is challenging in practice.

In distributed MU-XR, participation over heterogeneous access networks introduces remote transmission latency and jitter. This can disrupt the timing alignment between users' real-world actions and the virtual environment, reducing immersion and potentially causing discomfort (Warburton et al., 2023; Stauffert et al., 2020). In addition, heterogeneous network conditions and software capabilities can cause interaction-level desynchronization among users, thereby reducing interaction smoothness (Tuncer et al., 2019). Given these challenges, network resource management (NRM) is highly relevant to MU-XR scenarios, as it determines whether sessions remain stable and whether users maintain a consistent shared context in the face of changing network conditions (Trinh and Muntean, 2022).

However, few studies explicitly derive NRM requirements from the user perspective. Specifically, existing XR networking studies often infer system performance requirements from assumed 5G/6G capabilities (e.g., peak rates, minimum latency, and reliability) (Pindi and Velez, 2025), and there is still no systematic literature review (SLR) that synthesizes the requirements for NRM in MU-XR systems. Therefore, this paper systematically reviews the research progress on NRM in MU-XR. To narrow this gap, we adopt a top-down approach, starting with the MU-XR system, examining how it drives requirements for NRM in this cross-domain area, and providing concrete suggestions for future MU-XR networking research.

1.1 Contribution and research questions

This **top-down approach** aims to complement existing work on NRM for XR. We argue that this approach can provide application-oriented guidance on the requirements that future NRM techniques for XR systems should consider, thereby complementing the typical network-focused quality-of-service (QoS) performance metrics (delay, delay variation, throughput, and bandwidth).

a. This paper provides a systematic review of requirements for NRM concepts specific to MU-XR systems.
b. It presents a top-down, application-oriented model that reviews different MU-XR application types, their immersion techniques, multi-user network metrics, and derives the associated NRM requirements.
c. It identifies key challenges and potential opportunities (in Section 6) for future work on the MU-XR NRM.

In order to achieve these contributions, we conduct an SLR, where the review of the considered papers is guided by these research questions:

RQ 1: What are the primary applications of MU-XR systems, and which immersion techniques (XR and sensor modalities) are most commonly used in these applications?

RQ 2: What evaluation methods of UX and multi-user network performance metrics are used to evaluate network and media performance?

RQ 3: What network technologies and functions support MU-XR interactions?

RQ 4: What are the present challenges and opportunities in NRM for MU-XR systems?

In this paper, we use the term "network function" to refer to the blocks, tasks, services, or techniques involved in NRM. We note that modern networks include many network functions beyond NRM; however, our focus is limited to those that directly support NRM in XR systems. The reader is referred to (ETSI, 2023) for a general definition of network functions.

The remainder of this paper is organized as follows. Section 2 presents the system perspective of XR interactions and corresponding network requirements. Section 3 reviews related work, and Section 4 describes the method used in this SLR. Section 5 summarizes the reviewed studies and key findings. Section 6 provides an overall discussion, and Section 7 concludes the paper.

2 System view and model of human-centered Intelligent Realities

Intelligent Realities (IR) builds upon XR technologies to enable human–computer interaction by integrating physical and digital environments. It incorporates AI, Digital Twin (DT), Internet of Things (IoT), robotics, wearable devices, and advanced networks to enhance interaction efficiency and user experience (IEEE ICIR, 2025 Conference Committee, 2025; Hu Y. et al., 2024). The Human-Centered Intelligent Realities (HINTS) project (Sundstedt et al., 2023) places greater emphasis on technology "augmenting" rather than "replacing" human capabilities. HINTS proposes using multimodal techniques, including eye tracking, emotion recognition, and behavioral prediction, to enhance the richness and satisfaction of UX.

In contrast to the conventional focus of traditional XR on video quality and network performance, this research focuses on the relationship between NRM, media presentation, and network performance. It emphasizes smooth interaction and intelligent adaptability, ensuring that virtual environments can support collaborative interactions among multiple users in a natural, intelligent, and efficient manner. We constructed this IR model (Figure 1) to offer a comprehensive perspective on networked XR systems. The IR model connects features and requirements across *XR applications*, *user topology and interaction*, *immersion*, *networking-compute continuum*, and *NRM*. Specifically, the model follows a layered requirement logic. Application scenarios initially define collaborative objectives and determine the user topology. These interaction requirements introduce immersion-level constraints, including sensing support and synchronization. These constraints, in turn, affect the choices along the networking-compute continuum (e.g., processing in a device-, edge-, or cloud-based setting), and are eventually translated into actual NRM mechanisms. Therefore, the model helps researchers and readers understand, from a top-down perspective, how different applications shape underlying network requirements, enabling the

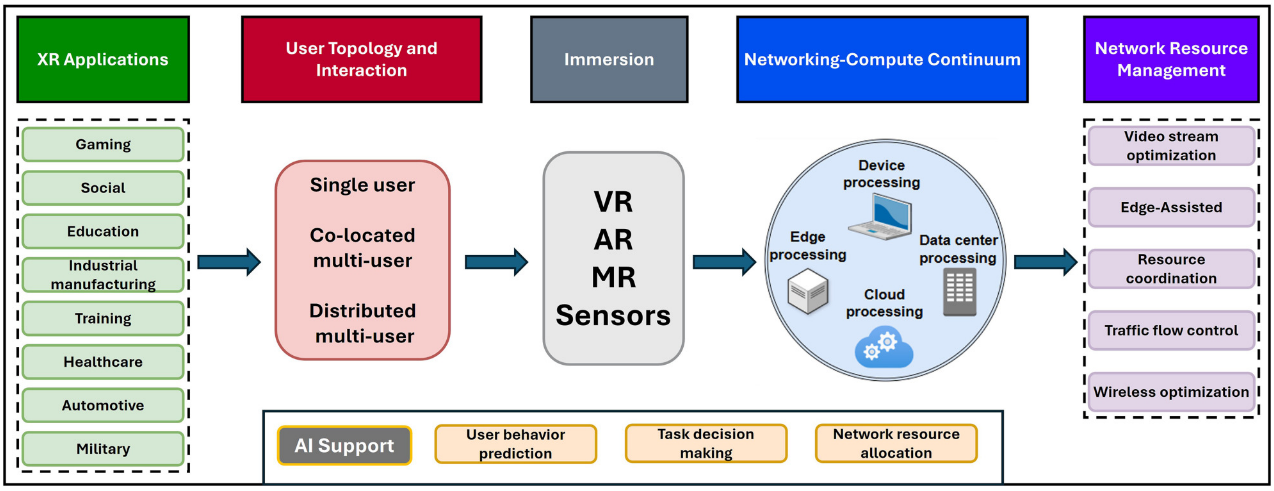

FIGURE 1
General system model of IR and core parts.

development of interaction strategies and networking solutions that enhance UX. As an example, 360LIVECAST (Zhao et al., 2025) considers live multi-user 360 video streaming, in which the continuity of interaction requires tightly regulated motion-to-photon (MTP) latency below 25 ms. This is an application-level requirement and limits user interaction and immersion, such as continuous head-motion tracking. To meet these constraints, computation is synchronized between the device and the edge using a hierarchical multicast design. This is translated into resource allocation and network-level traffic control, both of which are latency-sensitive. Each layer of the IR model has been detailed in the following subsections.

XR Applications encompass a wide range of areas, including gaming, social media, education, industrial manufacturing, training, healthcare, automotive, and military applications (Kourtesis, 2024). These areas have varying requirements related to immersion, interactivity, and scalability (Norambuena et al., 2025). Gaming requires responsiveness and stability in frames, while industrial and medical training require accuracy, synchronization, and reliability. These application-based requirements serve as the initial starting point for the model design of networked XR systems.

From the *User Topology and Interaction* perspective, XR applications are featured by user numbers and location patterns. Single-user, co-located multi-user, and distributed multi-user are typical classifications in XR applications. Therefore, interaction refers to the user behaviors between any two involved users, as well as between users and virtual world objects. In a single-user case, the interaction is between the individual and a virtual world. Local processing and rendering capabilities largely dictate system performance. However, with the development of collaborative features in XR, co-located multi-user experiences emerge, with data being exchanged over short-range, high-bandwidth networks such as Wi-Fi 6 and private 5G-based networks. Such environments require temporal synchronization and consistency of state between users (Sonkoly et al., 2023; Mallik et al., 2024). At the top level, distributed MU-XR is a system that has no space limitations because it connects users living in different regions and is connected through wide-area networks. This shift redefines the local performance parameters of latency, throughput, and fairness in the context of global networking challenges.

For *immersion*, to meet the growing interaction requirements, the immersion layer combines VR, AR, and MR, with various sensors (eye tracker, hand tracker, and so on) (Kourtesis, 2024). These elements receive continuous bodies of spatial, visual, and behavioral information, thus creating a connection between the physical and virtual worlds.

To maintain an immersive experience, the model requires not only immersion devices, but a *network-compute continuum* for executing IR applications (Patra et al., 2024). In a single-user scenario, most computation occurs on the device to minimize round-trip latency. When a large number of co-located users are involved, edge servers are used to handle rendering and synchronization functions, using locality to minimize delay and improve bandwidth efficiency. In distributed multi-user scenarios, cloud and data-center processing is necessary to coordinate global state, user status, and large-scale updates of 3D scenes (Skorin-Kapov et al., 2025). The MU-XR system becomes part of a dynamic continuum, whereby computation is dynamically coordinated to the latency sensitivity, network congestion, and workload complexity.

NRM ensures system performance and reliability (Fowler and Souihi, 2025). It encompasses mechanisms across the network stack, from the physical, transport, and network layers to the application layer. All these mechanisms work together to ensure low latency, stable throughput, and synchronized consistency in the presence of dynamic multi-user conditions. Moreover, *AI* enables predictive and adaptive intelligence, including predicting both user behavior and network load, enabling adaptive resource allocation.

Overall, this model encompasses application requirements, users, immersion technologies, the Networking-Compute continuum, and NRM. It illustrates the development of XR from a non-collaborative experience to a collaborative, intelligent, and network-integrated one. The Refined multi-user IR model section further elaborates on this model based on both our review findings and our own insights.

3 Related work

This SLR focuses on the MU-XR requirement for the NRM. However, existing reviews rarely address both MU-XR applications and UX, as well as XR NRM, in a single study. Our analysis reveals that relevant reviews can be categorized into two types: one centered on MU-XR applications and UX, and the other focused on XR NRM. The following sections summarize the consensus and shortcomings of each category, thereby establishing the necessity of this review.

3.1 Literature review on MU-XR

Existing MU-XR reviews predominantly focus on the usability and UX (e.g., immersion, presence, and collaborative efficiency) of various applications and systems, exploring their strengths and limitations.

Regarding application usability, Mathias et al. (2025) proposed a human-centered virtual world model for Architecture, Engineering, Construction, and Operations (AECO) to integrate XR with existing Building Information Modeling (BIM) data. Wang P. et al. (2025) reviewed XR-enabled remote collaboration tools, discussing their application usability and potential in manufacturing and training.

Regarding UX, several papers explored the collaborative role of XR technologies in multi-user interactions. For example, Caserman et al. (2020) reviewed full-body motion reconstruction in immersive VR and found that full-body avatars and body tracking enhanced presence and immersion in multi-user environments. Wang Y. et al. (2024) systematically identified factors influencing user engagement from a collaborative work perspective. They pointed out that distributed, multi-user, and multi-modal social participation remained insufficiently studied. In addition, Borhani et al. (2024) and Wang et al. (2026) specifically explored the impact of annotation techniques and sketch cues techniques on collaborative communication efficiency in remote interactions. The former argued that annotation technology improved user comprehension efficiency. However, research remained fragmented, making it hard to develop a framework that incorporated display devices, input methods, annotation types, and collaborative forms. The latter pointed out that sketch cue technology improved spatial referencing and collaborative communication efficiency, but lacked unified design specifications and evaluation frameworks. Moreover, most papers discussed enhancing UX through MU-XR in specific scenarios. Given the large number of such studies, we use education and training as an illustrative example. Jackson et al. (2025) summarized MU-XR's positive impact on student engagement and practical training. Similarly, Chen and Konomi (2022) conducted a systematic review specifically on "remote collaborative learning," further emphasizing experiential benefits and implementation challenges in cross-spatial collaborative learning.

Collectively, these MU-XR reviews were quite effective in answering questions such as "What is possible with MU-XR?" "What are the factors that affect the experience?" and "How can interaction design improve collaboration?" However, their common limitations were also apparent: Network infrastructure and resource management were usually addressed as external requirements, and discussions were often vague ("low latency, high bandwidth," etc.). There was a lack of translation of multi-user collaboration semantics (e.g., shared context consistency, fair interaction responsiveness, stable synchronous communication) into evaluable NRM objectives and network functions.

3.2 Literature review on network resource management for XR

Another category of reviews focuses on network and system implementation perspectives. XR network support methods were summarized around 5G/6G capabilities, media transmission, and resource allocation mechanisms. However, these studies generally assumed that XR mainly entailed transmission and streaming optimization problems.

Regarding standards and network evolution, Esswie and Repeta (2023) discussed the evolution of 3GPP standards toward "true XR support" in 5G/5G-Advanced/6G, noting that XR could not be adequately characterized by traditional eMBB/URLLC QoS classifications and that it required new service and wireless design paradigms. Chaudhari (2025) explained how to enable the tactile Internet using 6G to meet application needs and design considerations on the physical and media-transmission level.

Regarding media transmission optimization and network architecture, Wu et al. (2024) reviewed signal processing techniques for XR communications; Alsader et al. (2025) summarized client adaptation, SDN/MEC coordination, and AI/ML approaches from a QoE-driven adaptive video streaming perspective, noting that existing QoE modeling and network control still struggled to meet XR's stringent real-time and high-immersion demands reliably; Lv et al. (2024) proposed a QoS-based resource allocation model and discussed resource allocation algorithms in fog wireless access networks. In addition, some reviews addressed network simulation and system support. For example, Evgenieva et al. (2025) reviewed 6G network simulators and their potential XR applications, and Hatami et al. (2024) examined the difficulties of the Metaverse in real-time operation across a variety of applications. In specific applications, Pooyandeh et al. (2022) studied cybersecurity threats in the AI-enabled Metaverse and surveyed defense mechanisms. In industrial 6G and IoT, Mon et al. (2025) surveyed the practice of DTs with Industry 5.0, mentioning XR as an enabling technology.

Overall, network-side reviews provided relatively comprehensive coverage of the technical chain, including "standard capabilities, architectures, protocols, coding and adaptation, resource allocation, and security." However, their research subjects and evaluation dimensions were skewed: Most research focused on objective network performance measures (e.g., latency, jitter, throughput, reliability) and primarily aimed to meet or optimize them. They rarely went into detail about how experiential goals on the user side, such as immersion, interactive usability, and collaborative understanding, should inform the control of networks and resource management. Furthermore, existing reviews often implicitly assumed XR was used within a single-user or single-session streaming paradigm, and rarely provided a specific discussion of distributed multi-user XR. Crucially, they did not systematically characterize NRM strategies in the cross-regional, multi-user concurrent online conditions, while paying limited attention to the network architecture patterns that supported multi-user synchronization and shared context consistency. Consequently, important problems in multi-user situations, such as synchronization consistency, fairness, and experience stability under heterogeneous conditions, were not well represented in the network literature, lacking a systematic, requirement-driven framework or synthesis.

To address this gap, this SLR systematically analyzes NRM requirements in MU-XR systems, focusing on several critical user-centric issues, including synchronization, consistency, and experience stability in distributed collaborative scenarios.

4 Methods

The SLR follows the *PRISMA 2020 (Preferred Reporting Items for Systematic Reviews and Meta-Analyses)* reporting guidelines (Page et al., 2021) and transparently reports the study selection procedure and overall research process. It aims to ensure methodological rigor and standardization in the process.

4.1 Protocol and information sources

The planning phase focused on defining the research questions and scope, determining the search strings, and collecting relevant papers.

4.1.1 Research questions and scope

The objective of this paper is to systematically investigate NRM in application-oriented MU-XR systems (detailed motivations were discussed in the *Introduction*). We proposed four research questions as described in the *Introduction* (RQ1-RQ4) to explore applications, immersion techniques, and requirements for NRM in MU-XR.

4.1.2 Search strategy

We derived candidate search terms by identifying common expressions for multi-user collaboration, XR modalities, and network-related concerns from domain literature. We grouped synonyms and spelling variants into three anchor concepts: "Multi-user", "Network", and "XR", because each represents a core dimension of the topic and appears across relevant venues. We also considered related terms for each concept to ensure accuracy and comprehensiveness. Specifically, we found that "multi-user" has multiple variations in writing, including "multiple users" or "multi-user", so we used the fuzzy search string "multi* user" to encompass these variants. Given that our study specifically targets distributed multi-user environments, we extended the search string by including the terms "distributed users" and "remote users" to ensure comprehensive coverage of geographically separated user settings. The terms "VR", "AR", and "MR" as subsets of "XR" were also included in the search query. Since both acronyms and their full forms are commonly used in the literature, we incorporated all variants into the search string. For the final search element, we targeted the network domain in alignment with the scope of this study. As network-related research frequently involves lexical variations such as "networked" and "networking", we used the wildcard term "network*" to avoid omitting relevant publications. Based on these search keywords, the final search string was constructed, as presented in Table 1.

4.1.3 Information sources and time range

To collect comprehensive and relevant studies, we reviewed three digital libraries: Scopus, IEEE Xplore, and ACM Digital Library. The search period ranged from January 2015 to December 2025, which reflects the period when MU-XR and advanced networking gained popularity.

4.2 Study selection process

The study selection process involved identification, de-duplication, and two rounds of screening (title/abstract and full-text). To ensure the quality of the sources and the possibility of obtaining results, we restricted the search to peer-reviewed conference papers and journal articles in English.

4.2.1 Eligibility criteria

Detailed inclusion and exclusion criteria were applied to the identified literature, as outlined in Table 2. These criteria help minimize the influence of human subjectivity during the screening process, ensuring the final research remains highly relevant to our topic.

As our study focuses on network resource management in multi-user XR (MU-XR), the included literature must focus on MU-XR interaction and network technologies or functions. In addition, since this study adopts a joint perspective that bridges network mechanisms and user-level impact, we also include works that focus on user experience and evaluation metrics to reflect the user-side effects of network resource management.

Regarding the exclusion criteria, "Missing terms" refers to papers that do not clearly address MU-XR or network aspects in their core content. For example, studies focusing only on single-user VR systems without any network component were removed. "Keywords used as examples or background" indicates that MU-XR or network is mentioned only briefly in the introduction or related work, while the main contribution lies elsewhere. For example, a paper mentioning XR as an application scenario while primarily studying general system optimization was excluded. "Irrelevant topics" refers to works that are outside the scope of MU-XR interaction, user experience, or network technologies or functions, such as generic communication protocols without XR context. Finally, survey or position papers were excluded so that only original research with concrete technical contributions was retained.

4.2.2 Screening flow and results

4.2.2.1 Identification

The flow of documents through the selection process is illustrated in the *PRISMA flow diagram* (Figure 2). To carry out this work, we used Scopus (n = 504), IEEE Xplore (n = 207), and the ACM Digital Library (n = 83) to conduct a database identification, which resulted in an initial corpus of 794 records. We then removed 147 duplicated records observed in multiple sources, leaving 647 unique records to be screened.

TABLE 1 Search strings for various research elements.

Research elements	Search string
Multi-user	"Multi* user" OR "distributed user" OR "remote user"
XR	vr OR ar OR mr OR xr OR "virtual reality" OR "augmented reality" OR "mixed reality" OR "extended reality"
Network	network*

TABLE 2 Inclusion and exclusion criteria used for filtering MU-XR literature.

Inclusion criteria	Exclusion criteria
Focuses on MU-XR interaction	Missing terms
Focuses on user experience and evaluation metrics	Keywords used as examples or background
Focuses on network technologies or functions	Irrelevant topics
Written in English	Survey or position paper

4.2.2.2 First-round screening

The first round of screening, based on abstracts, titles, and keywords, resulted in the exclusion of 487 records. These were removed based on the exclusion criteria of missing terms or keywords that were used only as samples, background information, or were not the focus of the study. This process left 160 papers for full-text screening.

4.2.2.3 Second-round screening

Before the second-round screening, we searched for full-text access to the papers. After searching multiple avenues, we finally found that the full text of all 160 papers was available for retrieval. During the full-text screening stage, 104 records were excluded. These exclusions were based on criteria such as being irrelevant to the intersection of MU-XR and networking (e.g., single-user or non-network studies) or being survey/position papers.

4.2.2.4 Included

Following these systematic procedures, 56 primary studies were selected that met all eligibility criteria and formed the final corpus for the review.

4.3 Data extraction and synthesis

From the 56 finally included papers, key information was extracted and categorized based on the research questions: (1) MU-XR applications and Immersion techniques; (2) UX evaluation and network metrics; (3) Network technology functions and support. This structured extraction facilitated subsequent categorization and analysis.

We performed a quantitative frequency analysis to determine the technical focus and evaluation metrics in the studies. This strategy enabled us to examine the studies in more detail and make the evidence easier to follow. The synthesized results are presented in tables and figures that clearly summarize the technological methods, evaluation metrics, and research trends. This structured reporting makes the findings easy to understand and helps us identify research gaps and new directions in networked MU-XR.

4.4 Method limitations

We acknowledge that this method has several limitations despite successfully gathering, filtering, and analyzing relevant papers in accordance with the PRISMA 2020 guidelines.

1. The search string may not cover all relevant studies. Even though it has major terms and lexical differences, some of the most common terms in the field, such as "collaborative", "shared virtual environment", "social", and "telepresence", were not incorporated. Including these terms would substantially inflate the search results, with most retrieved records belonging to general HCI research rather than MU-XR network technology. Although their exclusion gives a more precise search, it might have missed a few partially relevant studies.
2. Although Scopus, IEEE Xplore, and ACM Digital Library cover a wide range, no combination of digital libraries is exhaustive. Other databases, such as "Web of Science" or "ScienceDirect", may contain additional relevant studies that were not accessed in this review.
3. Despite creating specific inclusion and exclusion criteria, the screening decisions (e.g., "Irrelevant topics") are based on interpretative judgment. Various researchers can set slightly different boundaries for what MU-XR network resource management research is.
4. For the analysis process, any bias that has been incorporated in the search and screening is likely to affect the further classification and synthesis of studies.

5 Results

5.1 Overview of included papers

Based on our analysis of 56 relevant papers, we summarized the overall findings in Table 3 to fully reveal the development trends and research priorities of networked MU-XR technology. We analyzed the year of publication of each paper. Since MU-XR is still in a stage of rapid development, examining its evolution over time and changes in research popularity can help us understand the technological maturity of this field and its future development direction. We studied the types of immersion techniques, including AR, VR, and XR, as well as the sensing modalities used in the reviewed studies. Since our review did not identify papers in which MR was the primary immersion technology, and studies mentioning both AR and VR generally refer to them under the broader umbrella of XR, we do not list MR as a separate category. Immersion is a key UX factor in XR and metaverse systems. It requires user data collection for system optimization and effective

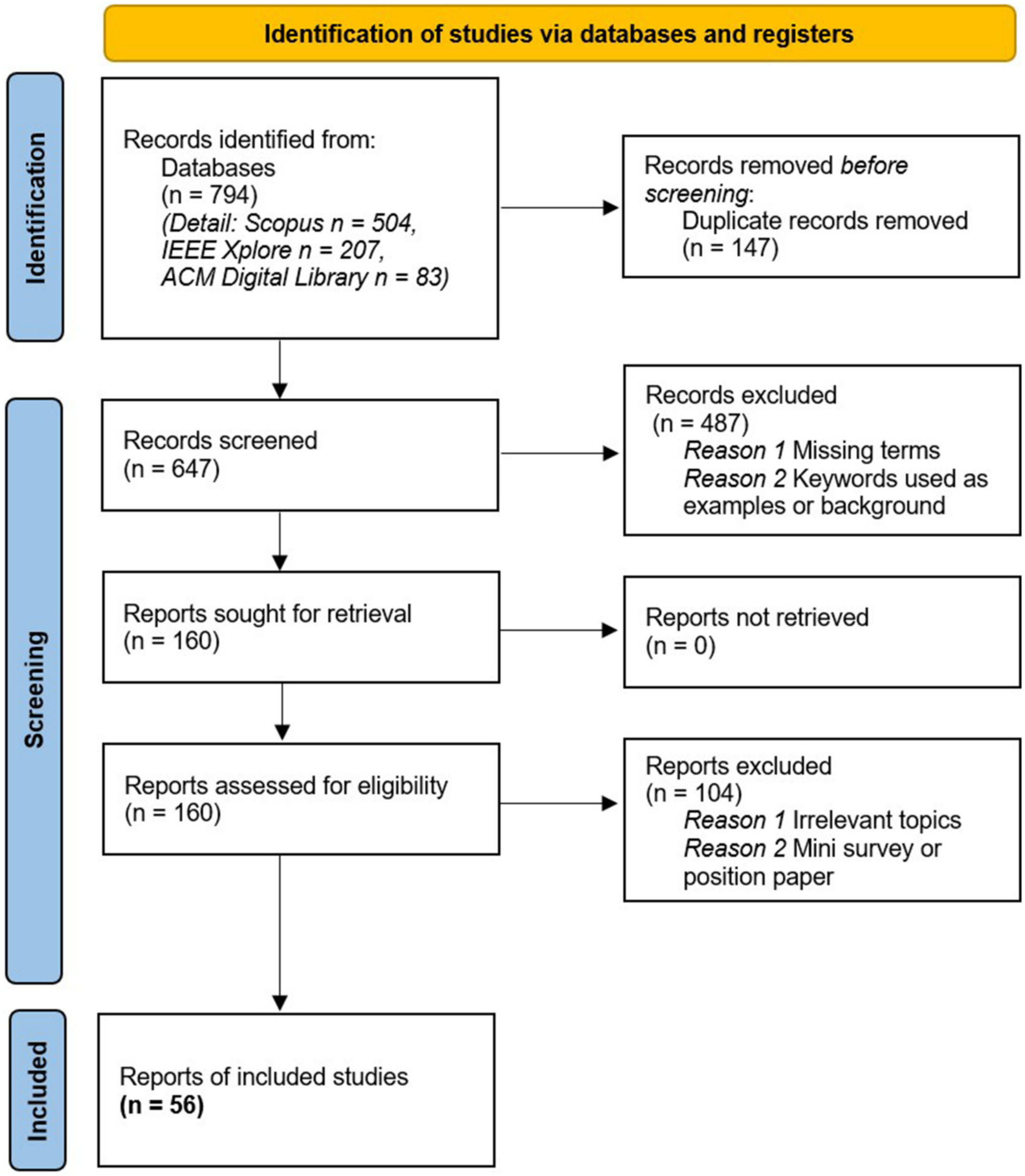

FIGURE 2
PRISMA diagram of the selection process from identification of sources to the final collection of studies, forming a corpus of 56 papers.

visual representation. We therefore summarize the immersion technologies used for visual representation (XR/AR/VR) and the sensors used for user data collection. To connect the user side and the network in MU-XR systems, we also examined the application domains of MU-XR. Finally, we analyzed the three areas mentioned in our research questions: (1) evaluation of UX, (2) Multi-user network performance metrics, and (3) network functions. Notably, during our classification and analysis process, we found that AI appeared frequently in the papers and was considered a major supporting technology for NRM across the system, e.g., for predicting distributed user behavior, network resource prediction, and dynamic allocation, as well as system decision-making. As a currently popular technology, AI can enhance the intelligence and scalability of MU-XR systems. Therefore, we list “AI functions” as a separate research area 4 and discuss the four areas in the next section.

As shown in Figure 3, early publications were sparse: only four appeared during 2015–2019, with none in 2016 or 2017, indicating limited attention in the initial stage. A clear increase emerged in 2020, which we attribute to the surge in demand for remote work and online collaboration during COVID-19. The number of publications decreased in 2021–2022, but it rose to 15 in 2023, the highest in the last 11 years. The output decreased by a minor margin in 2024 and 2025, though it remained high at 13 and 9 papers, respectively. In total, 52 out of the 56 papers (92.9% of the total number of papers) were published within the last 6 years (2020-2025). This publication trend indicates that MU-XR is a developing field, with strong research momentum, and may have

10.3389/fcmn.2026.1761635

TABLE 3 **Statistical classification of the 56 reviewed papers across immersion tech, applications, and areas.**

Year	Paper	Immersion tech				Application						Area		
		AR	VR	XR	Sensors	Meeting	Training	Medical	Gaming	Driving	Evaluation of user experience	MU network performance metrics	Network functions	AI functions
	Count	12	36	8	37	9	8	4	11	1	6	46	53	21
	Ratio	21.4%	64.3%	14.3%	66.1%	16.1%	14.3%	7.1%	19.6%	1.8%	10.7%	82.1%	94.6%	37.5%
2015	Bakri et al. (2015)			■								■	■	
2018	Ahn et al. (2018)		■									■	■	
	Elvezio et al. (2018)		■		①④			■				■		
2019	Cao et al. (2019)		■		①							■	■	■
2020	Apicharttrisorn et al. (2020)	■			⑤							■	■	
	Nath and Wu (2020)	■										■	■	
	Novotny et al. (2020)		■		①④⑦				■			■	■	
	Ran et al. (2020)	■			①⑤	■			■			■	■	
	Wang S. et al. (2020)		■		①							■	■	
	Wang C et al. (2020)	■										■	■	■
	Chakareski and Mastronarde (2020)		■		①	■	■					■	■	
2021	Aloqaily et al. (2021)		■		③							■	■	■
	Kamarianakis et al. (2021)		■		④			■	■			■	■	
	Santos et al. (2021)		■										■	
2022	Cai et al. (2022)	■			⑤			■	■			■	■	■
	Chen et al. (2022)		■		①							■	■	■
	Ren et al. (2022)	■			⑤				■			■	■	■
	Santos et al. (2022)		■										■	
	Zhang et al. (2022)	■			⑤							■	■	■

(*Continued*)

TABLE 3 Continued

Year	Paper	Immersion tech				Application					Area			
		AR	VR	XR	Sensors	Meeting	Training	Medical	Gaming	Driving	Evaluation of user experience	MU network performance metrics	Network functions	AI functions
2023	Alhilal et al. (2023)													
	Cai et al. (2023)													
	Chen and Guo (2023)				⑤									
	Chen et al. (2023)				①⑤									
	Han et al. (2023)				①									
	Huang et al. (2024)													
	Liu et al. (2023)													
	Okamoto et al. (2023)				①③⑤									
	Sonkoly et al. (2023)				①⑤									
	Van Damme et al. (2023)				①④									
	Yu et al. (2023a)													
	Yu H. et al. (2023)				①③⑤									
	Yu et al. (2023b)													
	Zhong et al. (2023)				①									
	Santos et al. (2023)				②⑨									
2024	Ha Huynh et al. (2024)				⑤									
	Sandeepa et al. (2024)				①⑧⑩									
	Van Damme et al. (2024)				①④⑤									
	Wang Z. et al. (2024)													
	Yang et al. (2024)				①									
	Sameri et al. (2024)				①④⑥⑧									
	Badnava et al. (2024)				①									
	Mahmoud et al. (2024)				①③									
	Chen et al. (2024)				①									
	Choi et al. (2024)				①									

(*Continued*)

10.3389/frcmn.2026.1761635

10.3389/frcmn.2026.1761635

TABLE 3 Continued

Year	Paper	Immersion tech				Application						Area		
		AR	VR	XR	Sensors	Meeting	Training	Medical	Gaming	Driving	Evaluation of user experience	MU network performance metrics	Network functions	AI functions
	Gao N. et al. (2024)				①									
	Okafor et al. (2024)				④❶									
	Tseng et al. (2024)													
2025	Casparsen et al. (2025)													
	Choi et al. (2025)				①									
	Lin et al. (2025)				①⑤⑧									
	Zhao et al. (2025)				①									
	Wang S. et al. (2025)				①									
	Huang and Song (2025)													
	Fowler and Souihi (2025)													
	Kuo et al. (2025)													
	Xu et al. (2025)													

Sensor types: ① Head pose tracker (HMD); ② Body pose/position tracker; ③ Eye/gaze tracker; ④ Hand/controller tracker; ⑤ RGB camera (video frames); ⑥ Depth sensor/LiDAR; ⑦ Microphone (audio); ⑧ Physiological sensors (heart rate/HRV, pulse oximeter (SpO_2), ⑨ GSR/EDA); ⑩ GPS receiver; Wheel speed sensor; ❶ Keyboard/mouse.

become one of the key directions of future research and applications. Further development of network infrastructure, device capacities, and XR rendering should unlock MU-XR's potential and enable deployment across more domains and use cases.

In Figure 4, VR appears in 64.3% of the reviewed studies. AR accounts for 21.4%, and XR accounts for 14.3%. This distribution shows that MU-XR research is most often evaluated in VR-based settings. Notably, VR reaches 86.4% (19/22) in 2024–2025, suggesting that VR has become the dominant recent trend, likely enabled by improved 360° video pipelines and stronger network support for performance. AR and XR are less frequently used in the reviewed prototypes and experiments. In addition, 66.1% of studies report at least one sensing modality.

Figure 5 shows that the head pose tracker is the most common modality (48.2%). It is followed by RGB camera/video frames (23.2%), hand/controller tracker (12.5%), eye/gaze tracker (7.1%), and physiological sensors (5.4%). The head pose tracker is prevalent because it directly reflects viewpoint dynamics in HMD-based VR. It is also a key input for prediction and adaptive synchronization. For example, Yang et al. (2024) used head-orientation signals to support delay-aware optimization in multi-user VR. Camera signals are also common because they provide environment-facing visual observations. These observations can feed into localization and scene understanding. For example, Apicharttrisorn et al. (2020) relied on camera frames to characterize and analyze system behavior under networked AR settings. Hand/controller tracker and eye/gaze tracker appear less often. They are mainly used to capture interaction intent and visual attention. Physiological sensors in this corpus include heart rate and heart rate variability (HR/HRV), pulse oximetry (SpO_2), and galvanic skin response/electrodermal activity (GSR/EDA). These signals are typically used as proxies for user state to support QoE-aware adaptation. In contrast, several modalities are only reported once (1.8%, 1/56 each). This includes depth/LiDAR, microphone, body pose/position tracker, GPS receiver, wheel speed sensor, and keyboard/mouse. They are usually scenario-specific signal inputs rather than common instrumentation.

Overall, the nearly two-thirds adoption rate of sensors confirms that user information acquisition is a practical component of MU-XR optimization. It also highlights the importance of immersive technologies in network-related MU-XR research. They provide key evidence for performance-oriented optimization.

We further analyzed the practical application scenarios in Figure 6. We included only cases that had been experimentally validated and could be applied in real-world environments. The results showed that MU-XR research mainly focuses on gaming scenarios, followed by meetings, collaborative training, and medical applications, while autonomous driving is relatively less explored, possibly due to its high safety requirements. This distribution likely reflects MU-XR's strengths, as these scenarios often require real-time collaboration among multiple users. Meetings and training rely on efficient information sharing and immersive remote communication, while gaming requires highly synchronized and low-latency multiplayer interactions. All three scenarios have strong market demand and broad commercial potential. Therefore, they have become the main research directions. In addition, application domains such as medical and autonomous driving demand extremely high real-time performance and stringent safety requirements. Although current research in these areas remains limited, advances in NRM and ultra-low-latency communication will make it increasingly feasible to support remote, near-zero-latency MU-XR interactions. Future research on MU-XR will further expand into more application scenarios.

Finally, we analyzed four research areas based on the research questions (in Figure 7). More than 90% of the reviewed papers focus on network functions (94.6%), followed by multi-user network performance metrics (82.1%), AI functions (37.5%), and UX evaluation (10.7%). This shows that most studies focused on network optimization for MU-XR, while research on multi-user experience remained limited. One reason was the lack of a well-established evaluation framework and unified assessment standards for MU-XR scenarios. It also reflects the weak linkage between UX evaluation and network-level design, highlighting the need for stronger integration between these two perspectives. Furthermore, AI appeared in 37.5% of the reviewed studies, and the proportion of studies applying this technology is expected to increase. The detailed analysis of these findings is presented in the following Results section.

5.2 Area 1: Evaluation of UX in MU-XR systems

In MU-XR systems, evaluating UX is an effective way to assess system quality. Among the 56 reviewed studies, six studies invited participants and conducted user studies. All of them used subjective questionnaire-based evaluations to measure UX. Two of them also incorporated physiological data to provide an objective assessment.

Three of these six studies adopted standardized questionnaires. Alhilal et al. (2023) used the NASA-Task Load Index (NASA-TLX) (Hart, 2006) to measure frustration and mental demand, highlighting that prolonged waiting times influenced user frustration. Lin et al. (2025) used the Simulator Sickness Questionnaire (SSQ) (Kennedy et al., 1993) to quantify UX in cybersickness and interaction smoothness. Furthermore, Van Damme et al. (2024) employed a simplified version of the SSQ, the Virtual Reality Sickness Questionnaire (VRSQ) (Kim et al., 2018), which focuses on oculomotor and disorientation indicators to specifically assess the severity of cybersickness. Furthermore, although these studies (Sandeepa et al., 2024; Van Damme et al., 2023; Sameri et al., 2024) did not use standardized questionnaires, they instead used self-designed 5-point or 7-point Likert questionnaires to evaluate UX. For example, Sandeepa et al. (2024) invited 257 participants to assess indicators such as functionality, responsiveness, and smoothness in a virtual cycling platform. Van Damme et al. (2023) compared immersion and pleasantness between using traditional controllers and haptic gloves in a collaborative pizza-making task. Finally, Sameri et al. (2024) examined UX factors such as immersion, collaboration smoothness, pleasantness, and ease of interaction, with a focus on multi-user metrics to evaluate UX.

In objective physiological evaluations, Lin et al. (2025) used Apple Watch and Empatica E4 wristbands to record heart rate variability (HRV) and galvanic skin response (GSR) to evaluate cybersickness. Similarly, Sameri et al. (2024) measured heart rate and GSR to capture users' real-time physiological reactions for evaluation. Both studies combined subjective and objective evaluation methods to form a more comprehensive framework to reflect the UX accurately.

In summary, research focusing on UX in MU-XR systems remains limited, with only 6 out of 56 studies conducting user studies. Most existing studies remain in the simulation or network performance analysis stage. In addition, only half adopted standardized questionnaires, and even those mainly targeted for single-user experiences rather than collaborative multi-user interactions. This highlights the need for developing a standardized multi-user experience questionnaire. Finally, only two papers combined objective physiological data for a comprehensive evaluation of the UX. Therefore, conducting user studies holds significant research potential in multi-user systems. In the future, researchers should focus on establishing multi-user standardized questionnaires and developing research methods that integrate subjective and objective approaches.

5.3 Area 2: Multi-user network performance metrics

Network performance metrics can be used to evaluate system performance objectively. When combined with user-experience assessment results, they help mitigate the errors introduced by user subjectivity in the evaluation process.

We analyzed the Multi-user network performance metrics reported in 56 papers. We found that almost every study, to some extent, either identified the network indicators that its system needed to monitor or reported network metrics measured in practice. Because resources in multi-user systems were often shared, many studies sought to improve resource allocation under constrained budgets. As a result, multi-user network metrics tended to emphasize overall system performance and scheduling behavior rather than a single-link perspective. We summarized the terminology used in the reviewed papers to describe multi-user network performance metrics in Table. 4. Some terms were reported with minor variations or inconsistencies across studies. To ensure clarity and comparability, we grouped them by their underlying intent and semantics, and adopted the most interpretable labels for each category. Notably, some studies evaluated multiple metrics simultaneously.

Across the reviewed studies, latency and throughput remained the primary metrics of interest: 25 papers investigated multi-user latency, and 18 examined multi-user throughput. Inter-user synchronization (4 papers) and fairness (3 papers) were reported less frequently. Global precision and global bandwidth were each discussed in two papers. Since Table 4 already provides brief descriptions of these metrics, we do not repeat formal definitions here. In the following subsections, we discuss each metric in detail and use boldface to highlight the specific metrics emphasized in the corresponding studies.

5.3.1 Multi-user latency

In multi-user latency optimization, researchers most commonly evaluated and improved system performance using total latency and average latency. For example, Aloqaily et al. (2021); Ha Huynh et al. (2024); Nath and Wu (2020); Wang C. et al. (2020) minimized the system-wide total latency under a fixed number of users. In particular, Nath and Wu (2020) introduced an AI-enabled optimization method that dynamically adjusted computational resource allocation to improve performance. In contrast, Cai et al. (2023); Chen and Guo (2023); Chen et al. (2022); Huang et al. (2024); Yu et al. (2023b); Zhang et al. (2022); Zhong et al. (2023); Okafor et al. (2024) used average-latency evaluation to identify stability boundaries. The evaluated latency terms in this group included RTT, mean response time, mean scheduling time, and processing/computation time. Moreover, Huang et al. (2024) additionally examined how the average delay changed as more users joined, thereby validating the scalability of the proposed SCAXR framework.

Four papers further introduced interaction latency and synchronization latency. Interaction latency was currently mostly operationalized in two-user settings. For instance, Apicharttrisorn et al. (2020) defined interaction latency as the elapsed time from User A placing a virtual object to User B observing it. Similarly, Sandeepa et al. (2024) measured the delay between User A initiating motion and User B receiving the corresponding update in a virtual cycling scenario. Elvezio et al. (2018) proposed a model to investigate the interaction between therapists and patients. In addition, Ahn et al. (2018); Bakri et al. (2015) discussed synchronization latency, referring to the delay incurred when the system enforced alignment among multiple users at a common time point or a shared state, which could require waiting and buffering.

To examine scalability and system limits, Alhilal et al. (2023); Cai et al. (2022); Chen et al. (2023); Yang et al. (2024) studied how latency evolved as the number of concurrent users increased, and how the user experience degraded accordingly. Although Zhao et al. (2025) did not vary the user population, it compared alternative transmission schemes and quantified their impact on the latency of concurrent 360° video streaming, while also probing the boundary conditions of the proposed design.

Finally, Ran et al. (2020); Santos et al. (2021); Van Damme et al. (2024) did not primarily focus on computing aggregate latency values, but instead analyzed multi-user latency distributions. For example, Ran et al. (2020) investigated inter-user latency disparities for the same virtual object under constant resource conditions. Santos et al. (2021) grouped users and compared latency across groups. Van Damme et al. (2024) reported latency distributions by link and configuration settings.

5.3.2 Multi-user throughput

Throughput, defined as the amount of data successfully delivered and effectively utilized at the receiver, reflects a system's practical delivery capability. It was widely adopted in 360° video studies. In our MU-XR review, we identified several throughput-oriented evaluation metrics, including aggregated throughput across users, inter-user throughput disparity, link- and role-specific throughput, and normalized or utility-based throughput.

Aggregated throughput was typically computed as the sum of per-user rates. It was often used to characterize peak load capacity by directly adding user rates, e.g., in Alhilal et al. (2023); Van Damme et al. (2023). In addition, Han et al. (2023); Okamoto et al. (2023) examined how the aggregate throughput load evolved under concurrent access as more users joined.

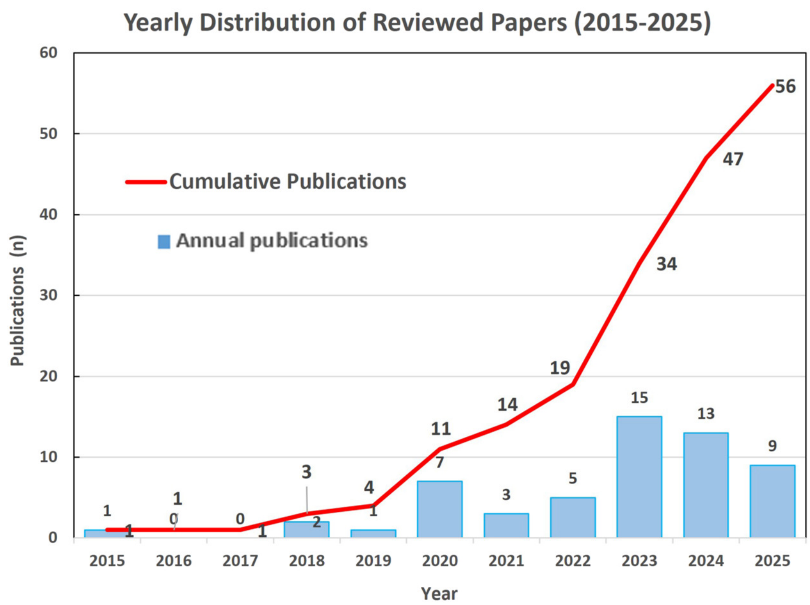

FIGURE 3
Yearly distribution of reviewed papers.

To capture the impact of heterogeneous network conditions, Badnava et al. (2024); Choi et al. (2025); Choi et al. (2024) analyzed throughput disparities across users. In particular, Choi et al. (2024) also investigated how per-user traffic load changed with increasing user scale.

For link- or role-based throughput, Van Damme et al. (2024); Sameri et al. (2024) reported throughput across different links or transmission directions/roles to identify bottlenecks, enabling targeted diagnosis of system limitations.

Normalized, theoretical-capacity, or utility-oriented throughput metrics were not primarily derived from packet-captured Mbps. Instead, throughput was formulated as an objective or a constraint in resource allocation, characterizing achievable capacity under limited resources. For example, Yu H. et al. (2023) proposed a time-varying normalized throughput metric for multi-user evaluation. Moreover, Liu et al. (2023); Chen et al. (2024); Gao N. et al. (2024); Lin et al. (2025) incorporated achievable (measured) or theoretical upper-bound throughput as a constraint to assess system performance. Although Chen et al. (2024) did not explicitly compute an upper bound, it used long-term average throughput as a constraint and was therefore included in this category. Finally, Zhong et al. (2023); Wang S. et al. (2025); Kuo et al. (2025) adopted scheduling-driven utility (e.g., social welfare) throughput metrics for evaluation, and notably, all three studies relied on joint metrics rather than throughput alone.

5.3.3 Inter-user synchronization

Single-user synchronization focuses on the synchronization time between a user and an object. In contrast, multi-user synchronization examines how inter-user synchronization affects overall system behavior. Studies Ha Huynh et al. (2024); Ran et al. (2020) investigated inter-user frame synchronization and used it as evidence for system optimization. Specifically, Ha Huynh et al. (2024) simulated heterogeneous delays across users and proposed an algorithm to minimize frame-time skew from the source to the server among users. In contrast, Ran et al. (2020) analyzed how spatial drift increased the number of coordinate-alignment failures, thereby degrading frame synchronization.

Studies Novotny et al. (2020); Ren et al. (2022) introduced the concept of a multi-user synchronization period. They treated the synchronization period as an interaction window for multi-user collaboration. An overly short period could lead to asynchronous interactions, whereas an overly long period could increase latency and waste bandwidth. Study Novotny et al. (2020) reduced the synchronization period from 100 m to 10 m and evaluated whether interactions remained smooth, aiming to identify an acceptable boundary. Conversely, Ren et al. (2022) defined the global synchronization time as the maximum interaction synchronization time across all edge nodes, and then evaluated system cost under the constraint that all users were synchronized, enabling further system optimization.

5.3.4 Fairness

Fairness directly affects user perception and ensures that experience disparities among users do not become excessive. Researchers typically evaluate fairness by integrating multiple network metrics into one or more fairness-related mathematical models. For example, Wang S. et al. (2020) proposed a QoE optimization model that includes metrics such as video quality, temporal jitter, and spatial jitter. They achieved QoE improvement and user fairness (using Jain's Fairness Index Jain et al. (1984)) optimization in multi-user 360° video services through a hierarchical caching strategy. On this basis, Liu et al. (2023) introduced a fairness index to measure the consistency of QoE for different users in multi-user VR scenarios. By reducing the QoE standard deviation, they achieved a more balanced distribution of quality.

10.3389/frcmn.2026.1761635

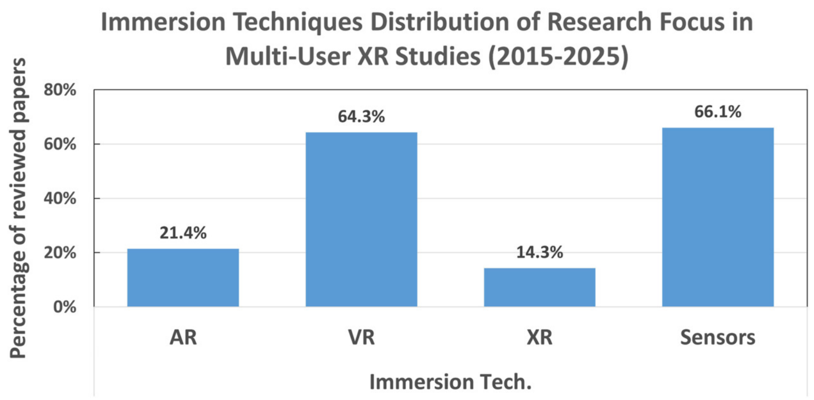

FIGURE 4
Immersion techniques distribution of research focus in MU-XR studies.

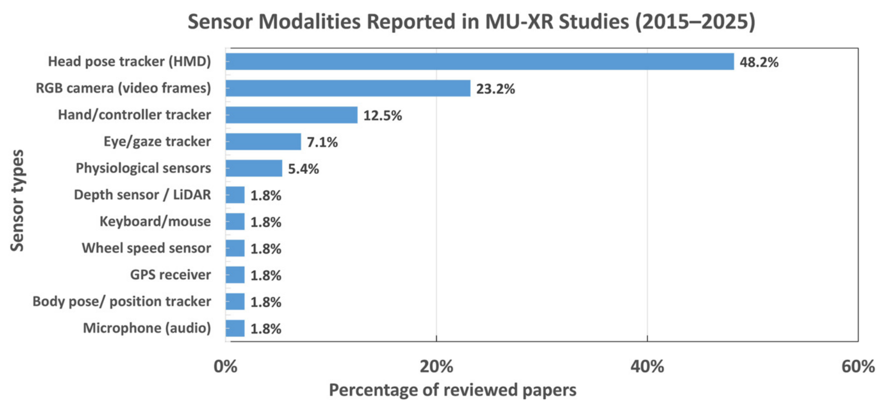

FIGURE 5
Sensor modalities in MU-XR studies.

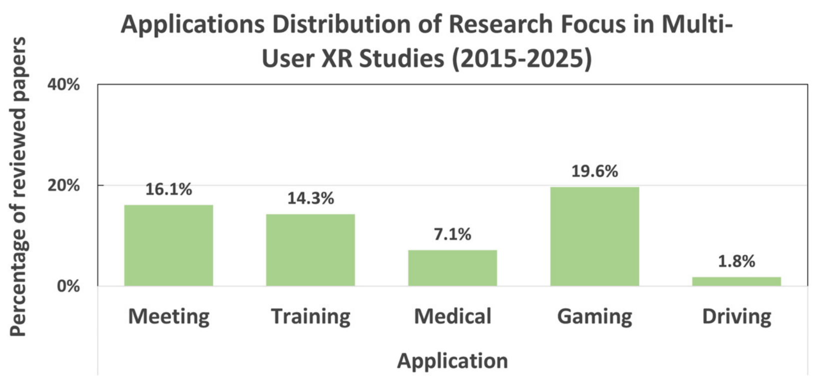

FIGURE 6
Application distribution of research focus in MU-XR studies.

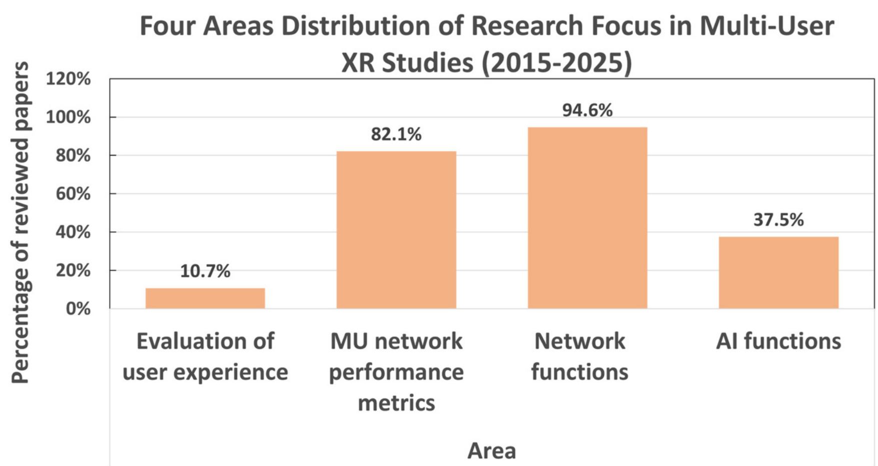

FIGURE 7
Four areas of distribution of research focus in MU-XR studies.

Furthermore, Badnava et al. (2024) proposed an algorithm to optimize overall fairness under limited bandwidth.

5.3.5 Global precision

Precision in multi-user settings was largely similar to that in single-user scenarios. However, because shared resources were limited, systems could not simultaneously maximize precision for all users. As a result, this metric was typically treated as a form of global precision, aimed at improving overall system acceptability. Study Cai et al. (2022) examined the precision of multi-user ID matching in industrial IoT. Study Chen et al. (2023) investigated spatial precision and localization error in AR applications. These studies indicated that measuring and optimizing feature-level precision helped improve collaboration efficiency.

5.3.6 Global bandwidth

In some multi-user network systems, network resources were shared among users, which made global bandwidth allocation a central research direction. Study Cao et al. (2019) proposed a global bandwidth saving ratio and combined it with a machine learning model to optimize video loading strategies. Study Mahmoud et al. (2024) examined bandwidth constraints and allocation decisions under multiple total-bandwidth budgets (ranging from 9 to 33 Mbps).

5.4 Area 3: Functional framework of network technologies

We summarized the network functions widely used in these papers. As the basic framework that interconnects users, it has a final impact on the interactions of multiple users, and the performance of the network environment directly defines the overall UX. Network functions are also the fundamental operational mechanisms that maintain the connectivity, synchronization, and fairness in MU-XR environments.

In accordance with the SLR, the following five categories of functions can be distinguished based on their primary purposes and areas of control: traffic flow control, resource coordination, edge-assisted performance optimization, video stream optimization, and wireless optimization, as summarized in Table 5.

Traffic flow control, as a core function of the communication architecture, was mentioned in 13 papers. Research primarily focused on enhancing network transmission efficiency through the development of new network protocols, congestion control mechanisms, and adaptive routing algorithms. In addition, some papers proposed hybrid transmission strategies (such as combining unicast and multicast) and hierarchical caching architectures to optimize data distribution, significantly improving the smoothness and stability of MU-XR scenarios.

In addition, 13 papers mention resource coordination. These papers significantly improve the transmission efficiency of system data by using distributed network technologies, SDN, etc. The decentralized design effectively reduces the computational pressure on the central server, improving the system's response speed and scalability.

12 papers mentioned Edge-assisted performance optimization, which can improve computational performance. Edge nodes reduce the end-device burden by performing computationally intensive tasks such as rendering and frame interpolation. In addition, intelligent task scheduling further relieves the system's computing pressure and improves overall performance.

Regarding video data transmission, 10 papers mentioned the Video Stream Optimization technology. When dealing with the large amount of video streaming data in multi-user interactions, researchers proposed a tile-based chunked transmission strategy and a frame interpolation algorithm, which effectively reduced the data transmission volume and computational pressure while meeting

TABLE 4 Categorization of Multi-user network performance metrics for MU-XR.

Catalog	Description	Count	Paper
Multi-user latency	Measures interaction and synchronization delays in MU-XR, using algorithm-defined metrics (e.g., total/mean delay) and multi-user latency distributions to evaluate system performance and scalability	25	Ahn et al. (2018); Alhilal et al. (2023); Aloqaily et al. (2021); Apicharttrisorn et al. (2020); Bakri et al. (2015); Cai et al. (2023); Cai et al. (2022); Chen and Guo (2023); Chen et al. (2022); Chen et al. (2023); Elvezio et al. (2018); Huang et al. (2024); Nath and Wu (2020); Ran et al. (2020); Sandeepa et al. (2024); Santos et al. (2021); Van Damme et al. (2024); Yang et al. (2024); Yu et al. (2023b); Zhang et al. (2022); Zhong et al. (2023); Okafor et al. (2024); Zhao et al. (2025); Huang and Song (2025); Wang C. et al. (2020)
Multi-user throughput	Evaluates the effective transmission of data volume for multiple users under resource constraints	18	Alhilal et al. (2023); Han et al. (2023); Liu et al. (2023); Okamoto et al. (2023); Van Damme et al. (2023); Van Damme et al. (2024); Yu et al. (2023a); Zhong et al. (2023); Sameri et al. (2024); Badnava et al. (2024); Choi et al. (2025); Chen et al. (2024); Choi et al. (2024); Gao N. et al. (2024); Lin et al. (2025); Wang S. et al. (2025); Huang and Song (2025); Kuo et al. (2025)
Inter-user synchronization	Studies the synchronization of users' video feeds in multi-user sessions and examines how the inter-user synchronization interval affects UX and the overall system	4	Ha Huynh et al. (2024); Novotny et al. (2020); Ran et al. (2020); Ren et al. (2022)
Fairness	Evaluates QoS/QoE equity among multiple users under shared constraints	3	Liu et al. (2023); Wang S. et al. (2020); Badnava et al. (2024)
Global precision	Assesses spatial/interaction correctness (e.g., localization or ID matching precision)	2	Cai et al. (2022); Chen et al. (2023)
Global bandwidth	Investigates the bandwidth allocation problem across the entire system, such as optimizing the system based on the global bandwidth savings ratio	2	Cao et al. (2019); Mahmoud et al. (2024)

users' high requirements for image quality and smoothness, thereby enhancing the immersive experience.

Five papers mentioned a network technology for wireless optimization. These studies solved the contention problem when multiple users share the channel by optimizing wireless network technology, which effectively reduces latency.

5.4.1 Traffic flow control

In multi-user network environments, multiple devices access and transmit data simultaneously, particularly in VR and XR applications, where each user must download a complete 3D scene from the server. The network usually experiences bandwidth congestion and transmission delays when many users load the XR environment simultaneously. To address this problem, several studies have proposed various optimization strategies to improve traffic flow efficiency.

First, Alhilal et al. (2023) proposed a solution to mitigate the impact of 3D model loading on the network performance by using low-poly modeling technology. They reduced the HKUST campus model from 126 MB to 45 MB, a 64% reduction in data volume, and the model loading time was reduced by more than 50% in actual tests. For traffic scheduling optimization, Santos et al. (2023) focused on critical decisions on Service Function Chaining (SFC) routing and deployment, supporting lifecycle management across multiple edge nodes. Santos et al. (2021) applied SFC and formulated a Mixed-Integer Linear Programming (MILP) model to calculate the optimal VR service chain path, ensuring balanced traffic distribution and reducing hotspot congestion. Santos et al. (2022) extended the SFC framework to optimize traffic through a QoS classification mechanism. In this approach, interactive data was assigned the highest priority, background data a lower priority, and video streams adopted a dynamic bitrate adjustment mechanism. Similarly, Casparsen et al. (2025) utilized the slicing mechanism and real-time resource scheduling of Open RAN networks to optimize the network architecture. Wang Z. et al. (2024) focused on optimizing traffic management in XR through transmission priority control. They used DRL to dynamically allocate wireless resources and optimize bandwidth utilization based on QoS requirements. Similarly, Sandeepa et al. (2024) proposed a low-latency communication scheme based on Message Queuing Telemetry Transport (MQTT), using a publish/subscribe (Pub/Sub) model to efficiently manage real-time data. In addition, Ran et al. (2020) and Badnava et al. (2024) employed adaptive AR communication technologies to reduce bandwidth consumption. Yu H. et al. (2023) combined AI-based scheduling with 6G deterministic networking (DetNet) to propose an intelligent traffic management solution for MU-XR resource optimization.

To improve XR network performance through protocol optimization, Bakri et al. (2015) were the first to apply

TABLE 5 Functional roles of network technologies supporting MU-XR systems (traffic flow, resource coordination, edge-assisted performance optimization, etc.).

Catalog	Description	Count	Paper
Traffic flow control	Focuses on optimizing routing, congestion control, and data distribution	13	Alhilal et al. (2023); Bakri et al. (2015); Ha Huynh et al. (2024); Ran et al. (2020); Sandeepa et al. (2024); Santos et al. (2021), Santos et al. (2022); Wang Z. et al. (2024); Yu et al. (2023a); Casparsen et al. (2025); Badnava et al. (2024); Okafor et al. (2024); Santos et al. (2023)
Resource coordination	Coordinates storage, compute, and network resources across nodes to enhance system scalability	13	Aloqaily et al. (2021); Cai et al. (2023); Han et al. (2023); Huang et al. (2024); Nath and Wu (2020); Novotny et al. (2020); Sonkoly et al. (2023); Van Damme et al. (2023); Wang S. et al. (2020); Yu et al. (2023b); Zhang et al. (2022); Choi et al. (2024); Wang C. et al. (2020)
Edge-assisted performance optimization	Utilizes edge computing to offload intensive tasks and reduce latency, enhancing real-time interactions	12	Apicharttrisorn et al. (2020); Cai et al. (2022); Chen and Guo (2023); Chen et al. (2023); Liu et al. (2023); Okamoto et al. (2023); Ren et al. (2022); Yang et al. (2024); Zhong et al. (2023); Choi et al. (2025); Fowler and Souihi (2025); Xu et al. (2025)
Video stream optimization	Improves video transmission and rendering strategies to reduce bandwidth and latency costs	10	Cao et al. (2019); Chen et al. (2022); Kamarianakis et al. (2021); Mahmoud et al. (2024); Gao N. et al. (2024); Tseng et al. (2024); Lin et al. (2025); Zhao et al. (2025); Wang P. et al. (2025); Chakareski and Mastronarde (2020)
Wireless optimization	Enhances multi-user access schemes for higher throughput and lower interference	5	Ahn et al. (2018); Yu et al. (2023b); Chen et al. (2024); Huang and Song (2025); Kuo et al. (2025)

TABLE 6 Summary of AI functional roles in MU-XR systems, including resource allocation, behavior prediction, and adaptive decision-making.

Catalog	Description	Count	Paper
Network resource allocation	Predicting computational and network demands with AI to enable adaptive resource allocation	10	Aloqaily et al. (2021); Chen et al. (2022); Wang Z. et al. (2024); Yang et al. (2024); Yu et al. (2023a); Yu et al. (2023b); Yu et al. (2023b); Casparsen et al. (2025); Huang and Song (2025); Wang C. et al. (2020)
User identification and behavior prediction	Using AI models to predict user gaze, hand, motionetc.	5	Cai et al. (2022); Cao et al. (2019); Ren et al. (2022); Zhang et al. (2022); Wang S. et al. (2025)
Task adaptive decision making	Applying AI-driven decision models to select optimal strategies or models dynamically under varying network conditions	6	Huang et al. (2024); Sonkoly et al. (2023); Badnava et al. (2024); Chen et al. (2024); Fowler and Souihi (2025); Xu et al. (2025)

HTTP/2 and QUIC in multi-user virtual worlds (MUVWs) to optimize the efficiency of the protocol in 3D web environments. Ha Huynh et al. (2024) used DRL to automatically select the most suitable transport protocol (TCP, QUIC, SCTP) to enable intelligent multi-user data synchronization. Okafor et al. (2024) optimized network-layer packet processing using the extended Berkeley Packet Filter (eBPF) and Express Data Path (XDP).

5.4.2 Resource coordination

Resource coordination was a widely used decentralized solution in MU-XR systems. Aloqaily et al. (2021) used blockchain to construct a distributed data-storage system and improve system security by ensuring only trusted nodes could join the network through a validation mechanism. This method improved resource utilization by about 46% and significantly reduced task response time. Similarly, Yu et al. (2023b) proposed a metaverse architecture that integrates Web3, MEC, and 6G based on blockchain, in which virtual service providers (VSPs) used smart contracts to achieve transparent channel allocation management.

Several studies leveraged multiple nodes to enable distributed task processing. Han et al. (2023), Nath and Wu (2020), and Sonkoly et al. (2023) proposed architectures to dynamically offload computing tasks based on different computing requirements and network conditions. Building on these techniques, Novotny et al. (2020) proposed the *Unity + Mirror* framework hybrid P2P-server architecture. The Mirror framework allowed servers and P2P devices to share synchronization while also providing a P2P direct connection mode.

To further reduce the dependence on the network, Van Damme et al. (2023) proposed a decentralized method for interactive synchronization that abandons the traditional server-client architecture. Similarly, Zhang et al. (2022) proposed a cooperative local caching scheme that prioritizes using local caches, directly reading local data when the cache is hit, and only requesting edge computation when the cache is not hit, thereby reducing the load on edge servers.

In task allocation, the system used an algorithm to assign data processing tasks to idle or lightly loaded nodes to reduce system load. For example, Wang C. et al. (2020) investigated server assignment and resource management for task chains in multi-user MAR, addressing a cross-node coordination problem. Cai et al. (2023) proposed a distributed queue system that could track and classify the replication status of data packets to achieve efficient data forwarding and processing. Similarly, Choi et al. (2024) proposed a dynamic monitoring and elastic deployment system. In addition, Huang et al. (2024) proposed a rendering mechanism in which a controller server manages device status and assigns tasks, coordinating multiple secondary servers to execute rendering tasks. Finally, Wang S. et al. (2020) studied a tile-based chunked transmission mechanism to dynamically distribute streaming tiles.

5.4.3 Edge-assisted performance optimization

Edge-assisted network optimization combines edge computing and network optimization, significantly improving network performance in multi-user, high-bandwidth applications.

Apicharttrisorn et al. (2020); Fowler and Souihi (2025) leveraged edge resources for real-time decision-making, offloading, and scheduling to enhance interaction performance. Cai et al. (2022) proposed a real-time ID-aware multi-device visual interaction system based on edge computing. In addition, Ren et al. (2022); Xu et al. (2025) used edge computing integrated with advanced network technologies to achieve multi-user interaction and computational optimization. Similarly, Chen and Guo (2023) proposed monitoring the load of computing resources and network conditions in real time and dynamically adjusting the task offload path. Chen et al. (2023) developed a framework for simultaneous localization and mapping (SLAM) of multi-user AR devices using edge computing. Okamoto et al. (2023) proposed caching and processing videos on edge servers, transmitting the tiles watched by the user based on the user's field of view (FoV), while Zhong et al. (2023) employed edge computing to enhance the transmission efficiency of multicast technology. Yang et al. (2024) implemented neural network interpolation (NFI) on edge servers for the first time.

Some studies combined edge computing with the cloud. Choi et al. (2025) studied adaptive task allocation and edge-cloud collaborative scheduling. In addition, Liu et al. (2023) proposed an edge-side compression control and cloud-side synchronization decisions to optimize frame synchronization. The study explored combining edge-side compression with cloud-side synchronization optimization had been combined to optimize QoE fairness.

5.4.4 Video stream optimization

Cao et al. (2019) proposed a 360° video streaming optimization system that combines joint viewport optimization with machine learning to predict users' head movements. They defined a joint viewport to estimate overlapping visual scenes among users. For shared areas, the server sent data simultaneously by multicasting instead of unicasting. Similarly, Chen et al. (2022); Gao N. et al. (2024); Zhao et al. (2025) also suggested that predicting the user FoV could reduce bandwidth consumption.

Mahmoud et al. (2024); Tseng et al. (2024); Lin et al. (2025); Chakareski and Mastronarde (2020); Wang S. et al. (2025), all proposed tile-based video streaming optimization strategies. They studied adaptive bandwidth optimization and improvements in transmission sequencing. The research of Mahmoud et al. (2024) and Tseng et al. (2024) also emphasized the importance of focusing on user satisfaction to optimize the system.

Finally, Kamarianakis et al. (2021) optimized the motion data stream in a VR collaborative environment using Dual Quaternions and Multivectors for efficient interpolation. Quaternions typically require only four dimensions, compared to nine in traditional matrices, thereby reducing computational load and storage requirements.

5.4.5 Wireless optimization

Five papers studied optimization of the WLAN physical layer to improve wireless VR transmission throughput and optimization of multi-user channel access. Ahn et al. (2018) proposed an optimized architecture for wireless VR networks that combined 802.11 ad/ay/ax to improve the transmission efficiency and stability of MU-XR interactions. In addition, Yu et al. (2023b) optimized the wireless VR transmission scheme by proposing a multi-user architecture combining NOMA, computation offloading, and DRL channel optimization. NOMA used power-domain multiplexing technology to enable multiple users to transmit data simultaneously on the same channel. It separated different user signals through the Successive Interference Cancellation (SIC) technique to improve spectrum utilization. Finally, Chen et al. (2024) optimized the wireless channel scheduling algorithm to achieve prediction-driven preemptive resource allocation. In wireless resource allocation, Huang and Song (2025) investigated OFDMA subchannel and rate-level resource allocation, while Kuo et al. (2025) focused on multi-user resource allocation and access schemes for Massive MIMO downlink systems.

In summary, these five categories form a hierarchical framework of network functions that bridge the communication, computation, and coordination layers of XR systems. Network functions serve as the management layer, coordinating transmission, computation, and resource allocation to ensure smooth and fair interaction. Studies on traffic flow control have focused on adaptive routing, congestion avoidance, and QoS-based prioritization, which ensure that throughput remains stable when multiple users share the same network resources. Resource coordination research focuses on distributed and software-defined architectures that dynamically allocate bandwidth and computation across nodes to balance the load. Edge-assisted frameworks take this coordination by offloading rendering and computation tasks closer to end-users, thus reducing transmission delay. Video stream optimization aims to strike a trade-off between visual quality and bandwidth expenditure by means of adaptive bitrate and tile-based segmentation. In all these functions, there is a common weakness in evaluation and control logic. Most mechanisms are optimized for either throughput or latency, rather than as a unified orchestration. However, not many studies have combined resource scheduling, flow control, and content adaptation into one management framework. Further efforts in the future should aim at integrated network-function management architectures consisting of SDN, edge intelligence, and real-time feedback mechanisms to facilitate adaptive resource orchestration of heterogeneous XR services.

5.5 Area 4: Roles of AI

While analyzing the 56 papers, we found that 21 papers have shown the application of AI techniques, with the corresponding categories and representative studies summarized in Table 6. Although this review primarily focuses on system-level mechanisms of networked MU-XRs rather than algorithmic implementations, it is noticeable that AI has shown strong performance in dynamic resource allocation and task decision-making. Therefore, we provide an overview of the key types of AI functions without a detailed explanation of how each algorithm works.

We identified three main aspects of the AI involved. Firstly, federated learning (FL) techniques, DRL, and its variants (such as Double DQN and Dueling DQN) are applied to dynamically predict network conditions, thereby optimizing the allocation of computing and communication resources and improving load balancing and system efficiency in multi-user network environments. 10 papers have shown such AI applications in network resource allocation. Second, we found that using AI to model and predict user behavior is another direction, supported by five papers. Specifically, Convolutional Neural Networks (CNNs), Long Short-Term Memory (LSTM) networks, and statistical models (such as ARIMA and logistic regression) are used to model user head movements, pose matching, and interaction behaviors. In addition, users' FOV and body behaviors can also be predicted to improve UX and optimize bandwidth allocation in XR systems. Third, AI has also demonstrated support for offloading strategies in decision-making tasks within XR systems, supported by six papers. Techniques such as Monte Carlo Tree Search (MCTS) and Deep Deterministic Policy Gradient (DDPG) have been used in these applications.

5.5.1 Network resource allocation

In predicting and allocating resources, AI technology predicts real-time network conditions and then dynamically distributes resources, optimizing bandwidth utilization and network latency that affects user interactions. Aloqaily et al. (2021) used an FL model to predict and allocate future loads in edge computing scenarios. Similarly, Wang C. et al. (2020) adopted AI to predict network-related computational requirements, such as task computational complexity and processing time. In addition, Yu H. et al. (2023) employed double DQN and continuous learning (CL) to address complex multi-user network environments and user requirements. The paper proposed that using a combination of these two techniques could reduce the overestimation problem of the original Q-learning algorithm. Yang et al. (2024) used Real-time Intermediate Flow Estimation (RIFE), which generates high-quality intermediate frames and quickly encodes and compresses them to improve the frame rate performance of VR videos. To address multi-dimensional resource allocation, Wang Z. et al. (2024); Huang and Song (2025) proposed a DRL-based method. They used multi-agent and multi-network structural methods to solve the allocation problem. In addition, Chen et al. (2022) used the Proximal Policy Optimization (PPO) algorithm to train the global adaptive bitrate allocation of videos. The system condition (bandwidth, buffering, and user FoV prediction) and edge cache information were integrated to automatically infer the quality level of the following video segments, ensuring visual quality and consistency among users. Furthermore, considering the multiple QoS requirements of a MU-XR network, Yu et al. (2023a) and Casparsen et al. (2025) proposed QoS-aware MU-XR frameworks. They used multiple DRL algorithms and Recurrent Neural Networks (RNN) to achieve multi-channel allocation, reducing latency and enhancing user satisfaction. In the following research (Yu et al., 2023b), the authors proposed the Hybrid Reward PPO algorithm, which was based on the PPO algorithm and decomposed the multi-user, high-dimensional objectives into multiple sub-objective rewards to better accommodate the diverse user requirements in terms of latency, energy consumption, and other aspects. HRPPO achieved a 45% higher cumulative reward in numerical simulations and converged faster than PPO.

5.5.2 User identification and behavior prediction

AI is widely used in user behavior recognition and prediction. Cai et al. (2022) used the CNN model Mask R-CNN to recognize the human body and extract keypoint data to capture real-time pose correspondence between users and the camera. For behavior prediction, Cao et al. (2019); Wang S. et al. (2025) used an AI model to train and predict head movements. This method accurately predicted the user FoV, thereby optimizing bandwidth utilization. In addition, Ren et al. (2022) applied the Autoregressive Integrated Moving Average (ARIMA) and Logistic Regression models for motion prediction to select appropriate keyframes. The study also analyzed the feasibility of deep learning methods such as Recurrent Neural Networks (RNN). Zhang et al. (2022) proposed the use of probabilistic statistical models, such as the Bernoulli Mixture Model (BMM) and Gaussian Mixture Model (GMM), to encode image/scene features using Fisher Vector (FV) feature encoding, in order to optimize object recognition on mobile devices and improve the processing efficiency of multi-user AR applications.

5.5.3 Task adaptive decision making

In dynamic networks, decisions are often determined in real-time as network conditions change. In task decision-making, Huang et al. (2024) proposed the use of a Generative Adversarial Network (GAN) model to compress or reconstruct holographic video in real time and perform inference computing and decision-making at the edge/cloud to reduce transmitted data and improve rendering efficiency. Similarly, Chen et al. (2024) proposed a task offloading optimization algorithm based on the Multi-Agent Deep Deterministic Policy Gradient (MADDPG) framework to enhance resource utilization. In addition, Sonkoly et al. (2023) used neural networks to select the best-performing output when multiple SLAM algorithms were running in parallel. Badnava et al.

(2024) proposed a DRL algorithm in which the agent made joint decisions on video bitrate selection based on video playback statistics (such as historical throughput, transmission time, and rendering time) and video information (e.g., remaining frame count). Finally, Fowler and Souihi (2025); Xu et al. (2025) investigated how to make more efficient decisions under resource constraints to enable more effective resource allocation.

AI techniques can be used to improve the adaptiveness of MU-XR systems in three primary areas: NRM, user behavior prediction, and task-level decision-making. In network management, AI predicts network congestion and bandwidth fluctuations, enabling proactive resource distribution and improving system stability and responsiveness. In predicting behaviors, AI processes gaze, motion, and physiological information to infer user intent, enabling pre-rendering and reducing latency. At the task level, AI selects optimal strategies for offloading and rendering based on real-time network and device conditions. In general, AI is increasingly transforming the operation of networked XR systems from reactive response to proactive prediction and deployment. These functions should be incorporated into intelligent cross-layer control structures in future research.

6 Discussion

This section interprets the findings of the systematic review by first identifying the *challenges and opportunities* in networked MU-XR systems, and then presenting *a refined system model* to outline multiple directions for future work.

6.1 Challenges and opportunities

This study took an application-oriented perspective to systematically examine NRM requirements in MU-XR. By analyzing the reviewed work across application types, immersive techniques, and four research areas, we identified the key challenges and emerging opportunities in networked MU-XR systems.

6.1.1 Challenges

Lack of UX-related experimentation remains a key challenge. Only 6 of the 56 reviewed papers conduct genuine user studies; the remainder largely rely on theoretical analysis or simulation. Human perception of network impairments is non-linear and highly sensitive to contextual and environmental changes (Vlahovic et al., 2022). Therefore, in networked MU-XR scenarios, objective network measurements alone are insufficient to accurately assess overall system effectiveness. Although we found that numerous papers in the XR field include user studies, far fewer do so when NRM technologies are involved. We believe this is due to the difficulty of setting up MU-XR architectures in heterogeneous networks and the challenge of controlling multi-user network conditions, which leads to low experimental repeatability.

QoE-related metrics for multi-user scenarios remain limited. The current evaluations still focus on basic network performance measures, particularly latency (25/56) and throughput (17/56). In contrast, QoE fairness is addressed in only 3 of 56 studies. Because we are studying the intersection field of multi-user XR and network, developing standardized QoE metrics or metric sets for multi-user scenarios would improve cross-study comparability and support stronger cumulative evidence.

Dynamic synchronization remains difficult. Only a few studies in the literature have addressed this. The experiments typically involve two-person groups or small groups rather than large-scale sessions. This is largely because current network capacity is still limited (Huang et al., 2023), and users must optimize their algorithms or models under constrained resources, so synchronization cannot be treated as the primary objective, which makes global, large-scale synchronization hard to implement and evaluate in practice.

Sensing modalities were limited, and full-body tracking inputs were rare. The majority of the research focused on head pose (as used in 360 video delivery and VR interaction). Moreover, headsets are usually paired with head tracking and cameras, and most VR systems include controllers. In comparison, body tracking and LiDAR are more likely to suffer occlusion and tracking loss in a multi-user environment. Additionally, heterogeneous hardware platforms vary in sampling rates and other signal properties, which complicates data processing. This complicates cross-system analysis and increases the implementation barrier.

6.1.2 Opportunities for advancing MU-XR networking

Therefore, we consider that beyond 5G (B5G) networks and next-generation 6G networks are a breakthrough for XR, where 5G networks have much higher bandwidth and lower latency modes, all of which directly match XR requirements (Hazarika and Rahmati, 2023). Researchers have already experimented with MU-XR on 5G infrastructure and reported improvements in capacity and latency (Peñaherrera-Pulla et al., 2024). In the future, the concepts of 6G, like deterministic networking and sophisticated network slicing, can ensure the network's end-to-end latency constraints and dedicated XR slices (Nasrallah et al., 2019). For example, one study integrated AI with a 6G deterministic network to control XR traffic intelligently, suggesting that future networks may provide hard QoS guarantees for XR traffic. At the same time, the ongoing development of edge/cloud computing as a network architecture provides some opportunities. MU-XR can offload computation and reduce latency by spreading servers further among users (Lee and Shin, 2023). Moreover, the network can be more adaptive, with predictive analytics that allocate resources before the network is congested or slows down due to increased latency (Casasnovas et al., 2024; Kumar and Raubal, 2021).

Another opportunity lies in the rising attention and investment in the XR industry. This implies that greater effort will likely be devoted to immersion techniques, such as developing lighter HMDs and other wearable devices. More specialized multimodal sensors for capturing user information are also expected to mature, enabling XR to expand into advanced application domains, including collaborative surgery, smart driving, and digital smart cities. In turn, these shifts in application requirements will further drive innovation and change in XR network design.

Overall, MU-XR networking still faces challenges in UX evaluation, QoE-related metrics, synchronization, and sensor applications. However, the potential of B5G and 6G, edge computing, and smart NRM offers a path to eliminate these issues. The research community is actively tackling these issues, and it is hoped that with further development, networks will be able to meet the high demands of large-scale MU-XR experiences. Additionally, the development of multimodal sensors will also drive the diversification of MU-XR applications.

6.2 Refined multi-user IR model

Building on the general IR system model and the core parts (Figure 1) outlined in Section 2, this section provides additional details to the model as updated in the new version (Figure 8) based on the findings from our SLR and our own insights. We follow the scope of 3GPP TR 26.928 3GPP (2022), which frames XR systems in terms of use cases, devices/form factors, client–network architectures and media processing, delivery procedures, and KPI/QoE metrics. We treat user topology/interaction and AI support functions as extensions, since TR 26.928 does not define them as standalone scope items, but they are critical for multi-user synchronization and practical optimization under limited network capacity.

6.2.1 XR applications

As application scenarios expand, users have increasing expectations for virtual environments (Wang et al., 2023). Focusing on user-centered scenarios, XR application requirements will increase and become multi-user instead of single-user. A single-user VR game or training simulation can easily run on a local computer, with the UX limited to interactions with the system. However, co-located collaborative XR (where two or more users occupy the same area) must be closely synchronized in real-time with the content and actions of each user (Sonkoly et al., 2024). Interaction among users significantly enhances immersion and enables many collaborative tasks within XR environments. However, distant interactions, such as global-scale multiplayer games or remote surgical procedures involving multiple participants, remain technically challenging. As the scale of distributed XR activities increases, the complexity of the underlying network and compute architecture also increases exponentially (Hatami et al., 2024). These massive-scale scenarios require high-precision, highly detailed scene rendering, as well as high-speed data exchange between hundreds of users.

6.2.2 User topology and interaction

Group size shapes the user experience during XR interactions. In the case of single-user deployment (i.e., playing a single-player VR game at home or participating in individualized immersive training), the system is mainly concerned with responding to and providing rich content around one individual. These usually allow people to move at their own pace without worrying about fitting in with others. However, in the context of MU-XR applications, new requirements for interaction arise. For example, co-located users may participate in an AR cooking game or a collaborative virtual design system in the same physical location, which enhances shared immersion and teamwork. These experiences enable richer communication, such as movement and gaze tracking, that are needed for multi-user collaboration. Moreover, in a distributed multi-user group, users from different geographic locations interact in the same virtual environment. This scenario enables people who are physically separated to work but can socialize together. On the other hand, it can reduce travel and provide access to shared experiences on a larger scale. For example, a team collaborating on a task can train together across different time zones, greatly improving both work efficiency and cost-effectiveness.

6.2.3 Immersion

To support multi-user interaction, immersive technologies are employed. AR, VR, and MR devices combine inputs from cameras, depth sensors, motion trackers, and more to capture each user's perspective (Kourtesis, 2024). In multi-user scenarios, these inputs must be integrated into a common processing pipeline. For example, when multiple AR users collaborate in the same physical space, they require a shared coordinate system and a global map of the environment to ensure that virtual objects remain aligned for everyone (Dhakal et al., 2022; Vidal-Balea et al., 2025). Different applications require different sensors to function effectively. We have reviewed the literature to establish that eye trackers, hand trackers, and haptic sensors are commonly used to capture user behavior in terms of interaction. Another application of physiological sensors involves the Galvanic Skin Response (GSR) and Brain-Computer Interface (BCI) systems, which can be used to test the emotional and cognitive responses of users. For example, they can be applied to identify stress levels or changes in attention during an XR experience. This broadens the scope and granularity of user data collection, and the analysis of user behavior begins to become fine-grained.

6.2.4 Networking-compute continuum

These requirements give rise to a networking–compute continuum. Computation is dynamically distributed across devices, edge servers, and cloud data centers based on the scale of interactions and sensitivity to latency (Wang and Zhao, 2022; Xu et al., 2025). Simpler XR applications run entirely on a powerful headset or local computer; however, more complex experiences offload heavy tasks to nearby or cloud resources. During co-located multi-user scenarios, shared engines (such as world map and physics) can be hosted on on-premises edge nodes, allowing a group of users to access them with minimal delays. In a distributed MU-XR scenario, cloud servers are utilized to store a global state and perform large-scale computations (e.g., rendering large virtual worlds or executing complex simulations for a large number of users simultaneously). Offloading to remote clouds, however, comes with network latency. An all-cloud solution would not be able to support real-time XR performance without high-performance, low-latency networks. As a result, today's XR systems employ hybrid architectures, with latency-sensitive interactions (such as motion tracking and viewpoint rendering) residing on-device or at the closest edge. At the same

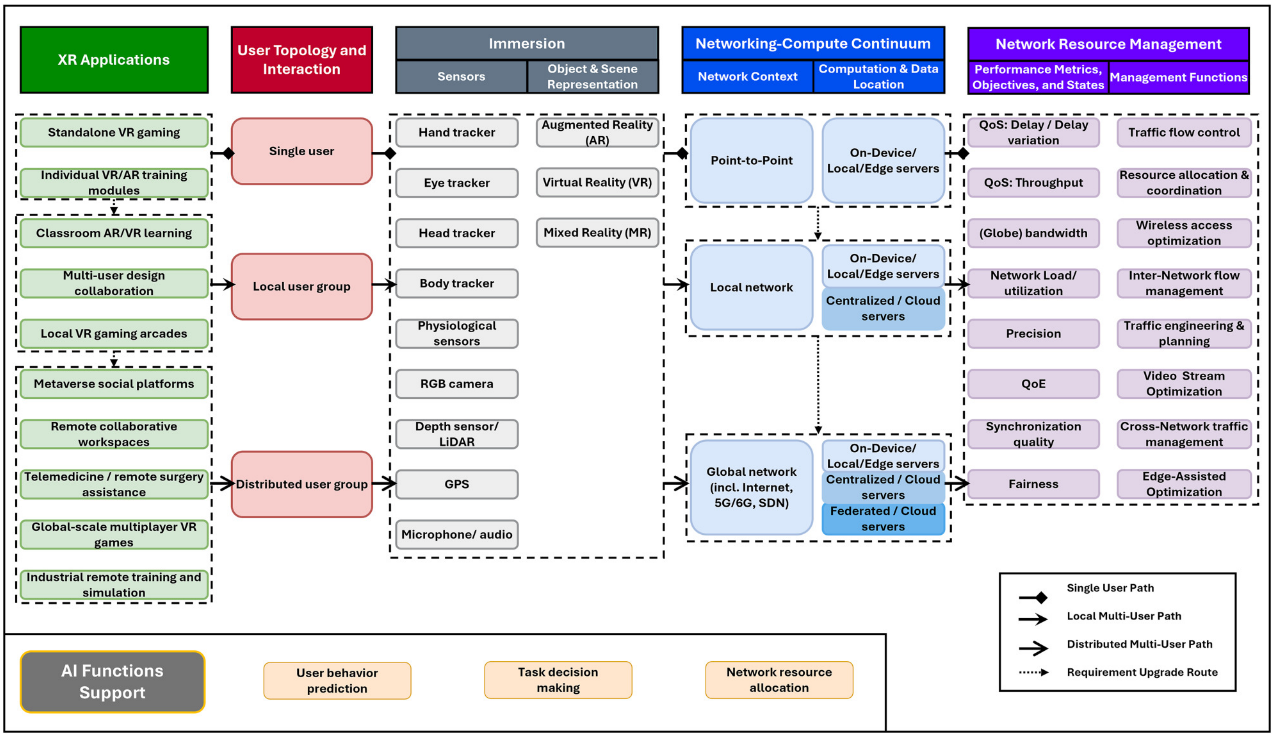

FIGURE 8
Refined multi-user IR system model.

time, less time-sensitive workloads (massive scene updates and AI-driven content generation) are sent to the cloud. This continuum is arranged in a manner that ensures every function operates at the right level, maintaining a balance between load and end-to-end latency at acceptable levels (Chen and Jiang, 2024). However, it is not easy to coordinate devices, multiple edge servers, and the cloud simultaneously. Even at scale, distributed XR architectures struggle to maintain synchronization and cope with network delays, highlighting the core problem of ensuring real-time performance across a distributed infrastructure.

6.2.5 NRM

Therefore, MU-XR depends on NRM functions. It implements QoS requirements using measures and control capabilities that are customized for immersive media. Important QoS factors are end-to-end latency, throughput of high-rate streams, and reliability (packet loss tolerance is very low with interactive XR). To achieve these network functions, such as traffic engineering and planning, cross-layer and cross-network management are required. For example, XR control traffic (e.g., head-tracking input) can be marked for ultra-low latency delivery, and large video frames can be compressed and potentially transmitted with lower priority to prevent channel congestion. The specialized synchronization protocols ensure that updates generated by different users or servers remain temporally synchronized, thereby preventing one user's view from falling behind that of another (Gao B. et al., 2024). The management of network resources requires very fine-grained and cross-layer optimizations. An example is 5G network slicing, which can assign a slice with guaranteed bandwidth and latency to an XR application (Abood et al., 2024). Additionally, edge servers can multicast shared data to local users, thereby avoiding duplicate traffic. Lastly, this stack features AI functions that help in prediction, decision-making, and resource optimization. Machine learning models are capable of forecasting network conditions and traffic dynamics, enabling the system to adjust proactively (Baidya and Moh, 2024). Real-time computational decisions, including how to distribute computational tasks among devices, edges, and the cloud, can also be made using AI-driven orchestration and optimized based on the criteria of lowest latency and bandwidth consumption. In summary, AI-assisted NRM is becoming one of the main solutions to ensure the high QoS required for achieving coordinated real-time XR experiences.

Overall, it is clear that managing network resources for MU-XR systems is fundamentally different from managing those for single-user XR. In a single-user case, the system can dedicate all network and computing resources to ensure a smooth experience, which primarily involves reducing the latency between the user and the dedicated server or computer. However, once multiple users communicate within the same virtual environment, the network must not only be capable of supporting much larger amounts of data, but also be able to synchronize these streams to ensure that all participants are synchronized. This implies that network delays or jitter impact not only an individual user's

10.3389/frcmn.2026.1761635

experience but also the consistency of UX. In addition, the whole system must ensure that the quality of a high-bandwidth stream for one user does not adversely affect the quality of other users. Moreover, MU-XR can involve geographically dispersed users, making it impossible to establish a local, low-latency network for everyone. To ensure real-time performance, the system must rely on edge/cloud infrastructure, with the network carefully optimized for minimal latency. This means that systems must not only reduce latency and maximize throughput, but also synchronize performance across many users while ensuring consistency and fairness. Single-user XR solutions cannot be readily scaled to a multi-user environment without significant network-centric improvements. In this way, MU-XR pushes networking technology to its limits, requiring solutions that extend well beyond those of single-user XR systems. The key difference is that multi-user interactions are more complex, and this aspect means that XR is no longer merely an individual real-time streaming issue but a distributed systems issue, in which the intelligent use of network resources is central to creating an immersive group experience.

We designed the partitioning to be as comprehensive as possible based on the reviewed evidence and our insights. However, it may still miss emerging MU-XR components (e.g., new sensing modalities, application patterns, or NRM functions) that remain understudied or not yet standardized. Future work should refine and extend the partitioning as the field matures and new evidence becomes available.

7 Conclusion and future work

This SLR identified and analyzed 56 studies on multi-user XR (MU-XR) systems. The studies were examined through four research questions addressing their relationship to NRM in MU-XR systems. The reviewed applications operate under stringent network-performance and resource constraints. As the immersion requirement increases and MU-XR remains a dominant focus, NRM must meet more specific and demanding requirements, especially for large-scale distributed sessions. Publication counts dipped slightly in recent years but remain high, indicating a strong research interest in MU-XR systems.

We further analyzed UX evaluation, network metrics, NRM functions, and AI functions, which indicate the connection between the network and the user-side. A key finding of this paper is that further research is needed to assess and better connect multi-user XR experiences with network-level NRM and QoS mechanisms. Although multi-user experiences and network performance are closely related, only 12.8% of the surveyed articles evaluate user experience when discussing NRM. We conclude that this gap needs to be closed to enable efficient resource management for highly immersive MU-XR systems. In addition, critical network metrics that impact UX include latency, jitter, synchronization, and stability. These metrics limit the scale of MU-XR user groups and the immersiveness of multi-user interaction, thereby influencing the development of XR applications. The use of B5G and 6G, edge computing and cloud computing, and intelligent NRM techniques enables the development of distributed, real-time XR systems.

Future research should focus on developing standardized evaluation methods for multi-user experience and validating them through more user studies. It should also examine how a wider range of immersive techniques can be applied across different scenarios. This introduces new demands on uplink resource management, since high-immersion multi-user applications require efficient transmission of user feedback data. In addition, future work should develop NRM algorithms and platforms that prioritize synchronization for specific user groups or for multi-user settings with particular immersion requirements. Finally, this paper proposes a multi-user XR system model (Figure 8) that emphasizes immersiveness-related features and aims to derive the boundary of NRM requirements from XR application patterns. At this stage, however, there is insufficient empirical UX evaluation evidence to support a one-to-one mapping between application patterns and NRM requirements. For this reason, the model should be continuously refined and updated as more empirical results become available. We argue that such iterative refinement can provide more practical guidance for aligning future intelligent resource management mechanisms with multi-user XR applications.

Data availability statement

The data analyzed in this study are derived from publicly available literature retrieved from Scopus, IEEE Xplore, and the ACM Digital Library. No specific datasets were generated or deposited in a public repository for this study.

Author contributions

BQ: Conceptualization, Data curation, Investigation, Methodology, Validation, Visualization, Writing – original draft, Writing – review and editing. KT: Conceptualization, Investigation, Methodology, Supervision, Visualization, Writing – review and editing. YH: Conceptualization, Investigation, Methodology, Supervision, Writing – review and editing.

Funding

The author(s) declared that financial support was received for this work and/or its publication. This work was supported and funded by the Knowledge Foundation, Sweden, through the Human-Centered Intelligent Realities (HINTS) Profile Project (contract 20220068).

Acknowledgements

The authors would like to express their sincere gratitude to the Human-Centered Intelligent Realities (HINTS) Profile Project for its support of this work.

Conflict of interest

The author(s) declared that this work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declared that generative AI was not used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Abood, M. S., Wang, H., Virdee, B. S., He, D., Fathy, M., Yusuf, A. A., et al. (2024). Improved 5G network slicing for enhanced QoS against attack in SDN environment using deep learning. *IET Commun.* 18, 759–777. doi:10.1049/cmu2.12735

Ahn, J., Kim, Y. Y., and Kim, R. Y. (2018). Virtual reality-wireless local area network: wireless connection-oriented virtual reality architecture for next-generation virtual reality devices. *Appl. Sci. Switz.* 8, 43. doi:10.3390/app8010043

Alhilal, A., Shatilov, K., Tyson, G., Braud, T., and Hui, P. (2023). "Network traffic in the metaverse: the case of social VR," in *Proceedings - 2023 IEEE 43rd international conference on distributed computing systems workshops*, 109–114.

Aloqaily, M., Bouachir, O., and Ridhawi, I. (2021). "Blockchain and FL-based network resource management for interactive immersive services," in *Proceedings - IEEE global communications conference* (GLOBECOM), 1–10. doi:10.1109/GLOBECOM46510.2021.9685091

Alsader, M., Barakabitze, A. A., and Mkwawa, I.-H. (2025). QoE-Driven adaptive video streaming: architectures, techniques, and future research challenges toward 6G networks. *IEEE Access* 13, 157408–157441. doi:10.1109/access.2025.3597058

Apicharttrisorn, K., Balasubramanian, B., Chen, J., Sivaraj, R., Tsai, Y.-Z., Jana, R., et al. (2020). "Characterization of multi-user augmented reality over cellular networks," in *Annual IEEE communications society conference on sensor, mesh and Ad hoc communications and networks workshops*.

Badnava, B., Chakareski, J., and Hashemi, M. (2024). Multi-task decision-making for multi-user 360° video processing over wireless networks. *Proc. Int. Conf. Multimedia Inf. Process. Retr.* MIPR, 294–300. doi:10.1109/MIPR62202.2024.00054

Baidya, T., and Moh, S. (2024). Comprehensive survey on resource allocation for edge-computing-enabled metaverse. *Comput. Sci. Rev.* 54, 100680. doi:10.1016/j.cosrev.2024.100680

Bakri, H., Allison, C., Miller, A., and Oliver, I. (2015). HTTP/2 and QUIC for virtual worlds and the 3D web? *Procedia Comput. Sci.* 56, 242–251. doi:10.1016/j.procs.2015.07.204

Bayro, A., Moon, H., Ghasemi, Y., Jeong, H., and Lee, J. Y. (2025). Object manipulation in physically constrained workplaces: Remote collaboration with extended reality. *IISE Trans. Occup. Ergonomics Hum. Factors* 13, 177–190. doi:10.1080/24725838.2025.2484731

Borhani, Z., Sharma, P., and Ortega, F. R. (2024). Survey of annotations in extended reality systems. *IEEE Trans. Vis. Comput. Graph.* 30, 5074–5096. doi:10.1109/TVCG.2023.3288869

Cai, X., Yang, Z., Dong, L., Ma, Q., Miao, X., and Liu, Z. (2022). "DiTing: edge assisted real-time ID-aware visual interaction for multi-user augmented reality," in *2022 IEEE 28th international conference on parallel and distributed systems (ICPADS)*, 737–744.

Cai, Y., Llorca, J., Tulino, A. M., and Molisch, A. F. (2023). Decentralized control of distributed cloud networks with generalized network flows. *IEEE Trans. Commun.* 71, 256–268. doi:10.1109/tcomm.2022.3225186

Cao, H., Lu, J., and Zong, N. (2019). "A multi-user 360-video streaming system for wireless network," in *Proceedings - vrcai 2019: 17th ACM SIGGRAPH international conference on virtual-reality continuum and its applications in industry*, 123–130.

Casasnovas, M., Michaelides, C., Carrascosa-Zamacois, M., and Bellalta, B. (2024). Experimental evaluation of interactive edge/cloud virtual reality gaming over Wi-Fi using Unity render streaming. *Comput. Commun.* 226–227, 107919. doi:10.1016/j.comcom.2024.08.001

Caserman, P., Garcia-Agundez, A., and Göbel, S. (2020). A survey of full-body motion reconstruction in immersive virtual reality applications. *IEEE Trans. Vis. Comput. Graph.* 26, 3089–3108. doi:10.1109/TVCG.2019.2912607

Casparsen, A., Soret, B., Jessen Nielsen, J., and Popovski, P. (2025). Near real-time data-driven adaptive RAN slicing of VR traffic in ORAN. *IEEE Open J. Commun. Soc.* 6, 4624–4637. doi:10.1109/ojcoms.2025.3571498

Chakareski, J., and Mastronarde, N. (2020). "Mobile-edge cooperative multi-user 360° video computing and streaming," in *2020 IEEE 22nd international workshop on multimedia signal processing (MMSP)*, 1–6. doi:10.1109/MMSP48831.2020

Chang, E., Lee, Y., Yoo, B., and Im, H. (2023). "Comparing context-sharing interfaces in XR remote collaboration," in *Proceeding - 2023 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)* (Institute of Electrical and Electronics Engineers Inc.), 719–720. doi:10.1109/VRW58643.2023.00202

Chaudhari, B. S. (2025). Enabling tactile internet *via* 6G: application characteristics, requirements, and design considerations. *Future Internet* 17, 122. doi:10.3390/fi17030122

Chen, R., and Guo, S. (2023). Look-ahead task offloading for multi-user Mobile augmented reality in edge-cloud computing. *IEEE Netw.* 37, 40–46. doi:10.1109/mnet.001.2200593

Chen, S., and Jiang, W. (2024). Online dynamic multi-user computation offloading and resource allocation for HAP-assisted MEC: an energy efficient approach. *J. Cloud Comput.* 13, 92. doi:10.1186/s13677-024-00645-5

Chen, J., and Konomi, S. (2022). "Utilization of XR technology in distance collaborative learning: a systematic review," in *Cross-cultural design. Applications in learning, arts, cultural heritage, creative industries, and virtual reality*. Editor P.-L. P. Rau (Springer International Publishing), 13312, 14–29. doi:10.1007/978-3-031-06047-2_2

Chen, Z., Zou, L., Tao, X., Xu, L., Muntean, G.-M., and Wang, X. (2022). EdgeVR360: edge-assisted multiuser-oriented intelligent 360-degree video delivery scheme over wireless networks. *Lect. Notes in Comput. Sci. Incl. Subseries Lect. Notes in Artif. Intell. and Lect. Notes in Bioinforma. 13605 LNAI*, 242–255. doi:10.1007/978-3-031-20500-2_20

Chen, J., Ramakrishnan, K., Dhakazl, A., and Ran, X. (2023). Networked architectures for localization-based multi-user augmented reality. *IEEE Commun. Mag.* 61, 104–110. doi:10.1109/mcom.003.2300275

Chen, J., Qin, X., Zhu, G., Ji, B., and Li, B. (2024). Motion-prediction-based wireless scheduling for interactive panoramic scene delivery. *IEEE Trans. Netw. Sci. Eng.* 11, 1566–1579. doi:10.1109/tnse.2023.3325420

Choi, H.-J., Komuro, N., and Kim, W.-S. (2024). Microservices-based resource provisioning for multi-user cloud VR in edge networks. *Electron. Switz.* 13, 3077. doi:10.3390/electronics13153077

Choi, H.-J., Kim, J.-H., Lee, J.-H., Han, J.-Y., and Kim, W.-S. (2025). Adaptive microservice architecture and service orchestration considering resource balance to support multi-user cloud VR. *Electron. Switz.* 14, 1249. doi:10.3390/electronics14071249

Dhakal, A., Ran, X., Wang, Y., Chen, J., and Ramakrishnan, K. K. (2022). "SLAM-share: visual simultaneous localization and mapping for real-time multi-user augmented reality," in *Proceedings of the 18th international conference on emerging networking EXperiments and technologies* (New York, NY, United States: Association for Computing Machinery) (CoNEXT '22), 293–306. doi:10.1145/3555050.3569142

Elvezio, C., Ling, F., Liu, J.-S., and Feiner, S. (2018). "Collaborative virtual reality for low-latency interaction," in *Adjunct proceedings of the 31st annual ACM symposium on user interface software and technology* (Berlin Germany: ACM), 179–181.

Esswie, A. A., and Repeta, M. (2023). "Evolution of 3GPP standards towards true extended reality (XR) support in 6G networks," in *2023 IEEE international black sea conference on communications and networking (BlackSeaCom)*, 7–14. doi:10.1109/BlackSeaCom58138.2023.10299766

ETSI (2023). *ETSI GR NFV 003 V1.7.1: network Functions Virtualisation (NFV); terminology for Main Concepts in NFV*. France: ETSI Group Report (GR) 003, European Telecommunications Standards Institute, Sophia Antipolis.

Evgenieva, E., Vlahov, A., Ivanov, A., Poulkov, V., and Manolova, A. (2025). A comprehensive survey of 6G simulators: Comparison, integration, and future directions. *Electronics* 14, 3313. doi:10.3390/electronics14163313

Fowler, S., and Souihi, S. (2025). Multi-user QoE optimization in virtual reality gaming using RL on Mobile edge networks. *Clust. Comput.* 28, 597. doi:10.1007/s10586-025-05271-7

Gao, B., Sheng, D., Zhang, L., Qi, Q., He, B., Zhuang, Z., et al. (2024). "STAR-VP: improving long-term viewport prediction in 360° videos *via* space-aligned and time-varying fusion," in *Proceedings of the 32nd ACM international conference on multimedia* (New York, NY, United States: Association for Computing Machinery), 24, 5556–5565. doi:10.1145/3664647.3681268

Gao, N., Wang, Y., Hua, X., Bi, T., and Jiang, T. (2024). 360° video multicast scheduling optimization in orthogonal and nonorthogonal scenarios. *IEEE Internet Things J.* 11 (10), 17128–17140. doi:10.1109/jiot.2024.3357133

GPP (2022). TR 26.928: extended reality (XR) in the context of 5G radio and network services. *Tech. Rep. 26.928, 3rd Gener. Partnersh. Proj. (3GPP). Release 17*.

Guertin-Lahoud, S., Coursaris, C. K., Sénécal, S., and Léger, P.-M. (2023). User experience evaluation in shared interactive virtual reality. *Cyberpsychology, Behav. Soc. Netw.* 26, 263–272. doi:10.1089/cyber.2022.0261

Ha Huynh, V. S., Qian, P., Wang, N., Udora, C., and Tafazolli, R. (2024). "User-intent aware transport-layer intelligence for frame synchronisation in multi-party XR application," in *2024 IEEE gaming, entertainment, and media conference (GEM)*, 1–6.

Han, B., Pathak, P., Chen, S., and Yu, L.-F. (2023). CoMIC: a collaborative Mobile immersive computing infrastructure for conducting multi-user XR research. *IEEE Netw.* 37, 124–131. doi:10.1109/mnet.126.2200385

Hart, S. G. (2006). Nasa-task load index (NASA-TLX); 20 years later. *Proc. Hum. Factors Ergonomics Soc. Annu. Meet.* 50, 904–908. doi:10.1177/154193120605000909

Hatami, M., Qu, Q., Chen, Y., Kholidy, H., Blasch, E., and Ardiles-Cruz, E. (2024). A survey of the real-time metaverse: challenges and opportunities. *Future Internet* 16, 379. doi:10.3390/fi16100379

Hazarika, A., and Rahmati, M. (2023). Towards an evolved immersive experience: exploring 5G- and beyond-enabled ultra-low-latency communications for augmented and virtual reality. *Sensors* 23, 3682. doi:10.3390/s23073682

Hu, Y., Tutschku, K., Boeva, V., Goswami, P., Abghari, S., and Sundstedt, V. (2024). "Towards an ethical and data privacy metrology for AI-Enriched human-centered XR systems," in *2024 IEEE International Conference on Metrology for eXtended Reality, Artificial Intelligence and Neural Engineering (MetroXRAINE) (St Albans, United Kingdom: IEEE)*, 119–124. doi:10.1109/MetroXRAINE62247.2024.10796457

Huang, Q., and Song, J. (2025). "Consistency-driven resource allocation in Mobile augmented reality systems," in *2025 international conference on communication, remote sensing and information technology (CRSIT)*, 196–200. doi:10.1109/CRSIT66101.2025.11141579

Huang, Y., Wang, H., Qiao, X., Su, X., Li, Y., Dustdar, S., et al. (2024). SCAXR: empowering scalable multi-user interaction for heterogenous XR devices. *IEEE Netw.* 38 (4), 250–258. doi:10.1109/MNET.2023.3320440

Huang, Z., Xiong, C., Ni, H., Wang, D., Tao, Y., and Sun, T. (2023). *Standard evolution of 5G-Advanced and future Mobile network for extended reality and metaverse 6*, 20–25. doi:10.1109/IOTM.001.2200261

IEEE ICIR 2025 Conference Committee (2025). *Ieee 4th international conference on intelligent reality (icir 2025)*. Available online at: https://icir.ieee.org/ (Accessed April 3, 2026).

ITU-R (2023). *Framework and overall objectives of the future development of IMT for 2030 and beyond. Recommendation ITU-R*, M.2160-0. Geneva, Switzerland: International Telecommunication Union.

Jackson, H., Thakur, S., and Olanrewaju, O. A. (2025). The potential of extended reality in engineering education: a systematic review. *Cogent Educ.* 12, 2541442. doi:10.1080/2331186x.2025.2541442

Jain, R., Chiu, D.-M. W., and Hawe, W. R. (1984). *A quantitative measure of fairness and discrimination for resource allocation in shared computer systems*. DEC Research Report TR-301. Hudson, MA: Digital Equipment Corporation.

Kamarianakis, M., Lydatakis, N., and Papagiannakis, G. (2021). Never 'Drop the Ball' in the operating room: an efficient hand-based VR HMD controller interpolation algorithm, for collaborative, networked virtual environments. *Lect. Notes in Comput. Sci. Incl. Subseries Lect. Notes in Artif. Intell. Lect. Notes in Bioinforma.* 13002, 694–704. doi:10.1007/978-3-030-89029-2_52

Kennedy, R. S., Lane, N. E., Berbaum, K. S., and Lilienthal, M. G. (1993). Simulator sickness questionnaire: an enhanced method for quantifying simulator sickness. *Int. J. Aviat. Psychol.* 3, 203–220. doi:10.1207/s15327108ijap0303_3

Kim, H. K., Park, J., Choi, Y., and Choe, M. (2018). Virtual reality sickness questionnaire (VRSQ): motion sickness measurement index in a virtual reality environment. *Appl. Ergon.* 69, 66–73. doi:10.1016/j.apergo.2017.12.016

Kourtesis, P. (2024). A comprehensive review of multimodal XR applications, risks, and ethical challenges in the metaverse. *Multimodal Technol. Interact.* 8, 98. doi:10.3390/mti8110098

Krodel, T., Schott, V., Mayer, A., and Ovtcharova, J. (2024). "Impact of XR-Enabled collaboration in businesses—an economic, ecological, and social perspective,"*Stud. Syst. Decis. Control*, 440, 767–777. doi:10.1007/978-3-031-42085-6_66

Kumar, N., and Raubal, M. (2021). Applications of deep learning in congestion detection, prediction and alleviation: a survey. *Transp. Res. Part C Emerg. Technol.* 133, 103432. doi:10.1016/j.trc.2021.103432

Kuo, W. H., Hsu, M. C., and Wu, H. C. (2025). Novel downlink multiuser resource-allocation scheme for providing layer-encoded multimedia streams using massive MIMO transmissions, *IEEE Trans. Netw. Serv. Manage.*, 22, 5959–5971. doi:10.1109/TNSM.2025.3595142

Lee, E.-S., and Shin, B.-S. (2023). Enhancing the performance of XR environments using fog and cloud computing. *Appl. Sci.* 13, 12477. doi:10.3390/app132212477

Lin, C.-W., Yang, D.-N., and Liao, W. (2025). Mobile tile-based 360° video multicast with cybersickness alleviation. *IEEE Trans. Mob. Comput.* 24, 3423–3440. doi:10.1109/tmc.2024.3514852

Liu, H., Yang, P., Wang, X., and Liu, W. (2023). "A QoE-Fair synchronized transmission scheme for edge-assisted interactive virtual reality," in *Ieee infocom 2023 - IEEE conference on computer communications workshops (INFOCOM WKSHPS)*, 1–6.

Lv, W., Chen, J., Cheng, S., Qiu, X., and Li, D. (2024). QoS-driven resource allocation in fog radio access network: a VR service perspective. *Math. Biosci. Eng.* 21, 1573–1589. doi:10.3934/mbe.2024068

Mahmoud, M., Rizou, S., Panayides, A., Kantartzis, N., Karagiannidis, G., Lazaridis, P., et al. (2024). Optimized tile quality selection in multi-user 360° video streaming. *IEEE Open J. Commun. Soc.* 5, 7301–7316. doi:10.1109/ojcoms.2024.3498334

Mallik, A., Xie, J., and Han, Z. (2024). "A performance analysis modeling framework for extended reality applications in edge-assisted wireless networks," in *2024 IEEE 44th international conference on distributed computing systems (ICDCS)*, 726–737. doi:10.1109/ICDCS60910.2024.00073

Mathias, P. M., Javier, M. S., and Felipe, M. L. R. (2025). Virtual worlds in AECO operations: towards a human-centric framework. *Automation Constr.* 180, 106529. doi:10.1016/j.autcon.2025.106529

Mon, B. F., Hayajneh, M., Ali, N. A., Ullah, F., Ullah, H., and Alkobaisi, S. (2025). Digital twins in the IIoT: current practices and future directions toward industry 5.0. *Comput. Mater. and Continua* 83, 3675–3712. doi:10.32604/cmc.2025.061411

Nasrallah, A., Thyagaturu, A. S., Alharbi, Z., Wang, C., Shao, X., Reisslein, M., et al. (2019). Ultra-low latency (ULL) networks: the IEEE TSN and IETF DetNet standards and related 5G ULL research. *IEEE Commun. Surv. and Tutorials* 21, 88–145. doi:10.1109/COMST.2018.2869350

Nath, S., and Wu, J. (2020). "Dynamic computation offloading and resource allocation for multi-user Mobile edge computing," in *Globecom 2020 - 2020 IEEE global communications conference*, 1–6.

Norambuena, N., Ortega, J., Muñoz-La Rivera, F., Covarrubias, M., Valín Rivera, J. L., Ramírez, E., et al. (2025). Integrating digital twins of engineering labs into multi-user virtual reality environments. *Appl. Sci.* 15, 3819. doi:10.3390/app15073819

Novotny, A., Gudmundsson, R., and Harris, F. C. (2020). A Unity framework for multi-user VR experiences. *EPiC Ser. Comput.* 69, 13–21.

Okafor, O., Esposito, F., and Pecorella, T. (2024). "EdgeVerse: multi-user virtual reality *via* edge computing and eBPF," in *Proceedings of the 2024 20th international conference on network and service management: AI-powered network and service management for tomorrow's digital world, CNSM 2024*.

Okamoto, T., Ishioka, T., Shiina, R., Fukui, T., Ono, H., Fujiwara, T., et al. (2023). "Edge-assisted multi-user 360-Degree video delivery," in *2023 IEEE 20th consumer communications and networking conference (CCNC)*, 194–199.

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., et al. (2021). The prisma 2020 statement: an updated guideline for reporting systematic reviews. *J. Clin. Epidemiol.* 134, 178–189. doi:10.1016/j.jclinepi.2021.03.001

Patra, A., Pandey, A., Hassija, V., Chamola, V., and Mishra, R. P. (2024). A survey on edge enabled metaverse: applications, technological innovations, and prospective trajectories within the industry. *IEEE Access* 12, 125125–125144. doi:10.1109/ACCESS.2024.3452184

Peñaherrera-Pulla, O. S., Damsgaard, S. B., Yanakiev, B., Mogensen, P., Fortes, S., and Barco, R. (2024). Cloud VR on 5G: a performance validation in industrial scenarios. *IEEE Open J. Commun. Soc.* 5, 3641–3657. doi:10.1109/OJCOMS.2024.3413328

Pindi, N. R., and Velez, F. J. (2025). Traffic scheduling and resource allocation for heterogeneous services in 5G new radio networks: a scoping review. *Smart Cities* 8, 168. doi:10.3390/smartcities8050168

Pooyandeh, M., Han, K.-J., and Sohn, I. (2022). Cybersecurity in the AI-Based metaverse: a survey. *Appl. Sci.* 12, 12993. doi:10.3390/app122412993

Ran, X., Slocum, C., Tsai, Y.-Z., Apicharttrisorn, K., Gorlatova, M., and Chen, J. (2020). "Multi-user augmented reality with communication efficient and spatially consistent virtual objects," in *CoNEXT 2020 - Proceedings of the 16th international conference on emerging networking EXperiments and technologies*, 386–398.

Ren, P., Qiao, X., Huang, Y., Liu, L., Pu, C., Dustdar, S., et al. (2022). Edge AR X5: an edge-assisted multi-user collaborative framework for Mobile web augmented reality in 5G and beyond. *IEEE Trans. Cloud Comput.* 10, 2521–2537. doi:10.1109/tcc.2020.3046128

Sameri, J., Van Damme, S., Schwarzmann, S., Wei, Q., Trivisonno, R., De Turck, F., et al. (2024). "Collaborative cooking in VR: effects of network distortion in multi-user virtual environments," in *MMSys 2024 - proceedings of the 2024 ACM multimedia systems conference*, 509–515.

Sandeepa, D., Gimhana, D., Lakpriya, A., Jayaweera, L., Dias, D., and Hemachandra, K. T. (2024). "Competitive virtual cycling platform design for low latency networks," in *2024 IEEE wireless communications and networking conference (WCNC)*, 1–6.

Santos, J., van der Hooft, J., Vega, M. T., Wauters, T., Volckaert, B., and De Turck, F. (2021). "Efficient orchestration of service chains in fog computing for immersive media," in *2021 17th international conference on network and service management (CNSM)*, 139–145.

Santos, H., Rosário, D., Cerqueira, E., and Braun, T. (2022). "Multi-criteria service function chaining orchestration for multi-user virtual reality services," in *GLOBECOM 2022 - 2022 IEEE global communications conference* (GLOBECOM), 6360–6365. doi:10.1109/GLOBECOM48099.2022.10000827

Santos, H., Martins, B., Rosário, D., Cerqueira, E., and Braun, T. (2023). "Mobility-aware service function chaining orchestration for multi-user augmented reality," in *2023 IEEE 48th conference on local computer networks (LCN)*, 1–9. doi:10.1109/LCN58197.2023.10223348

Skorin-Kapov, L., Matijasevic, M., Zarko, I. P., Kusek, M., Huljenic, D., Skarica, V., et al. (2025). Shared presence *via* XR communication and interaction within a dynamically updated digital twin of a smart space: conceptual framework and research challenges. *Appl. Sci.* 15, 8838. doi:10.3390/app15168838

Sonkoly, B., Nagy, B. G., Dóka, J., Kecskés-Solymosi, Z., Czentye, J., Formanek, B., et al. (2023). "Towards an edge cloud based coordination platform for multi-user AR applications built on open-source SLAMs," in *NOMS 2023-2023 IEEE/IFIP network operations and management symposium*, 1–6.

Sonkoly, B., Nagy, B. G., Dóka, J., Kecskés-Solymosi, Z., Czentye, J., Formanek, B., et al. (2024). An edge cloud based coordination platform for multi-user AR applications. *J. Netw. Syst. Manag.* 32, 40. doi:10.1007/s10922-024-09809-9

Stauffert, J.-P., Niebling, F., and Latoschik, M. E. (2020). Latency and cybersickness: impact, causes, and measures. A review. *Front. Virtual Real.* 1, 582204. doi:10.3389/frvir.2020.582204

Sundstedt, V., Boeva, V., Zepernick, H.-J., Goswami, P., Cheddad, A., Tutschku, K., et al. (2023). "HINTS: human-centered intelligent realities," in *35th annual workshop of the swedish artificial intelligence society SAIS 2023*, 9–17.

Trinh, B., and Muntean, G.-M. (2022). "A deep reinforcement learning-based resource management scheme for SDN-MEC-supported XR applications," in *2022 IEEE 19th annual consumer communications & networking conference (CCNC)* (Institute of Electrical and Electronics Engineers Inc.), 790–795. doi:10.1109/CCNC49033.2022.9700522

Tseng, H.-W., Yang, T.-T., and Yang, Y. (2024). "An adaptive transmission strategy for tiled 360-Degree VR videos in NOMA systems," in *Proceedings of the ACM symposium on applied computing*, 1765–1771.

Tuncer, D., Charalambides, M., and Pavlou, G. (2019). Management application interactions in software-based networks. *IEEE Netw.* 33, 149–155. doi:10.1109/mnet.2019.1800276

Van Damme, S., Van de Velde, F., Sameri, M. J., De Turck, F., and Vega, M. T. (2023). "A haptic-enabled, distributed and networked immersive System for multi-user collaborative virtual reality," in *Proceedings of the 2nd international workshop on interactive EXtended reality* (New York, NY, USA: Association for Computing Machinery), 23, 11–19. doi:10.1145/3607546.3616804

Van Damme, S., Sameri, J., Schwarzmann, S., Wei, Q., Trivisonno, R., De Turck, F., et al. (2024). Impact of latency on QoE, performance, and Collaboration in interactive multi-user virtual reality. *Appl. Sci. Switz.* 14, 2290. doi:10.3390/app14062290

Vidal-Balea, A., Blanco-Novoa, O., Fraga-Lamas, P., and Fernández-Caramés, T. M. (2025). Off-Cloud anchor sharing framework for multi-user and multi-platform mixed reality applications. *Appl. Sci.* 15, 6959. doi:10.3390/app15136959

Vlahovic, S., Suznjevic, M., and Skorin-Kapov, L. (2022). A survey of challenges and methods for quality of experience assessment of interactive VR applications, *J. Multimodal User Interfaces, Surv. Challenges Methods Qual. Exp. Assessment Inter. VR Applications* 16, 257–291. doi:10.1007/s12193-022-00388-0

Wang, Y., and Zhao, J. (2022). "A survey of Mobile edge computing for the metaverse: architectures, applications, and challenges," in *2022 IEEE 8th international conference on collaboration and internet computing (CIC)*, 1–9. doi:10.1109/CIC56439.2022.00011

Wang, C., Zhang, S., Qian, Z., Xiao, M., Wu, J., Ye, B., et al. (2020). Joint server assignment and resource management for edge-based MAR system, *IEEE. ACM. Trans. Netw., Jt. Serv. Assign. Resour. Manag. Edge-Based MAR Syst.* 28, 2378–2391. doi:10.1109/TNET.2020.3012410

Wang, Y., Su, Z., Zhang, N., Xing, R., Liu, D., Luan, T. H., et al. (2023). A Survey on metaverse: fundamentals, security, and privacy. *IEEE Commun. Surv. and Tutorials* 25, 319–352. doi:10.1109/COMST.2022.3202047

Wang, P., Wang, S., Billinghurst, M., Yang, H., Zhang, X., Luo, R., et al. (2026). A critical analysis of XR remote collaboration using sketch cues. *J. King Saud Univ. Comput. Inf. Sci.* 38, 17. doi:10.1007/s44443-025-00409-4

Wang, P., Yang, H., Billinghurst, M., Zhao, S., Wang, Y., Liu, Z., et al. (2025). A survey on XR remote collaboration in industry. *J. Manuf. Syst.* 81, 49–74. doi:10.1016/j.jmsy.2025.05.016

Wang, S., Tan, X., Li, S., Xu, X., Yang, J., and Zheng, Q. (2020). "A QoE-based 360°video adaptive bitrate delivery and caching scheme for C-RAN," in *Proceedings - 2020 16th international conference on mobility, sensing and networking, MSN 2020*, 49–56.

Wang, S., Tan, X., Liu, Z., Tai, Z., Sun, M., Yang, J., et al. (2025). Adaptive computing and multicasting optimization for live 360-Degree video streaming, *IEEE Trans. Veh. Technol., Adapt. Comput. Multicasting Optim. Live 360-Degree Video Streaming* 75, 3190–3204. doi:10.1109/TVT.2025.3605289

Wang, Y., Gong, D., Xiao, R., Wu, X., and Zhang, H. (2024). A systematic review on extended reality-mediated multi-user social engagement. *Systems* 12, 396. doi:10.3390/systems12100396

Wang, Z., Wang, R., Wu, J., Zhang, W., and Li, C. (2024). Dynamic resource allocation for real-time cloud XR video transmission: a reinforcement learning approach. *IEEE Trans. Cognitive Commun. Netw.* 10, 996–1010. doi:10.1109/tccn.2024.3352982

Warburton, M., Mon-Williams, M., Mushtaq, F., and Morehead, J. R. (2023). Measuring motion-to-photon latency for sensorimotor experiments with virtual reality systems. *Behav. Res. Methods* 55, 3658–3678. doi:10.3758/s13428-022-01983-5

Wu, Y., Xu, M., Zhai, G., and Zhang, W. (2024). Physical layer signal processing for XR communications and systems. *Sci. China Inf. Sci.* 67, 221301. doi:10.1007/s11432-023-4122-4

Xu, C., Chen, Z., Tao, M., and Zhang, W. (2025). Wireless multi-user interactive virtual reality in metaverse with edge-device collaborative computing. *IEEE Trans. Wirel. Commun.* 24 (7), 6135–6150. doi:10.1109/TWC.2025.3551938

Yang, S., Yang, P., Chen, J., Ye, Q., Zhang, N., and Shen, X. (2024). Delay-Optimized multi-user VR streaming *via* end-edge collaborative neural frame interpolation. *IEEE Trans. Netw. Sci. Eng.* 11, 284–298. doi:10.1109/tnse.2023.3296511

Yu, H., Shokrnezhad, M., Taleb, T., Li, R., and Song, J. (2023). Toward 6G-Based metaverse: supporting highly-dynamic deterministic multi-user extended reality services. *IEEE Netw.* 37, 30–38. doi:10.1109/mnet.004.2300101

Yu, W., Chua, T. J., and Zhao, J. (2023a). Mobile edge computing and AI enabled Web3 metaverse over 6G wireless communications: a deep reinforcement learning approach. *IEEE Veh. Technol. Conf.* 2023-June, 1–10. doi:10.1109/VTC2023-Spring57618.2023.10199534

Yu, W., Chua, T. J., and Zhao, J. (2023b). Virtual reality in metaverse over wireless networks with user-centered deep reinforcement learning. *IEEE Int. Conf. Commun.* 2023-May, 6639–6644. doi:10.1109/ICC45041.2023.10278715

Zhang, W., Han, B., and Hui, P. (2022). SEAR: scaling experiences in multi-user augmented reality. *IEEE Trans. Vis. Comput. Graph.* 28, 1982–1992. doi:10.1109/TVCG.2022.3150467

Zhao, L., Zhou, Q., Chen, B., and Nahrstedt, K. (2025). "360LiveCast: a low-latency and bandwidth-efficient multicast framework for live 360 video," in *2025 IEEE 8th international conference on multimedia information processing and retrieval (MIPR)*, 226–232.

Zhong, L., Chen, X., Xu, C., Ma, Y., Wang, M., Zhao, Y., et al. (2023). A multi-user cost-efficient crowd-assisted VR content delivery solution in 5G-and-Beyond heterogeneous networks. *IEEE Trans. Mob. Comput.* 22, 4405–4421. doi:10.1109/tmc.2022.3162147