Quality of Experience Enhancement in Wireless Metaverse: A Resource Optimization Scheme

Maher Guizani‡, Latif U. Khan†, Mohammad A. Islam‡

[mahergzani@gmail.com, latif.khan2@gmail.com, mislam@uta.edu]
‡ University of Texas at Arlington, USA.
† Mohamed Bin Zayed University of Artificial Intelligence, United Arab Emirates.

***Abstract*—The rapid advancement of metaverse applications in wireless environments necessitates efficient resource management to enhance Quality of Experience (QoE). This paper presents a novel framework for optimizing wireless resource allocation within the metaverse to optimize QoE using convex optimization and matching theory. We formulate a QoE optimization problem that considers packet error rate (PER) and immersive experience. Our problem also enables us to make a tradeoff between immersive experience and PER while computing QoE. The formulated problem is an NP-hard mixed-integer non-linear programming (MINLP) problem, which we address through decomposition and convex optimization scheme, matching theory, and BSUM. Specifically, for a solution, our proposed model integrates matching theory, block-successive upper-bound minimization (BSUM), and convex optimization to optimize for the association, transmit power allocation, and resource allocation. Finally, numerical results are provided.**

***Index Terms*—Internet of Things, convex optimization, metaverse, matching theory.**

I. Introduction

The emergence of the wireless metaverse presents unprecedented opportunities for immersive virtual experiences, driven by advancements in augmented reality (AR), virtual reality (VR), and mixed reality (MR) [1], [2]. These applications demand ultra-low latency, high data rates, and seamless connectivity to ensure a high Quality of Experience (QoE) for users. Specifically, one can say that the Wireless Metaverse is an emerging paradigm that integrates virtual and augmented reality experiences with wireless networking and edge computing technologies. Ensuring optimal resource allocation to maximize Quality of Experience (QoE) is a fundamental challenge in this domain. However, achieving optimal QoE in wireless metaverse environments is challenging due to limited network resources, dynamic wireless conditions, and computational constraints [3]–[5]. To address the challenges of wireless metaverse, resource optimization plays a critical role in maximizing QoE while ensuring efficient utilization of network and computing resources. Traditional resource allocation techniques are often inadequate in handling the complex and dynamic nature of metaverse applications, necessitating novel optimization frameworks that leverage artificial intelligence (AI), edge computing, and network-aware strategies.

Various works [1]–[6] considered metaverse and wireless systems. In [6], the authors provide a broad vision of Metaverse-enabled wireless systems, discussing key enablers such as edge computing, artificial intelligence (AI)-driven resource management, and immersive service delivery. The study highlights the importance of low-latency communication and adaptive resource allocation for enhancing user experience. Additionally, it outlines a potential network architecture that leverages intelligent resource allocation to balance computational loads across cloud and edge nodes. Optimizing task offloading strategies is critical for ensuring smooth and immersive Metaverse interactions. The work in [7] explores dynamic task offloading mechanisms that distribute computational tasks between cloud servers and edge computing nodes. The study introduces a framework that dynamically adjusts resource allocation based on network conditions and user mobility, improving system efficiency while minimizing latency. Furthermore, the authors propose a cost-aware offloading mechanism that balances computational demands and energy consumption. On the other hand, maximizing QoE in Metaverse environments requires intelligent resource allocation policies that consider user experience metrics such as responsiveness, visual quality, and interaction fluidity. The study in [8] discusses QoE-driven resource allocation techniques that leverage machine learning to predict user demand patterns and dynamically allocate bandwidth and computational resources. The authors demonstrate how AI-powered predictive modeling can proactively adjust network parameters to optimize service delivery. On the other hand, other works also considered resource allocation in metaverse [9]–[11]. Different from the existing works, we consider quality of experience (QoE). The summary of our contributions is as follows:

- We propose a novel framework for wireless metaverse and consider QoE in terms of immersive experience and PER.
- The formulated problem is a mixed-integer non-linear programming (MINLP) problem and difficult to solve using traditional convex optimization schemes. Therefore, we decompose the main problem into simple problems and then use block-successive upper-bound minimization (BSUM), matching theory, and relaxation-based schemes.
- Finally, we present numerical simulations and conclude our work.

The rest of the paper is structured as follows: Section II presents our proposed optimization framework, detailing its

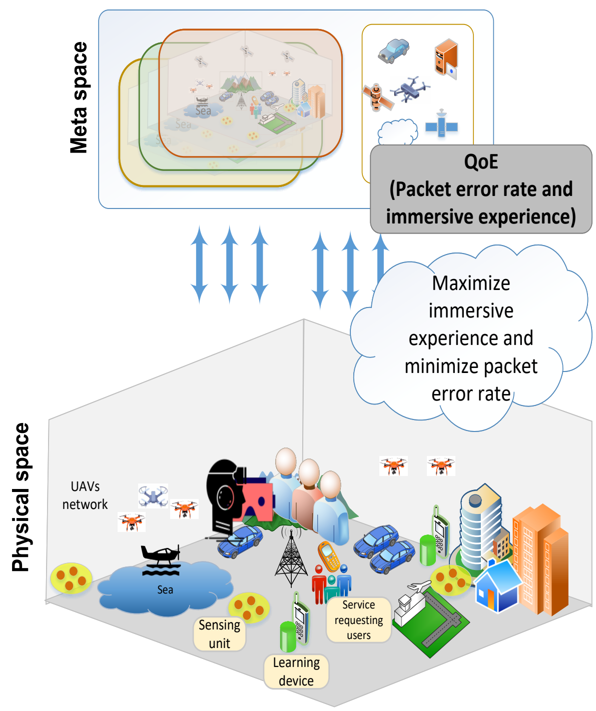


Fig. 1: Proposed system model.

key components and methodologies. Section III provides the proposed solution approach. Section IV discusses the implementation and performance evaluation, followed by results and analysis in Section V. Finally, Section VI concludes the paper with future research directions.

II. System Model and Problem Formulation

Consider a metaverse-empowered wireless system, as illustrated in Fig. 1. There are two spaces: (a) a physical space and (b) a meta space. All entities in the system are in physical space. On the other hand, the meta space is based on a virtual model of the physical world. These spaces interact with each other for enabling various operations. The interaction is via wireless or wired media. Let us consider a set $\mathcal{G}$ of G entities in the physical space that wants to communicate physical space states with the meta space. For instance, it can be the set of AR/VR headsets that can sense the environment and then share the sensed data with the meta space. The states of the physical space should be communicated with the meta space timely for up to date operation. To do so, there is a need for deployment of a massive number of sensors. These sensors can be modeled by considering a continuous distribution of sensors similar to the work in [12]. Next, these continuously distributed sensors can be divided into discrete units and then perform communication modeling. Let us consider a distribution function, $f(a, b, c)$, where a, b, and c denote the longitude distance, latitude distance, and height, respectively. This distribution function will be used for computing latency. Next, we present the communication model.

A. Metaverse Communication Model

For communication between meta space and the physical space, a set $\mathcal{O}$ of O resource blocks that are orthogonal in nature. As the number communication end-nodes will be massive in the foreseeable future, therefore, we need to plan the efficient use of the communication resources. To do so, in our system model, we propose the reuse of resource blocks for metaverse sensing units. For a resource block o, we can write the throughput as follows.

$$\Re_g = T_o\left(\log_2\left(1+\frac{p_g h_{g,b}}{\sum_{c\in\mathcal{C}} p_c h_{c,b} + N_o^2}\right)\right),\ \forall g\in\mathcal{G},\quad (1)$$

where $\sum_{c\in\mathcal{C}} p_c h_{n,b}$ and N_o^2 denote the interference due to other users already using the resource block o and noise, respectively. T_o is the bandwidth of the resource block o. For a sensing unit g, the transmission latency associated with a meta space b using a resource block o, can be given by:

$$\Xi_g(\boldsymbol{\kappa},\boldsymbol{\upsilon},\boldsymbol{p}) = \left(\frac{\kappa_{(g,b)}\upsilon_{(g,o)}\chi_g\epsilon\eta(x,y,z)}{T_o\left(\log_2\left(1+\frac{p_g h_{g,b}}{\sum_{c\in\mathcal{C}} p_c h_{c,b}+N_o^2}\right)\right)}\right),\quad \forall g\in\mathcal{G},\quad (2)$$

where $\kappa_{(g,b)}$ and $\upsilon_{(g,o)}$ are the association and resource allocation variables, respectively. $\kappa_{(g,b)} = 1$ when the sensing unit g is associated with meta space b and $\kappa_{(g,b)} = 0$, otherwise. Similarly, the resource allocation variable $\upsilon_{(g,o)} = 1$ when the resource block o is assigned to a sensing unit g and $\upsilon_{(g,o)} = 0$, otherwise. p_g is the transmit power. In our system, a single sensing unit must be assigned to a maximum of one meta space.

$$\sum_{b=1}^{B} \kappa_{(g,b)} \leq 1\ \forall g\in\mathcal{G},\quad (3)$$

Every meta space has a certain capacity to serve the sensing entity by processing the data received from them. To process the sensing unit data, there is a need for computing and storage resources. Meanwhile, there are limitations on the availability of computing and storage resources at the meta space. Therefore, one must ensure the association should not exceed the serving capacity of the meta spaces (e.g., ζ_b of meta space b).

$$\sum_{g=1}^{G} \kappa_{(g,b)} \leq \zeta_b\ \forall b\in\mathcal{B},\quad (4)$$

On the other hand, the transmit power should be within a range:

$$p_{\min} \leq p_g \leq p_{\max},\ \forall g\in\mathcal{G},\quad (5)$$

Meanwhile, the summation of transmit power of all devices should not exceed the maximum limit.

$$\sum_{g=1}^{G} p_g \leq p_{\text{MAX}}. \tag{6}$$

On the other hand, communication resource blocks are limited and thus, there is a need for constraints. A single resource block should be assigned to only one sensing unit.

$$\sum_{g=1}^{G} \upsilon_{(g,o)} \leq 1 \;\forall o \in \mathcal{O}, \tag{7}$$

Similar to (7), there should be a constraint on the maximum number of resource blocks a single sensing unit can get. This is because of limited availability of the resource blocks. One can write this constraint as:

$$\sum_{o=1}^{O} \upsilon_{(g,o)} \leq 1 \;\forall g \in \mathcal{G}, \tag{8}$$

In our sensing for the metaverse system model, we want to ensure the quality of service (QoS) in order to get accurate and well synchronized operations of the metaverse. The QoS for a system can be defined in many ways, such as throughput, outage probability, reliability, and latency, among others. On the other hand, we are also interested in increasing the QoE when we consider metaverse. There are many ways to define the QoE for metaverse. For instance, we can consider packet error rate, immersive experience, or both packet error rate and immersive experience. Here, we consider the packet error rate and immersive experience as metrics for QoE.

B. QoE for Wireless Metaverse

In a metaverse, there is a wide variety of sensors (e.g., AR/VR headsets, autonomous car LIDARs, and surveillance cameras) that sense the physical world and then share the sensed data(e.g., movements, gestures, and environmental conditions) with the meta space. A visual sensing data (i.e., specifically movements and gestures) in addition to other types of data (e.g., time series data of traffic flow in ITS) need careful attention during transmission to the meta space. End-devices in the physical space sense the environment and share it with the meta space mainly for two purposes: (a) synchronization of the meta space and the physical space for updated operation and (b) for task offloading (e.g., rendering functions performed at the meta space) to meta space during limited computing power at end-devices. Mostly, the sensed data in metaverse applications consist of visual data [13]. For such kind of visual data, measuring the service quality has a subjective nature. Traditional QoS metrics (e.g., latency and reliability) are not only sufficient for sensing in the metaverse. Therefore, there is a need for other QoE metrics in addition to QoS metrics. For QoE metrics in metaverse, we should focus on human-centric utilities (e.g., human perception and feelings). In addition, QoE should also focus on the packet error rate. Packet error rate is important due to the fact that receiving intact packets of data will enable us with a good experience and vice versa. Different works considered different metrics for QoE [14]. The work in [15] considered the packet error rate (PER) as a QoE metric for interactions between the virtual environment and the end-users. We consider a QoE metric that accounts for PER and immersive experience.

$$\mathcal{Q}(\boldsymbol{p},\boldsymbol{x},\boldsymbol{z}) = \beta\left(1-\exp\left(\frac{-\vartheta(\sum_{y\in\mathcal{Y}_r} h_y^r P_y^r + \sigma^2)}{p_n h_{n,m}^r}\right)\right) + (1-\beta)\frac{1}{I_g \ln\left(\frac{R}{R_{\text{min}}}\right)} \;\forall g \in \mathcal{G}, \tag{9}$$

where β is a scaling constant that controls the proportion of immersive experience and PER in the calculation of QoE. In (9), we have two terms: (a) immersive experience and (b) packet error rate. We want to maximize the immersive experience (i.e., $I_g \ln\left(\frac{R}{R_{\text{min}}}\right)$). Meanwhile, we want to minimize the packet error rate, (i.e., $\left(1-\exp\left(\frac{-\vartheta(\sum_{y\in\mathcal{Y}_r} h_y^r P_y^r+\sigma^2)}{p_n h_{n,m}^r}\right)\right)$). Therefore, we write QoE as a summation of PER term and $(1/\ immserive\ experience\ term)$ as we want to optimize the QoS by minimizing the PER and the other term. Next, we present our problem formulation.

C. Problem Formulation

In this section, we present the problem formulation that is based on optimizing the QoE as follows.

$$\mathbf{P}: \underset{\kappa,\upsilon,\boldsymbol{p}}{\text{minimize}} \quad \mathcal{Q}(\boldsymbol{\kappa},\boldsymbol{\upsilon},\boldsymbol{p}) \tag{10}$$

subject to:

$$\sum_{b=1}^{B} \kappa_{(g,b)} \leq 1 \;\forall g \in \mathcal{G}, \tag{10a}$$

$$\sum_{g=1}^{G} \kappa_{(g,b)} \leq \zeta_b \;\forall b \in \mathcal{B}, \tag{10b}$$

$$p_{\text{min}} \leq p_g \leq p_{\text{max}}, \;\forall g \in \mathcal{G}, \tag{10c}$$

$$\sum_{g=1}^{G} p_g \leq p_{\text{MAX}}, \tag{10d}$$

$$\sum_{g=1}^{G} \upsilon_{(g,o)} \leq 1 \;\forall o \in \mathcal{O}, \tag{10e}$$

$$\sum_{o=1}^{O} \upsilon_{(g,o)} \leq 1 \;\forall g \in \mathcal{G}. \tag{10f}$$

$$\kappa_{(g,b)} \in \{0,1\}, \upsilon_{(g,o)} \in \{0,1\}. \tag{10g}$$

Constraints (10a) and (10b) are about the association. Constraints (10c) and (10d) set the limits of transmit power allocation. (10e) and (10f) set limits on the allocation of wireless resources. Constraint (10g) sets the binary nature of the association and resource allocation variables. Problem **P** is a MINLP and its solution is challenging. Therefore, it is not trivial to solve problem **P** directly. To address this, we will use a decomposition scheme that will be presented in the next section.

III. Proposed Solution

In this section, we present our decomposition-based solution. We use an iterative scheme that is based on solving a single problem by keeping the other variables fixed. Then, the other variables are solved. First, we write an association problem as follows:

$$\mathbf{P}-\kappa: \ \underset{\kappa}{\text{minimize}} \quad \mathcal{Q}(\boldsymbol{\kappa}) \tag{13}$$

subject to:

$$\sum_{g=1}^{G} \kappa_{(g,b)} \leq 1, \ \forall b \in \mathcal{B}, \tag{11a}$$

$$\sum_{g=1}^{G} \kappa_{(g,b)} \leq \zeta_b, \ \forall g \in \mathcal{G}, \tag{11b}$$

$$\kappa_{g,b} \in \{0,1\}, \ \forall g \in \mathcal{G}. \tag{11c}$$

To solve problem **P**-κ, we use a one-to-many matching due to its powerful feature of solving combinatorial problems with low complexity compared to exhaustive search schemes. For matching [16], first, we need to define the preference profile that is based on a QoE cost function.

$$\mathcal{P}_g(b) = \mathcal{Q}(\kappa) \tag{12}$$

Now, we write wireless resource allocation problem as.

$$\mathbf{P}-\upsilon: \ \underset{\upsilon}{\text{minimize}} \quad \mathcal{Q}(\boldsymbol{\upsilon}) \tag{13}$$

subject to:

$$\sum_{g=1}^{G} \upsilon_{g,o} \leq 1, \ \forall b \in \mathcal{B}, \tag{13a}$$

$$\sum_{o=1}^{O} \upsilon_{g,o} \leq 1, \ \forall g \in \mathcal{G}, \tag{13b}$$

$$\sum_{o=1}^{O} \upsilon_{g,o} \leq 1, \ \forall g \in \mathcal{G}, \tag{13c}$$

$$\upsilon_{g,o} \in \{0,1\}, \ \forall g \in \mathcal{G}. \tag{13d}$$

In the resource allocation problem, the variable υ is binary. Therefore, it is challenging to solve. Therefore, we transform the binary variable into a continuous variable and then apply convex optimization. Later, we will transform the continuous variable into a binary variable. Next, we write the transmit power allocation problem as follows:

Algorithm 1 Association algorithm

Next to computing preference profile, there is a need for matching algorithm, whose summary is given in algorithm 1.

Inputs
Devices preference profile $\mathcal{P}_g, \forall g \in \mathcal{G}$
Meta spaces $\mathcal{B}$
Output
Matching function $\mu^{(t)}$
Initialization step
$\mathcal{P}_g{}^{(0)} = \mathcal{P}_g, \forall g \in \mathcal{G}$
$\mu^{(t)} \triangleq \{\mu(g)^{(t)}, \mu(b)^{(t)}\}_{g \in \mathcal{G}, b \in \mathcal{B}} = \emptyset$, $\mathcal{K}_b{}^{(t)} = \emptyset$, $t = 0$
Matching step
repeat
  $t \leftarrow t+1$
  for $g \in \mathcal{G}$, propose b according to $\mathcal{P}_g^{(t)}$ **do**
    while $b \notin \mu(g)^{(t)}$ and $\mathcal{P}_g^{(t)} \neq \emptyset$ **do**
      if $q_b^{(t)} \geq l_g^b$ **then**
        $\mu(b)^{(t)} \leftarrow \mu(b)^{(t)} \cup g$
        $q_b^{(t)} \leftarrow \mathrm{q}_b^{(t)} - l_g^b$
      else
        $\mathcal{K}'^{(t)}_b = \{g' \in \mu(b)^{(t)} | g \succ_b g'\}$
        $g_{lp} \leftarrow$ the least preferred g' $\in \mathcal{K}'^{(t)}_b$
        while $\mathcal{K}'^{(t)}_b \neq \emptyset \cup q_b^{(t)} \leq l_g^b$ **do**
          $\mu(b) \leftarrow \mu(b) \setminus \{g'\}$
          $\mathcal{K}'^{(t)}_b \leftarrow \mathrm{K'}^{(t)}_b \setminus \{\mathrm{g}_{lp}\}$
          $q_b^{(t)} = q_b^{(t)} + l_g^b$
          $g_{lp} \leftarrow$ the least preferred g' $\in \mathcal{K}'_{b^{(t)}}$
        end while
        if $q_b^{(t)} \geq l_g^b$ **then**
          $\mu(b) \leftarrow \mu(b) \cup \{g\}$
          $q_b^{(t)} = q_b^{(t)} - l_g^b$
        else
          $g_{lp} \leftarrow g$
        end if
        $\mathcal{K}_b^{(t)} = \{g \in \mathcal{G} | g_{lp} \succ_b g\} \cup \{g_{lp}\}$
        for $k \in \mathcal{K}_b^{(t)}$ **do**
          $\mathcal{P}_g \leftarrow \mathcal{P}_g \setminus k$
        end for
      end if
    end while
  end for
until $\mu^{(t)} = \mu^{(t-1)}$

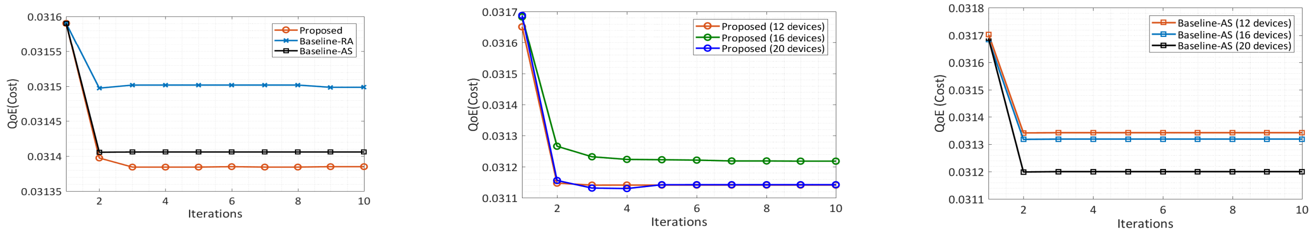
$$\mathbf{P}-\mathbf{p}: \ \underset{\boldsymbol{p}}{\text{minimize}} \quad \mathcal{Q}(\boldsymbol{p}) \tag{14}$$

subject to:

$$p_{\min} \leq p_g \leq p_{\max}, \ g \in \mathcal{G}, \tag{14a}$$

$$\sum_{g=1}^{G} p_g \leq p_{\mathrm{MAX}}. \tag{14b}$$

Problem **P-** is a non-convex optimization. Therefore, we apply BSUM that is based on addition of a proximal term to the objective function and then minimize it [1].



(a) (b) (c)

Fig. 2: (a) QoE (Cost) for various baselines, (b) QoE (cost) for proposed schemes using various metaverse entities, and (c) QoE (cost) for Baseline-AS using various metaverse entities.

IV. Numerical Results

In this section, we present numerical results to show the effectiveness of the proposed scheme. An area of 1000×1000 is used for analysis. Meta spaces are deployed at the network edge servers located at SBS. All values are computed by taking average of many runs (e.g, 40). As our formulated problem is novel and different from the existing works, therefore, we can not directly compared it with existing works. For a fair comparison, we consider baselines. These baselines are *Baseline-A* and *Baseline-R*. *Baseline-A* uses BSUM for power allocation, matching for association, and random resource allocation. For *Baseline-R*, random association, relaxation-based resource allocation, and BSUM-based power allocation are used.

Fig. 2a shows the performance of various schemes in terms of QoE(cost). For the proposed scheme and various baselines, it is evident that proposed scheme outperforms both baselines. This shows the superiority of the proposed scheme. The reason for this improvement lies in joint resource allocation, association, and power allocation. In both baselines, one of the three parameters is chosen random that is why their performance is degraded compared to proposed scheme. Now, we study the effect of varying numbers of metaverse entities on the performance of the proposed scheme. For various number of metaverse entities, the performance of the proposed scheme remains stable. Similar to Fig. 2b, Fig. 2c illustrates the performance of baseline-AS for various number of metaverse entities. It is evident from Fig. 2c that baseline-AS has stable performance for various numbers of metaverse entities.

V. Conclusion

In this paper, we have presented a novel framework for metaverse. We considered QoE for metaverse communication in terms of immersive experience and PER. Our framework offers a tradeoff between immersive experience and PER while computing QoE. We formulated a problem that consider association, resource allocation, and transmit power allocation as optimization variables. For transmit power allocation, we used a BSUM-based solution, whereas for resource allocation, we used a relaxation-based solution. For association, we considered a matching game-based solution. We concluded that our proposed framework and solution can serve as guidelines for various works based on metaverse.

References

[1] L. U. Khan, M. Alghfeli, M. Guizani, N. Saeed, and S. Muhaidat, "Tcs: A joint task offloading, communication, and sensing framework for vehicular metaverse," *Authorea Preprints*.

[2] L. U. Khan, M. Guizani, H. Jung, and M. A. Serhani, "A sampling and communication framework for synchronization in metaverse," in *2024 International Conference on Ubiquitous Networking (UNet)*, vol. 10. IEEE, 2024, pp. 01–05.

[3] W. C. Ng, W. Y. B. Lim, J. S. Ng, Z. Xiong, D. Niyato, and C. Miao, "Unified resource allocation framework for the edge intelligence-enabled metaverse," in *ICC 2022-IEEE International conference on Communications*. IEEE, 2022, pp. 5214–5219.

[4] H. Du, J. Wang, D. Niyato, J. Kang, Z. Xiong, X. Shen, and D. I. Kim, "Exploring attention-aware network resource allocation for customized metaverse services," *IEEE network*, vol. 37, no. 6, pp. 166–175, 2022.

[5] T. Baidya and S. Moh, "Comprehensive survey on resource allocation for edge-computing-enabled metaverse," *Computer Science Review*, vol. 54, p. 100680, 2024.

[6] L. U. Khan, Z. Han, D. Niyato, M. Guizani, and C. S. Hong, "Metaverse for wireless systems: Vision, enablers, architecture, and future directions," *IEEE Wireless Communications Magazine*, 2024.

[7] C.-F. Liu, M. Bennis, M. Debbah, and H. V. Poor, "Dynamic task offloading and resource allocation for ultra-reliable low-latency edge computing," *IEEE Transactions on Communications*, vol. 67, no. 6, pp. 4132–4150, 2019.

[8] H. Du, J. Liu, D. Niyato, J. Kang, Z. Xiong, J. Zhang, and D. I. Kim, "Attention-aware resource allocation and qoe analysis for metaverse xurllc services," *IEEE Journal on Selected Areas in Communications*, vol. 41, no. 7, pp. 2158–2175, 2023.

[9] P. Si, J. Zhao, H. Han, K.-Y. Lam, and Y. Liu, "Resource allocation and resolution control in the metaverse with mobile augmented reality," in *GLOBECOM 2022-2022 IEEE Global Communications Conference*. IEEE, 2022, pp. 3265–3271.

[10] Y. Han, D. Niyato, C. Leung, C. Miao, and D. I. Kim, "A dynamic resource allocation framework for synchronizing metaverse with iot service and data," in *ICC 2022-IEEE International conference on Communications*. IEEE, 2022, pp. 1196–1201.

[11] Z. Long, H. Dong, and A. El Saddik, "Human-centric resource allocation for the metaverse with multiaccess edge computing," *IEEE Internet of Things Journal*, vol. 10, no. 22, pp. 19 993–20 005, 2023.

[12] O. Hashash, C. Chaccour, W. Saad, K. Sakaguchi, and T. Yu, "Towards a decentralized metaverse: Synchronized orchestration of digital twins and sub-metaverses," *arXiv preprint arXiv:2211.14686*, 2022.

[13] J. N. Njoku, C. I. Nwakanma, G. C. Amaizu, and D.-S. Kim, "Prospects and challenges of metaverse application in data-driven intelligent transportation systems," *IET Intelligent Transport Systems*, vol. 17, no. 1, pp. 1–21, 2023.

[14] L. Feng, X. Jiang, Y. Sun, D. Niyato, Y. Zhou, S. Gu, Z. Yang, Y. Yang, and F. Zhou, "Resource allocation for metaverse experience optimization: A multi-objective multi-agent evolutionary reinforcement learning approach," *IEEE Transactions on Mobile Computing*, 2024.

[15] P. Yang, T. Q. Quek, J. Chen, C. You, and X. Cao, "Feeling of presence maximization: mmwave-enabled virtual reality meets deep reinforcement learning," *IEEE Transactions on Wireless Communications*, vol. 21, no. 11, pp. 10 005–10 019, 2022.

[16] J. A. Rassen, A. A. Shelat, J. Myers, R. J. Glynn, K. J. Rothman, and S. Schneeweiss, "One-to-many propensity score matching in cohort studies," *Pharmacoepidemiology and drug safety*, vol. 21, pp. 69–80, 2012.