MULTIMEDIA COMMUNICATIONS TECHNICAL COMMITTEE
http://www.comsoc.org/~mmc

MMTC Communications - Frontiers

Vol. 19, No. 3, May 2024

CONTENTS

SPECIAL ISSUE ON Quality of Experience in the Metaverse

Guest Editors: Francesca Marcello[1] and Luca Resti[2]
[1]University of Cagliari, Italy, francesca.marcello@unica.it
[2] University of York, UK, luca.resti@york.ac.uk

The Metaverse, refers to a collective virtual shared space that is persistently online and active, created by the convergence of virtually enhanced physical reality and physically persistent virtual reality. It encompasses augmented reality, virtual reality, and the internet as we know it, all within a 3D virtual universe.

As the next frontier in multimedia telecommunications technology, the Metaverse promises to revolutionize how we communicate and interact. It aims to create a space where users can interact with a computer-generated environment and other users in real-time. This is not limited to text or voice communication but extends to full-body avatars that can express non-verbal cues, creating a more immersive and expressive form of communication. In the Metaverse, communication isn't just about exchanging information; it's about sharing experiences. Users can explore digital worlds together, participate in interactive events, collaborate in shared workspaces, or socialize in virtual communities. This level of interaction can create a sense of presence and social connection that surpasses any existing communication technology. However, the development of the Metaverse also presents new challenges. Ensuring a high Quality of Experience (QoE) is crucial for user engagement. This includes not only technical aspects such as latency, resolution, and frame rate but also factors like user interface design, accessibility, and safety measures.

Thus, the Special Issue (SI) on Quality of Experience in Metaverse Communications aims to face innovative approaches to realize robust, secure and interactive Metaverse environments, making the communications through the Metaverse more user friendly and network-efficient.

The first paper titled "*Introducing MetaInterTwin - A Novel Architecture for Metaverse and IoT Integration*" proposes a novel architecture designed to bridge the gap between Internet of Things (IoT) devices and the metaverse. Sensor data collected from IoT devices such as thermostats and fitness trackers can be used to produce a digital counterpart of the IoT device within the metaverse known as a Digital Twin (DT). However, this requires data from each physical device to be synchronized for each of its DTs. Additionally, DTs may exist in separate, cloud-based instances of the metaverse, known as subverses, meaning that an efficient and low-latency synchronization approach is required to maintain a high QoE for users.

The authors propose MetaInterTwin as a virtualization layer for efficiently managing the DTs across different subverses. It is shown that the proposed approach greatly reduces both latency and device CPU load, when compared to processing the synchronization entirely on the device.

Next, in "*Metaverse in Tourism: Enabling Technologies and Challenges through the Sardinia Use Case*" the authors explore the potential found in the application of the metaverse paradigm to tourism, and the related networking, computing and ethical challenges that arise with it. The Italian island of Sardinia is presented as an exemplary case study, illustrating how immersive technologies can enable remote virtual tours using Virtual Reality (VR) headsets, and enhance in-person tours via Augmented Reality (AR).

The authors explore the conflict between the need for detailed, historically accurate virtual environments and the practical limitations imposed by the limited computing capabilities of low-power wearable devices. The challenge is further compounded by the complex distributed network infrastructure needed to guarantee a high Quality of Service (QoS) and QoE in remote historical sites. Finally, the regulatory and ethical challenges of metaverse-enabled tourism are illustrated, as this innovative technology also carries the risk of disrupting the balance of the existing tourism ecosystem.

Concluding the round-up of papers for this special issue is "*Robust NeRF-based Digital Twins*", which introduces a robust approach for generating digital twins from web-sourced images using Neural Radiance Fields (NeRFs). The use of web-sourced images can lead to inaccuracies in the resulting reconstructed 3D scenes, if the chosen ground-truth data set contains digitally altered or AI-generated material. To address this, the authors have developed an iterative approach that builds and progressively refines coherent 3D models, filtering out data containing signs of potential tampering.

Once the algorithm has converged on a final model, the next challenge lies in transmitting the data to the users. To this end, an adaptive video streaming approach is proposed, which leverages edge nodes for caching and processing images before employing a machine learning-driven radio resource management scheme. The result is an efficient architecture that can adapt to varying bitrate requirements, maintaining both a high QoS and QoE in the process.

Francesca Marcello is an Assistant Professor at the Department of Electrical and Electronic Engineering at the University of Cagliari (Italy). She received her PhD in Electronic and Computer Engineering from the University of Cagliari in February 2023. She is a member of the Net4U Lab (https://net4u.diee.unica.it/) since 2018, a member of the Italian University Consortium for Telecommunications (CNIT) since 2019, and a member of the IEEE since 2019. Her research interests are mainly related to the Internet of Things, Smart Buildings, Human Behaviour Analysis and Quality of Experience. Her activity focuses on human behaviours recognition and prediction inside buildings for energy and comfort management systems and for improving well-being and quality of life.

Luca Resti received a B.Sc. in Engineering of Computing Systems from Politecnico di Milano, and a M.Sc. in Information and Communications Engineering from the Friedrich-Alexander University of Erlangen-Nürnberg. After working as an associate researcher in the Conversational AI Research group at the Fraunhofer Institute for Integrated Circuits (IIS), he is currently a PhD candidate at the department of Computer Science at the University of York in the UK. His research – which is part of the Sound Interactions in the Metaverse doctoral training program at the AudioLab in collaboration with EA SEED - explores how procedural audio and interactive elements in XR applications affect quality of experience, with the goal of developing novel audio quality assessment methodologies.

Introducing MetaInterTwin - A Novel Architecture for Metaverse and IoT Integration

Simone Porcu[1], Claudio Marche[2]
[1]DIEE, University of Cagliari, simone.porcu@unica.it
[2]DIEE, University of Cagliari, claudio.marche@unica.it

Abstract

The Metaverse has seen significant interest and development in recent years, with major tech companies, such as Meta and Apple, leading the charge. However, a key challenge remains: integrating real-world Internet of Things (IoT) devices into the Metaverse and enabling interaction within this virtual space. In this sense, this research presents "MetaInterTwin", an innovative architecture designed to bridge the gap between the IoT and the Metaverse. Specifically, the proposed architecture introduces a novelty virtualization layer, which ensures real-time synchronization between the physical device and its digital counterpart, optimizing resource consumption and reducing latency across multiple metaverses, namely subverses. The Quality of Experience Influence Factors are furthermore analyzed to evaluate the architecture performances while managing different IoT objects, and a QoE objective model is proposed to understand how the architecture can maintain good performance in terms of perceived quality.

Keywords: Metaverse, Internet of Things, Quality of Experience.

1. Introduction

The past decade has grown interest in Virtual Reality (VR) and Augmented Reality (AR) technologies. These immersive technologies have undergone significant advancements in the last two years, laying the groundwork for the revolutionary concept of the Metaverse. The Metaverse is a collective virtual shared space that is persistently online and active. It results from the convergence of virtually enhanced physical reality and physically persistent virtual reality, encompassing augmented reality, virtual reality, and the internet as we know it, all within a 3D virtual universe.

The concept of the Metaverse has sparked widespread interest and curiosity. The most significant initiative in this regard has been undertaken by the company formerly known as Facebook, which has rebranded itself as Meta and is investing heavily in the realization of its concept of Metaverse. Meta has proposed its vision of the Metaverse, introducing software and headsets to facilitate immersion into this virtual universe. In 2023, Apple followed the case, launching its own VR/AR headset and presenting its interpretation of the Metaverse. However, in general, there is a noticeable gap in information available to end-users about the operational methodology of the Metaverse. Specifically, it remains unclear whether users can integrate their Internet of Things (IoT) devices, such as smartphones and home automation sensors, into the Metaverse environment and interact with them within this virtual space.

Given the lack of information on this aspect, there is a need to understand how users perceive the Quality of Experience (QoE) when using headsets in a Metaverse scenario that incorporates real-life IoT objects. To this, the QoE is defined as the degree of delight or annoyance of the user of an application or service. It results from the fulfillment of his or her expectations with respect to the utility and/or enjoyment of the application or service in the light of the user's personality and current state [1]. The way the QoE allows to monitor the perceived quality is the most suitable to lead the developing a Metaverse infrastructure, but there is still the need to understand which parameter must be monitored to understand the perceived quality.

In recent years, the IoT has been introduced into the context of the Metaverse, enabling it to analyze and interact with the real world and to provide users with immersive mixed-reality environments. In this sense, the Metaverse helps map the real-time IoT data from the physical reality into the virtual world, acting as a supplement for the experiential interface of the users [2]. Several scenarios and applications have seen improvements due to the use of IoT in the Metaverse. One of the most important is represented by the healthcare scenario, in which all the IoT medical devices attached to the users' body elicit a response in the virtual world and can be used to track and sense personal body information [3]. Moreover, another application is depicted by the educational scenario, where concepts of visualization learning are increasingly being adopted to revolutionize traditional teaching methods and create immersive environments for students [4]. However, from these works, it is not clear how to make interoperable the IoT objects within the different Metaverses that have been released in the last year.

Moved by this reason, this research paper aims to explore this not-enough-explored territory, providing an innovative architecture named "MetaInterTwin", capable of managing different IoT objects and making them interoperable between different metaverses. The QoE Influence Factors (IFs) are analyzed to evaluate the architecture performances while managing different IoT objects that belong to different metaverses.

2. Metaverse architecture: the Subverse

The Metaverse is described as a "post-reality universe", a continuous and enduring multi-user environment that blends physical reality with digital virtuality. This concept is built on the convergence of technologies that facilitate multisensory interactions with virtual environments, digital objects, and individuals, such as Virtual Reality (VR) and Augmented Reality (AR) [5]. According to this definition, the Metaverse is prototyped as a singular virtual space where various users can interact with each other or perform tasks alone or together. However, individual companies are developing their own interpretation of the Metaverse, which contradicts the term "universe" that implies the existence of only one Metaverse. According to this, we propose a new terminology: we refer to these individual metaverses as "subverses" and use the term "Metaverse" to denote the collective universe encompassing all these subverses. As reported in Figure 1, the Metaverse is the container entity that allows to operate different subverses. As the Figure 1 highlights, the subverses are not able to interact with each other. This particular choice has been made since the architectural differences of each subverse make interoperability impossible.

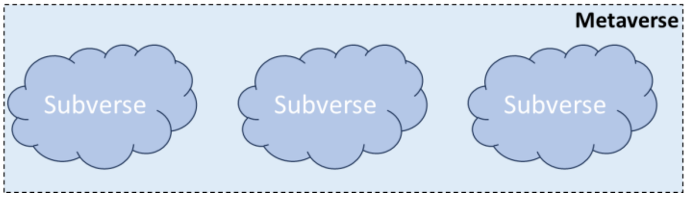


Figure 1 – Metaverse containing various subverses.

3. Scenario

In the scenario reported in Figure 2, we consider a subject who possesses *n* IoT objects. These objects range from smart home devices like thermostats and refrigerators to wearable technology like fitness trackers. The subject's goal is to integrate these IoT objects into various subverses, thus creating different Digital Twin (DT) copies, one for each subverse. These IoT objects have the ability to collect and analyze data, and they can be remotely monitored and controlled. Thus, the data collected by the IoT objects could be used to influence the state of each subverse. For example, if the subject's smart thermostat collects data about the subject's preferred room temperature, this data could be used to adjust the temperature in the subject's virtual home in one subverse. However, the temperature must also change in the other subverses. This example is explanatory of the interoperability problems that need to be addressed. To this, a robust Metaverse architecture would need to be developed to manage the interoperability between IoT objects and the different subverses.

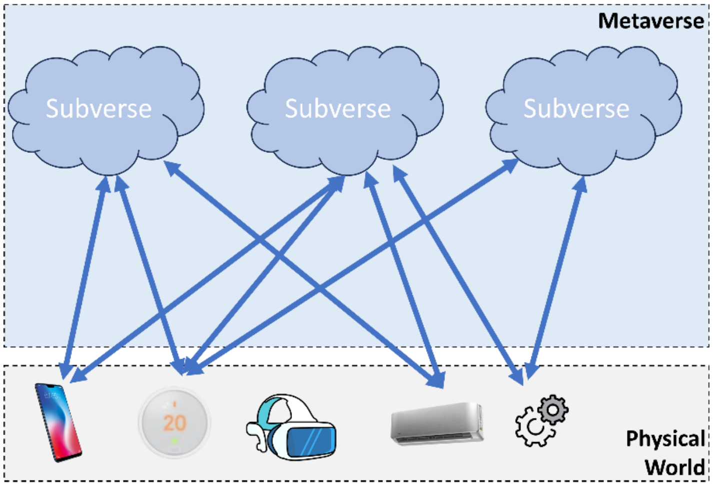


Figure 2 – Scenario.

According to this scenario, in the following sections, we are going to introduce our proposed solution that contains a virtualization layer in charge to manage the DT of each IoT object. The virtualization layer manages the information exchanged between each DT and each subverse. To effectively test the efficiency of the virtualization layer, we compared latency and resource usage, during the information exchange, using our proposed solution and interfacing directly the IoT object to the *n* subverses. The performances were also evaluated in terms of perceived QoE.

4. Proposed Solution

The IoT and Metaverse cooperation has brought benefits in many important scenarios, such as entertainment, socialization, and smart environments. For this reason, this section aims to propose our general mixed IoT and Metaverse architecture, highlighting the importance of the proposed Virtualization Layer, used to share all the digital counterparts (Digital Twins), and making them available to the different subverses.

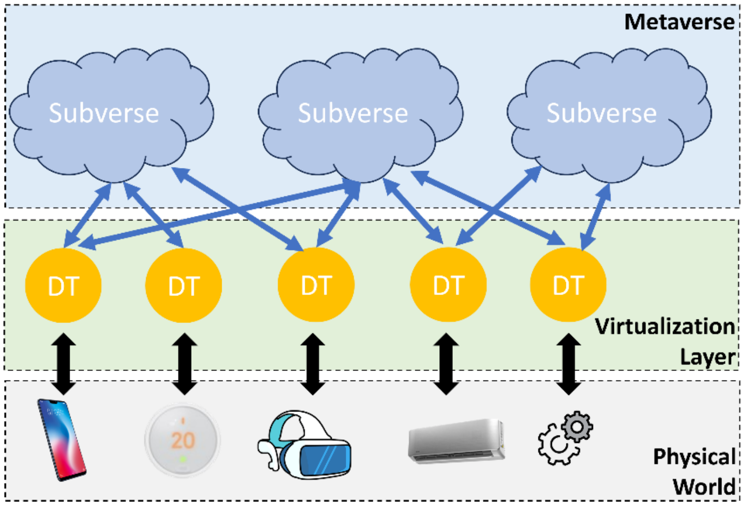


Figure 3 – "MetaInterTwin" architecture.

MetaInterTwin consists of a three-layer architecture, as illustrated in Figure 3: Physical World, Virtualization layer and Metaverse. The bottom layer comprises devices, such as smartphones or virtual reality headsets, tasked with sensing the physical environment and providing data to the layers above. The DTs serve as digital representations of real-world entities enhanced with virtual capabilities, detailing their characteristics and the services they can offer. The digital entity can be part of different subverses environments, letting the DT be editable by each subverse. The proposed Virtualization layer is a management layer that makes the various DTs interoperable through the subverses, providing efficient resource utilization and latency reduction across multiple subverses. Moreover, this layer ensures real-time synchronization between the physical device and its digital counterpart, also sharing the device state changes with the different subverses. Consequently, all subverses are promptly informed of these changes, eliminating the need for redundant queries and updates. This not only minimizes latency but also optimizes resource consumption, enabling a more seamless and responsive user experience across various subverses. Finally, the upper layer, i.e. the Metaverse, completely installed on the Cloud, serves as a virtual universe, encompassing a multitude of proprietary subverses. Each subverse owned and managed independently, offers different digital experiences and realities. However, it also necessitates efficient synchronization mechanisms to ensure seamless interactions and transitions between different subverses. The MetaInterTwin goal lies exactly in solving the previous concerns, providing a solution to the subverses interoperability while managing IoT objects.

5. Performance Evaluation

As explained before, the QoE introduce the possibility to obtain a subjective measure that captures the perceived quality of using an application or a service [1]. Its subjectivity enables the development of objective quality models based on monitored IFs. However, these IFs vary depending on the application or system under observation [6]. In the context of the Metaverse, it is not properly defined which are the IFs that have to be considered to monitor the perceived QoE.

Therefore, this section aims to identify the parameters that can be considered to analyze the QoE in a Metaverse/IoT scenario. These parameters will provide a more comprehensive understanding of the user experience within this increasingly interconnected digital landscape. Finally, the evaluation of a reference scenario illustrates how the proposed architecture improves the performance of managing different IoT objects with respect to a classical IoT/Metaverse solution.

5.1 Influence Factors

Traditional influence factors, such as human, system, and context factors, play a significant role in shaping the QoE. These include the user’s characteristics, the performance of the system, and the user’s environment. In order to analyze and shape a service based on the QoE, understanding these factors is crucial for optimizing the user experience in the Metaverse.

Traditional IFs are those that affect the QoE of any multimedia-based application or service, such as the Metaverse but they can be considered as well for the IoT context.

They include:

- Human factors: These are the characteristics of the user, such as sex, age, affective state, and discomfort symptoms (e.g., cybersickness).
- System factors: These are the aspects related to the network infrastructure, the hardware, and the multimedia content. These represent the various Quality of Service (QoS) factors, which are popular in the multimedia context. For instance, network QoS factors like packet loss, latency, and jitter have effects on QoE for IoT services.
- Context factors: These are the factors that describe the user’s environment, such as physical, temporal, social, economic, task, and technical characteristics. They affect the user’s expectations, preferences, and attention.

According to [7], it is clear that the quality of a multimedia IoT context can be evaluated using the above mentioned IFs. However, it is important to note that the rapid evolution of IoT is making way for the development of several IoT applications that require minimal or no human involvement in the data collection, transformation, knowledge extraction, and decision-making process. Therefore, evaluating their quality is challenging in the absence of any human involvement or feedback.

According to this, the authors in [8] highlighted that the System IFs are the most impactful, and are those with less interaction required from the user. Thus, the next session is focused on the evaluation of the performance of our proposed architecture based on the System IFs, in order to understand how this solution can outperform the common IoT approaches. Moreover, a QoE objective model is proposed to deeply understand how the proposed architecture can improve the user perceived quality while managing various IoT objects into different subverses.

5.2 Evaluation of a Reference Scenario

This section illustrates the enhancement in performance management of IoT objects within the Metaverse, achieved through the proposed architecture, compared to a conventional IoT/Metaverse solution. According to the previous explained scenario, we set a test where an IoT object, represented by a Raspberry Pi 3 Model B+, is interfaced with diverse environmental sensors, such as a DHT22 for capturing temperature and humidity data, and a PIR sensor to detect the presence of people. This device communicates and transmits the sensed data to the Metaverse. Specifically, we implement two distinct approaches. In the first approach, i.e. the Classical Approach, the device communicates directly with all the subverses, while in the second approach, which employs the proposed architecture, we design the novel virtualization layer that interfaces its managed DTs with all the subverses. This innovative layer serves as a conduit for data exchange across the subverses, enhancing the overall efficiency and performance of the system. As follows, we analyze the communication performances by monitoring the cumulative latency and the CPU Load as a resource usage parameter.

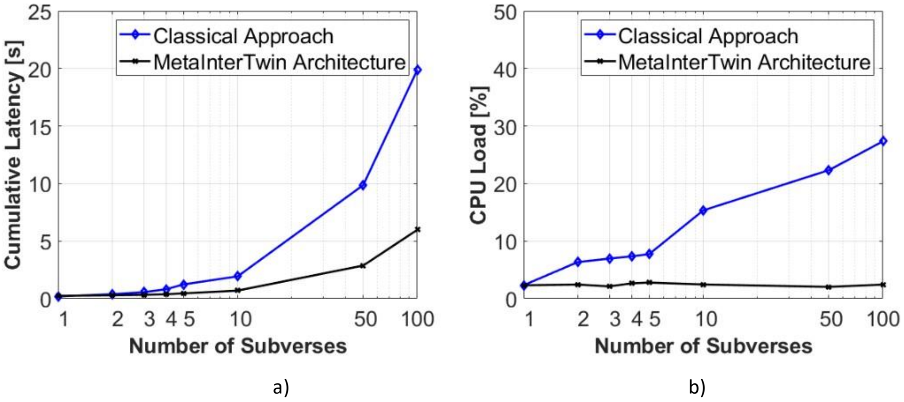


Figure 4 – a) Latency and b) Device CPU Load while running the Classic Approach and the Proposed Solution.

Considering this setup, Figure 4 illustrates how the proposed solution overcomes the classical approach in terms of both performances. Specifically, as illustrated in the left graph, our architecture is able to outperform the classical approach in terms of cumulative latency, expressed as the time needed to update all the subverses. This is due to the fact that in the classical approach, the device is responsible for updating all its Digital Twins in each subverse, and as a result, by incrementing the number of subverses, the time needed to update all the virtual counterparts increases exponentially. On the other hand, in the proposed architecture, the device communicates only with its unique DT, placed in the Cloud, which manages the communication with all the subverses. Moreover, as illustrated in the right graph, for a growing number of subverses, the CPU load of the device increases since it is responsible for updating all the DTs in each subverse. Meanwhile, in the proposed solution, the CPU load is constant, and the number of subverses does not affect the device resources.

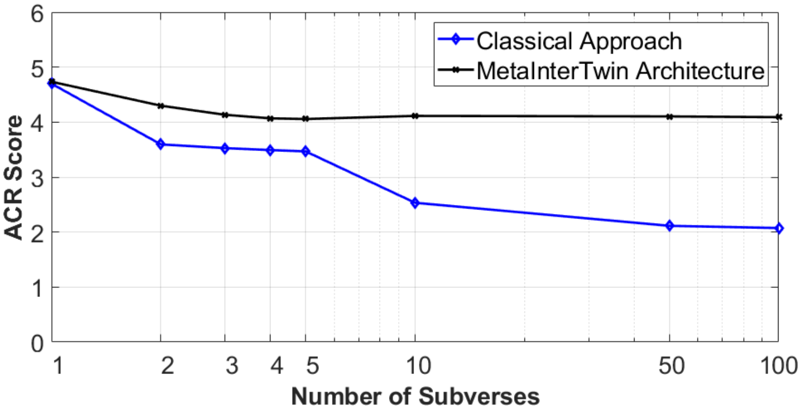


Figure 5 QoE estimation based on latency and usage of resources.

To understand how the quality changes when increasing the communication with the subverses, we propose an objective QoE modelling function, to understand the gain in terms of quality of using our proposed approach. The objective function maps the quality according the ACR scale (from 1 to 5 where 1 means bad quality and 5 means excellent quality) and it is described according to the following equation:

$$QoE = 0{,}06 * Latency - 0.12 * CPU\ load + 4{,}38$$

Comparing the estimated values of QoE for both approaches, is clear how much stable and performant is the proposed architecture. It is notable that even increasing the number of subverses, the perceived quality slightly decreases without going below a medium quality of 3. To evaluate these results, in future works we will execute a subjective test to validate the performance of the proposed architecture.

6. Conclusions

In conclusion, this research introduces the "MetaInterTwin" architecture, addressing the challenge of integrating IoT devices into the Metaverse. Specifically, to optimize the perceived quality, especially considering the IoT-related IFs, such as latency and CPU load, the implementation of a novelty virtualization layer is proposed. The proposed layer enables real-time synchronization between physical IoT devices and their digital counterparts across multiple sub-metaverses. Simulations demonstrate enhanced performance compared to a classic IoT/Metaverse approach. The three-layer architecture of the architecture, comprising the Physical World, Virtualization layer, and Metaverse, proves better results in terms of resource utilization and reducing latency, providing an improved QoE.

References

[1] A. Perkis *et al.*, ‘QUALINET White Paper on Definitions of Immersive Media Experience (IMEx)’, 2020, doi: 10.48550/ARXIV.2007.07032.
[2] K. Li *et al.*, ‘When Internet of Things meets Metaverse: Convergence of Physical and Cyber Worlds’, 2022, doi: 10.48550/ARXIV.2208.13501.
[3] V. Shah and A. Khang, ‘Metaverse-Enabling IoT Technology for a Futuristic Healthcare System’:, in *Advances in Computational Intelligence and Robotics*, A. Khang, V. Shah, and S. Rani, Eds., IGI Global, 2023, pp. 165–173. doi: 10.4018/978-1-6684-8851-5.ch008.
[4] X. Chen, Z. Zhong, and D. Wu, ‘Metaverse for Education: Technical Framework and Design Criteria’, *IEEE Trans. Learn. Technol.*, pp. 1–12, 2023, doi: 10.1109/TLT.2023.3276760.
[5] S. Mystakidis, ‘Metaverse’, *Encyclopedia*, vol. 2, no. 1, pp. 486–497, Feb. 2022, doi: 10.3390/encyclopedia2010031.

[6] S. Porcu, A. Floris, and L. Atzori, ‘Quality of Experience in the Metaverse: An Initial Analysis on Quality Dimensions and Assessment’, in *2022 14th International Conference on Quality of Multimedia Experience (QoMEX)*, Lippstadt, Germany: IEEE, Sep. 2022, pp. 1–4. doi: 10.1109/QoMEX55416.2022.9900897.

[7] A. Floris and L. Atzori, ‘Managing the Quality of Experience in the Multimedia Internet of Things: A Layered-Based Approach’, *Sensors*, vol. 16, no. 12, p. 2057, Dec. 2016, doi: 10.3390/s16122057.

[8] K. Fizza *et al.*, ‘QoE in IoT: a vision, survey and future directions’, *Discov. Internet Things*, vol. 1, no. 1, p. 4, Dec. 2021, doi: 10.1007/s43926-021-00006-7.

Authors Bio-data

Simone Porcu is an Assistant Professor at the Department of Electrical and Electronic Engineering at the University of Cagliari (Italy). He received a bachelor's degree in computer science from the University of Cagliari, Italy, in 2015, a master's degree in computer science in 2017 and a Ph. D. in Electronic and Computer Engineering from the University of Cagliari in 2021. He also obtained the Doctor Europaeus title for his research period abroad at the Technische Universitat Berlin (September 2018-March 2019). Simone has been a member of the Net4U Lab (https://net4u.diee.unica.it/) at the University of Cagliari since 2017, a member of the Italian University Consortium for Telecommunications (CNIT) in 2018, and a member of the IEEE from 2018.

Claudio Marche is an Assistant Professor at the University of Cagliari, where he is a member of the Net4U research group. He received his PhD in Electronic and Computer Engineering in 2023 and the M.Sc. degree in Telecommunication Engineering with full marks in 2018 at the University of Cagliari. Since graduation, he has been working as a Researcher in the Department of Electrical and Electronic Engineering at the University of Cagliari. In 2021, he has been a visiting student at the Inria Centre, in Lille (France). Currently, he is a member of the editorial board for the Frontiers in Energy Efficiency Journal. His current research interests include the Internet of Things, Social Networks and Trust Management techniques.

Metaverse in Tourism: Enabling Technologies and Challenges through the Sardinia Use Case

Kien Nguyen[1], *Md Sabbir Ahmed*[2], *Claudia Carballo Gonzalez*[3], *Chiara Cattani*[4], *Kaleem Arshid*[5]
[1] *Department of Computer Science, University of Würzburg, Würzburg, Germany, kien.nguyen@uni-wuerzburg.de*
[2] *Department of Information Engineering, University of Pisa, Pisa, Italy, sabbir.ahmed@dii.unipi.it*
[3] *Department of Electrical and Electronic Engineering, University of Cagliari, Sardinia, Italy, claudia.carballogonz@unica.it*
[4] *Department of Social and Political Sciences, University of Cagliari, Sardinia, Italy, c.cattani@studenti.unica.it*
[5] *Department of Marine, Electrical, Electronic and Telecommunications Engineering, University of Genova, Italy, kaleem.arshid@edu.unige.it*

Abstract

The metaverse is a virtual world with the potential to revolutionize tourism. Tourism stakeholders may attract tourists from all over the world by establishing an authentic and immersive virtual place in the world, allowing them to experience that place's past without ever leaving their homes, and preserving and promoting the culture and legacy of the place for future generations. However, numerous obstacles must be overcome before the metaverse may be fully realized in that place. Content generation, technical execution, and regulation are among the hurdles. This article outlines two prospective metaverse use cases for deployment, describes some of the constraints that must be overcome, and examines several enabler technologies that can help to overcome these challenges.

Keywords: Metaverse, Edge AI, Sustainable Tourism, Extended Reality

1. Introduction

5G and beyond networks will enable a wide range of advancements in media service delivery [1], [2]. In such a context, the future of communication technology appears to be centered around immersive communications, encompassing extended reality (XR), haptic, and holographic applications [3]. Then, with the transition from two-dimensional (2D) to three-dimensional (3D) displaying methods and the addition of enhanced multisensorial content, the lines between the physical and virtual worlds will become increasingly blurred. As a result, immersive 3D communications are expected to significantly impact the communication industry and our lifestyle in the coming years, paving the way for the metaverse paradigm [4].

The metaverse is defined as “the post-reality universe, a perpetual and persistent multiuser environment merging physical reality with digital virtuality. It is based on the convergence of technologies that enable multisensory interactions with virtual environments, digital objects, and people, such as virtual reality (VR) and augmented reality (AR)” [5].

The metaverse applications provide a virtual world where people can interact with the digital environment and other users in real-time [6]. These applications are characterized by diverse key performance indicators (KPIs), imposing tight quality of service (QoS) requirements in terms of ultra-high reliability, data rate, energy efficiency, scalability, and very low latency. Then, the end-to-end (E2E) network deployment must be dotted with enough capabilities to deliver high quality immersive applications and not impair user quality of experience (QoE). Integrating technologies like unicast/multicast/broadcast delivery [7], slicing [8], edge-cloud computing, big data, and artificial intelligence (AI) [9] is crucial to making the metaverse possible. Moreover, the cooperation of terrestrial and non-terrestrial networks (TNs-NTNs) must be effectively managed to satisfy the “always best connected” (ABC) paradigm with adequate differentiated traffic management [10].

On the other hand, in addition to the technical challenges previously mentioned, the metaverse environment presents several issues related to security, privacy, ethics, and content creation, among others [11], [12]. Consequently, academia, industry, and regulatory entities must work together to identify and handle these challenges effectively.

Along with the evolution of the metaverse, new capabilities and groundbreaking services will emerge, directly impacting the entertainment, health, industry, educational, transport, and tourism sectors. Particularly, the metaverse offers new avenues for sustainable tourism [13], addressing accessibility challenges in showcasing historical sites. Sustainable tourism is crucial in managing resources and communities. Efforts include carrying capacity studies, reducing footprints, promoting ecotourism, and environmental protection [14]. Additionally, reduced visitation during the COVID-19 pandemic impacted tourism, revealing potential opportunities for sustainable tourism [13].

For travel and tourism customers, entering metaverse platforms connects their minds with the virtual worlds as if they are in an airplane, a hotel, or a particular destination [15], [16]. The metaverse does not intend to replace physical travel. It creates a desire to travel and enhances tourists’ real experience on-site or remotely.

This short paper focuses on the metaverse and its potentialities for tourism, enabling technologies, and challenges. To comply with our goal, we propose deploying two metaverse use cases on Sardinia. This Italian island must be discovered by travelers worldwide, avoiding seasonal overtourism [17]. The Sardinian community needs a solution to attract tourists by showing archaeological sites, ancient traditions, and breathtaking landscapes in attractive and sustainable ways. In such a context, the metaverse is a great opportunity to approach Sardinia’s huge and fascinating heritage. Metaverse for travelers is a time-travel journey that can be explored and that can take users through different eras, from the Nuragic age to the Roman Empire and even to the present day.

The remainder of the paper is structured as follows. Section II presents two metaverse use cases for sustainable tourism in Sardinia. Section III examines some of the enabling technologies to implement the proposed metaverse applications. Then, Section IV discusses content, technical, and regulation challenges for the metaverse tourism implementation. Finally, Conclusions are drawn in Section V.

2. Looking inside Sardinia: Use Cases

The tourism in Metaverse project consists of two significant use cases aimed at engaging users both inside and outside Sardinia as shown in Figure 1.

The first scenario involves remote exploration through VR technologies: individuals physically distant from Sardinia can have an immersive experience in the metaverse, overcoming economic or health limitations and simultaneously stimulating audience engagement. The local tourism system ‘Sardegna Turismo’ integrates the access link to this platform, allowing users to explore the region easily and completely free of charge. This user-friendly experience is accessible directly from the browser without any expensive VR devices and any kind of disorders such as motion sickness related to them. On the other side, they can opt for a more immersive experience using a personal third-party VR headset like Oculus or a complete VR set with sensors. The platform offers a range of location and historical period options; users can select their preferences to travel through time and space, experiencing an interactive and engaging journey. A virtual guide interacts directly with participants, enhancing the experience by providing real-time information and answering questions. This is an innovative way to explore the environment and history of Sardinia. It is further an opportunity to stimulate real tourism, encouraging users to physically visit the places they have explored virtually.

The second scenario involves AR technologies: tourists physically visiting Sardinia can enhance their exploration. This service is directly accessible on-site at an archaeological zone, a natural area, or a village. Local guides offer a more innovative and engaging tour by overlaying digital information and 3D content onto the real space. Moreover, the use of VR headsets or personal smartphones allows a more inclusive approach, enabling participants to tailor the experience to their preferences. For example, some people may prefer the use of VR headsets for complete immersion, while others may choose for their smartphones for greater flexibility. Visitors can explore the area with the support of a local guide or independently for a customized experience. The visitor is no longer a passive listener but actively involved in the tourist tour by interacting directly with the context enriched with 3D content and educational activities. Exploring places with AR makes the experience stimulating and memorable, allowing users to interact with the real context, creating a seamless synergy between the digital and physical worlds and reducing the motion sickness disorder associated with VR headsets.

In both cases, the main goal is promoting Sardinia, its territory, and its historical-cultural heritage in the metaverse platform through VR and AR technologies: an innovative and accessible way for delivering an unforgettable and sustainable tourist experience.

3. Enabling Technologies

The E2E network deployment must be dotted with enough capabilities to enable the proposed metaverse use cases. In such a context, storage capacity, computational complexity, data movement cost, power consumption, and privacy issues are critical concerns. Additionally, the slicing paradigm arises as an effective technology to add flexibility, prioritization, and isolation by creating several logical network slices (NSs) utilizing a shared physical infrastructure. Consequently, the metaverse applications can be accommodated in diverse slices and enjoy differentiated traffic management regarding QoS constraints and network load.

As Figure 1 shows, we consider a heterogeneous radio access network (RAN) infrastructure to deliver the proposed applications effectively. For example, the users accessing Sardinian tourist sites through the first use case can use the TN connectivity (e.g., Wi-Fi access point (AP) or macro-base station (BS)). They must have a Service Level Agreement (SLA) that guarantees fulfilling the latency, throughput, and other QoS requirements for an adequate user perception.

For the second use case, the TN-NTN combination is crucial to achieve a satisfactory QoE and to optimize the radio resources following the ABC paradigm. Particularly, several archaeological sites can be situated in remote zones where the TN deployment is unaffordable. Then, unmanned aerial vehicles (UAVs) acting as aerial BSs represent an economical solution to provide the service thanks to their flexibility and rapid deployment. Moreover, the metaverse application can be mapped into a multicast slice to simultaneously satisfy many users located in the same area through effective resource management.

On the other hand, we aim to leverage multi-access edge computing (MEC) to process our metaverse tasks locally. The primary objective is to reduce E2E latency and develop a solution that is both cost-effective and sustainable. Our solution utilizes existing local resources, such as desktops or laptops, to host the MEC, which can perform image processing and deep learning tasks, including model training and inference. By utilizing MEC, computational tasks are offloaded from the remote cloud to the edge, resulting in reduced data transmission latency and enhanced real-time data processing capabilities through the integration of deep learning. Additionally, the study focuses on cost-effectiveness by utilizing existing local resources, reducing dependency on expensive cloud infrastructure. Sustainability is also prioritized through a reduction in energy consumption via localized computing and decreased network bandwidth requirements.

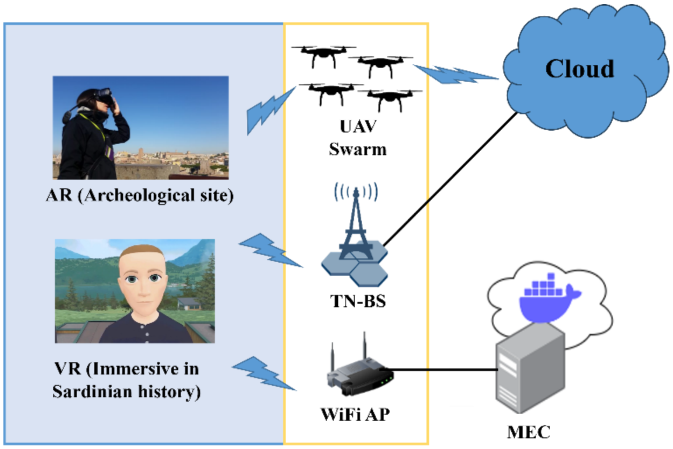


Figure 1. Metaverse traveler's use cases

4. Challenges

To realize our vision of a virtual and immersive metaverse for the use case of Sardinia, two primary challenges relating to content and technical implementation are mentioned in the following.

4.1 Integrity versus Immersion

From a non-technical point of view, the success of historical-specific or virtual-world experiences in general must pay attention to how immersive the users' experience is. Immersion is defined through three key concepts: sight, sound, and touch [18]. While the last one has not yet been technically realized, the first two have been demonstrated widely in the gaming industry. In this area, historical perspectives have been considered and utilized widely. A clear example comes from the series Assassin's Creed from Ubisoft, of which many medieval and ancient cities/locations (e.g., Paris in 1798, London in 1868, Egyptian B.C) had been reconstructed with relatively historical accuracy. To achieve that, the company has worked intensively with more than 70 archaeologists and guided tours covering a vast range of aspects of life and ritual at the time [19]. Partly inspired by Ubisoft's approach, the DECIMA project at the University of Toronto has created a fully interactive 3D urban map of Florence, Italy, in the 16th century [20].

Several factors that contribute significantly to immersion in gaming, such as ancient human accents and dialects, old music, and extinct animal sounds, may pose challenges for adoption in virtual tourism. The gaming and tourism industries have different purposes and target audiences, making it difficult to align tasks between them. Gaming focuses on entertainment and often does not claim to be historically accurate, giving developers the freedom to creatively fill in missing data from history, especially ancient history. Additionally, gamers are often young people who prioritize entertainment over historical accuracy. In contrast, tourism is closely tied to a country's culture and politics [21] and, therefore, must strictly protect the integrity of information. Tourists have a wide range of motivations, including entertainment and education. Inaccurate historical information can mislead people's understanding of a culture.

Crafting an immersive virtual world for tourism poses a unique challenge in preserving historical integrity. While the gaming industry's expertise in virtual world creation can serve as a solid foundation, historical data must be handled with utmost care and presented accurately, authentically, and engagingly. This requires close collaboration with archaeologists and historians to ensure that the virtual world faithfully represents the past without compromising historical accuracy or authenticity.

4.2 Technical implementation

Recently, the birth of many graphic engines and rendering platforms, e.g., Unreal Engine 5, Unity, Roblox, and Nvidia Omniverse [22], has enabled computer graphics to be ultrarealistic. However, applying them to VR/XR-enabled tourism faces many challenges from both computing and networking. This is due to the intrinsic of the tourist use case comprising remote location, mobility of users, and the limitation of wear devices. In detail, the computing challenges primarily come from the computing power requirement to render such a detailed environment in real-time with a smooth and high QoE. However, most of the virtual immersive worlds nowadays are rendered through powerful, stationary computers and servers. VR/XR glasses, if applicable, normally connect to these computers, thus offloading the entire processing to the computer. However, this is not applicable to the metaverse where mobility is allowed for users to freely roaming in an augmented world, for example, in a ruined city or excavating historical site. Equipping computing power right at the glasses will meet the challenge of battery power and the device's weight. On the other hand, cloud computing and wireless communication can fulfill the demand for powerful computation while maintaining device weight and mobility. However, the metaverse, which relies on the cloud, is entirely dependent on the communication path between the distant datacenter and local devices. This makes it prone to transmission latency and bandwidth constraints.

One solution to address this challenge is applying edge-cloud/MEC [23], where data processing is performed at nearby micro data centers or stations instead of relying solely on the cloud. This approach reduces latency and bandwidth usage in the core network by minimizing the distance data needs to travel. However, as the Metaverse aims to seamlessly integrate the physical and virtual worlds through digital twins (DTs) [24], which incorporates mixed data formats, such as haptics, videos, human-to-environment interactions, and human-to-human interactions, other stringent requirements need to be realized. This ambitious vision necessitates large bandwidth and ultra-reliable low-latency communication (URLLC) to support immersive content delivery and real-time interaction. To achieve these stringent requirements, reliable networking infrastructure is essential. A comprehensive communication overhaul is needed, encompassing both high-layer protocols and low-layer communication mechanisms. From higher-layer network orchestration, network slicing with dedicated network slices can be allocated for Metaverse traffic, ensuring prioritized access and resource allocation for latency-sensitive applications. Network virtualization with Network functions virtualization (NFV) and software-defined networking (SDN) [25] can be employed to create flexible and agile networks that can adapt to the dynamic demands of the Metaverse. From a low-layer networking perspective, managing diverse mobility patterns within Sardinia's varied geography is a challenge for the metaverse, necessitating adaptable network infrastructure to accommodate users stationary or moving. Employing high-frequency bands (e.g., millimeter–wave and terahertz) is crucial for achieving high data transfer rates and quality VR/AR experiences. However, these bands pose challenges like signal attenuation over distances, signal penetration obstacles, and interference from physical objects, requiring careful consideration for optimal implementation [26].

By addressing these technical challenges, we can pave the way for a truly immersive and seamless Metaverse experience, blurring the lines between the physical and virtual worlds.

4.3 Regulation and social effects

With its immense potential to collect vast amounts of data, including body-based, spatial, and biometric information, the metaverse raises significant privacy and sovereignty concerns [27]. The introduction of digital assets like cryptocurrencies and non-fungible tokens (NFTs) within the metaverse further exacerbates these concerns, as these assets can be exploited for illicit activities such as money laundering [28]. To address these challenges, regulatory measures have been initiated by the Italian government and European authorities [29]. However, the metaverse's nascent and rapidly evolving nature presents a significant obstacle to effective regulation. Ensuring historical accuracy within the metaverse is crucial to prevent misinterpretation or oversimplification of historical narratives. The metaverse's impact on physical tourism is a complex issue. While it could stimulate virtual tourism, there is a risk that it might deter in-person visits, potentially affecting local businesses [21]. Addressing data privacy and ethical considerations, particularly concerning sensitive historical data, demands careful and compliant integration within the metaverse.

5. Conclusions

To create a realistic and immersive virtual Sardinia, advanced 3D modeling and cultural and linguistic authenticity would be required. Furthermore, the content structure must be scalable in order to support expansion and new discoveries. The metaverse has the potential to completely transform tourism in Sardinia. Sardinia may benefit from its tremendous prospects by tackling the difficulties described above and embracing the power of the metaverse.

The findings of our article can help Sardinia establish a metaverse tourism strategy. This strategy can help to keep Sardinia at the forefront of metaverse tourism. These popular uses may allow visitors to tour ancient ruins and historical places without fear of causing damage or disrupting local communities. Furthermore, through interactive encounters and simulations, virtual tourists could learn about Sardinia culture and heritage. Tourists might connect with other travelers and discuss their Sardinia experiences. As a result, the metaverse has the potential to increase the accessibility, sustainability, and inclusiveness of tourism. By implementing a well-thought metaverse tourism strategy, Sardinia may position itself as a leader in this new and interesting industry.

References

[1] R. Liu, R. Y.-N. Li, M. Di Renzo, and L. Hanzo, "A vision and an evolutionary framework for 6G: Scenarios, capabilities and enablers," arXiv preprint arXiv:2305.13887, 2023.
[2] S. A. A. Hakeem, H. H. Hussein, and H. Kim, "Vision and research directions of 6g technologies and applications," Journal of King Saud University-Computer and Information Sciences, 2022.
[3] X. S. Shen, J. Gao, M. Li, C. Zhou, S. Hu, M. He, and W. Zhuang, "Toward immersive communications in 6G," Frontiers in Computer Science, 2023.

[4] M. Xu, W. C. Ng, W. Y. B. Lim, J. Kang, Z. Xiong, D. Niyato, Q. Yang, X. S. Shen, and C. Miao, "A full dive into realizing the edgeenabled metaverse: Visions, enabling technologies, and challenges," IEEE Communications Surveys & Tutorials, 2022.

[5] S. M. Mystakidis, "Encyclopedia 2022, 2, 486–497," 2022.

[6] S. Porcu, A. Floris, and L. Atzori, "Quality of experience in the metaverse: An initial analysis on quality dimensions and assessment," in 2022 14th International Conference on Quality of Multimedia Experience (QoMEX), IEEE, 2022.

[7] C. C. Gonzalez, E. F. Pupo, J. Montalban, S. Pizzi, E. Iradier,´ and M. Murroni, "Hybrid Terrestrial-Airborne Connectivity for Unicast/Broadcast Services Beyond 5G," in 2023 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB), 2023.

[8] C. C. Gonzalez, E. F. Pupo, L. Atzori, and M. Murroni, "Dynamic Radio´ Access Selection and Slice Allocation for Differentiated Traffic Management on Future Mobile Networks," IEEE Transactions on Network and Service Management, 2022.

[9] M. K. Shehzad, L. Rose, M. M. Butt, I. Z. Kovacs, M. Assaad, and´ M. Guizani, "Artificial intelligence for 6G networks: Technology advancement and standardization," IEEE Vehicular Technology Magazine, 2022.

[10] C. C. Gonzalez, S. Pizzi, M. Murroni, and G. Araniti, "Multicasting´ over 6G Non-Terrestrial Networks: A Softwarization-based Approach," IEEE Vehicular Technology Magazine, 2023.

[11] Y. Wang, Z. Su, N. Zhang, R. Xing, D. Liu, T. H. Luan, and X. Shen, "A survey on metaverse: Fundamentals, security, and privacy," IEEE Communications Surveys & Tutorials, 2022.

[12] H. Ning, H. Wang, Y. Lin, W. Wang, S. Dhelim, F. Farha, J. Ding, and M. Daneshmand, "A Survey on the Metaverse: The State-of-the-Art, Technologies, Applications, and Challenges," IEEE Internet of Things Journal, 2023.

[13] H. Go and M. Kang, "Metaverse tourism for sustainable tourism development: Tourism agenda 2030," Tourism Review, 2023.

[14] S. M. Rasoolimanesh, S. Ramakrishna, C. M. Hall, K. Esfandiar, and S. Seyfi, "A systematic scoping review of sustainable tourism indicators in relation to the sustainable development goals," Journal of Sustainable Tourism, 2023.

[15] S.-p. Tsai, "Investigating metaverse marketing for travel and tourism," Journal of Vacation Marketing, 2022.

[16] D. Buhalis, D. Leung, and M. Lin, "Metaverse as a disruptive technology revolutionising tourism management and marketing," Tourism Management, 2023.

[17] N. Kouroupi and T. Metaxas, "Can the Metaverse and Its Associated Digital Tools and Technologies Provide an Opportunity for Destinations to Address the Vulnerability of Overtourism?," Tourism and Hospitality, 2023.

[18] J. D. N. Dionisio, W. G. B. Iii, and R. Gilbert, "3d virtual worlds and the metaverse: Current status and future possibilities," ACM Computing Surveys (CSUR), 2013.

[19] A. Webster, "Assassin's Creed Origins' new educational mode is a violence-free tour through ancient Egypt." https://www.theverge.com/2018/2/20/17033024/ assassins-creed-origins-discovery-tour-educational-mode-release.

[20] E. L. Pecile, "Respecting historical spatial integrity: Building a historical 3d florence and avoiding the video game," DISEGNARECON, 2018.

[21] M. Schmitt, "Metaverse: Implications for business, politics, and society," SSRN Electron. J., 2023.

[22] "NVIDIA Omniverse Machinima App" https://www. nvidia.com/en-us/omniverse/apps/machinima/. [Accessed 30-11-2023].

[23] Y. C. Hu, M. Patel, D. Sabella, N. Sprecher, and V. Young, "Mobile edge computing—a key technology towards 5g," ETSI white paper, 2015.

[24] M. Batty, "Digital twins," 2018.

[25] J. Matias, J. Garay, N. Toledo, J. Unzilla, and E. Jacob, "Toward an sdn-enabled nfv architecture," IEEE Communications Magazine, 2015.

[26] S. Suyama, T. Okuyama, Y. Kishiyama, S. Nagata, and T. Asai, "A study on extreme wideband 6G radio access technologies for achieving 100 Gbps data rate in higher frequency bands," IEICE Transactions on Communications, 2021.

[27] R. Di Pietro and S. Cresci, "Metaverse: security and privacy issues," in 2021 Third IEEE International Conference on Trust, Privacy and Security in Intelligent Systems and Applications (TPS-ISA), IEEE,

2021.
[28] B. Hammi, S. Zeadally, and A. J. Perez, “Non-fungible tokens: a review,” IEEE Internet of Things Magazine, 2023.
[29] “Blockchain & Cryptocurrency Laws and Regulations — Italy —GLI — globallegalinsights.com.” https://www.globallegalinsights.com/ practice-areas/blockchain-laws-and-regulations/italy. [Accessed 16-11-2023]

Authors Bio-data

Kien Nguyen received the B.Sc. degree in Telecommunications and Electronics Engineering and the M.Sc. degree in Telecommunications Engineering from Hanoi University of Science and Technology in 2019 and 2021, respectively. He is a currently a Ph.D. student at the Computer Networks department, Institute of Computer Science, University of Würzburg, Würzburg, Germany under the support of German Academic Exchange Service (DAAD). His research interests include edge-cloud computing, serverless computing and future Internet technologies such as SDN and NFV.

Md Sabbir Ahmed received his B.Sc. degree in Computer Science and Engineering from Brac University, Bangladesh. Later, he moved to Italy with a scholarship for his M.Sc. degree in Computer Science and Networking, jointly offered by the Sant'Anna School of Advanced Studies - Pisa and the University of Pisa. After completing his M.Sc. degree, he received research grants and spent several years working in the field of high-performance computing and artificial intelligence at the Department of Information Engineering (DII) of the University of Pisa and the National Institute of Nuclear Physics (INFN), Italy. Currently, he is pursuing his Ph.D. in information engineering at the University of Pisa. His research interests include tactile internet, edge computing, medical IoT, AI in healthcare, and telerobotic surgery.

Claudia Carballo Gonzalez´ (Student Member, IEEE) received the B.Sc. degree in Telecommunications and Electronics Engineering and the M.Sc. degree in Telecommunications and Telematics from the Havana University of Technologies (CUJAE) in 2015 and 2020, respectively. She is a Ph.D. student at the Department of Electrical and Electronic Engineering (DIEE/UdR CNIT), University of Cagliari, 09123, Sardinia, Italy. Her research interests include 5G and beyond networks, O-RAN, slicing, multicast/broadcast, and artificial intelligence.

Chiara Cattani received the Bachelor's Degree in Design/Industrial Design in 2010 and the First Level Master Degree in Communication Design for Sustainability in 2013 from the University of Sassari. She is a student at the Department of Social and Political Sciences, attending Social Communication and Innovation Master Degree (University of Cagliari). She is focused in 3D graphic communication for marketing purpose.

Kaleem Arshid, currently a research assistant at the University of Genova, Italy. He did his MPhil in Computer Science from Preston University in Kohat, Pakistan, in 2017, and his Bachelor of Science in Information Technology from the University of Azad Jammu and Kashmir in Muzaffarabad, Pakistan, in 2013. From 2017 to 2019, he served as a Lecturer in the Department of Computer Science at the University of Sargodha, Pakistan. His research interests span across Cognitive Radio Networks, 5G, 6G Machine Learning, Artificial Intelligence, and UAV Assisted Wireless Networks.

Robust NeRF-based Digital Twins

E. Fontes Pupo[1], Y. Gholizade Atani[2], N. Malekzadeh Goradel[3], L. Savant Aira[4], and G. Zingarini[5]
[1]Fellow, IEEE and Department of Electrical and Electronic Engineering, University of Cagliari, Cagliari, Italy, e.fontespupo@studenti.unica.it
[2] Department of Mathematics and Computer Science, University of Catania, Catania, Italy, yaser.gholizade@phd.unict.it
[3]Department of Information engineering and mathematics, University of Siena, Siena, Italy, n.malekzadehgorad@student.unisi.it
[4]Department of Electronics and Telecommunications, Polytechnic University of Turin, Turin, Italy, luca.savant@polito.it
[5]University of Naples Federico II, giada.zingarini@unina.it

Abstract

The envisaged 5G and beyond networks offer unprecedented breakthroughs in media service delivery, new capabilities, and disruptive use cases. The next realm in the future of broadband communications is immersive communications, blurring the boundaries of the physical and the virtual worlds. This work presents an image-based 3D reconstruction using Neural Radiance Fields [1] that creates digital twins based on web images that may include fake or perturbed ones. Perturbed images may disrupt the 3D reconstruction process; hence, it is necessary to filter them out. We propose an iterative reconstruction method where incremental 3D models are updated sequentially adding images. The final 3D model is deployed to the user through a dynamic multicast adaptive video streaming in an edge cache-assisted network. Our solution includes an ML-based multi-beam multicasting radio resource management (RRM) subject to the viewer's field-of-view (FoV).

Keywords: 3D scenes, metaverse, adversarial attack, NeRF, Digital Twins, edge caching, multicasting.

1. Introduction

The metaverse offers transformative advantages for diverse experiences, such as virtual trips and events. It erases geographic barriers, rendering cultural destinations and attractions universally accessible. Users can engage intimately with various elements through augmented and virtual reality technologies, leading to immersive encounters. Robust social interactions within the metaverse cultivate global communities, facilitating enriched cross-cultural communication. Museums and other institutions can leverage virtual offerings to expand their revenue streams. This inclusive and captivating metaverse fosters cultural appreciation, leveling the playing field for all. Ultimately, the metaverse introduces a new era of exploration and enjoyment, marked by heightened accessibility, community-building, immersive engagements, and sustainable solutions across various domains.

3D reconstruction is a cornerstone technology within the Metaverse, empowering the creation of lifelike virtual environments, interactive objects, and immersive experiences that closely emulate reality. Through methods like photogrammetry or LiDAR scanning, it captures and replicates physical spaces, offering invaluable applications in gaming, virtual tourism, and realistic scenario simulations. Nonetheless, attaining precise 3D reconstructions from 2D images presents challenges; variations in lighting, occlusions, and complex surfaces can impact the accuracy of the models. Furthermore, ensuring accurate depth representation remains a hurdle, especially in expansive environments. Techniques such as stereo vision or photogrammetry encounter difficulties in capturing intricate depth details in certain scenarios. Robustness in 3D reconstruction signifies the capability of reconstruction algorithms to deliver dependable and accurate outcomes amidst diverse conditions and complexities. This resilience is pivotal to maintaining the fidelity and precision of reconstructed 3D models overcoming issues like data variations, environmental challenges, or inherent noise.

Within image-based 3D reconstruction methods, photogrammetry is a widely acknowledged technique capable of generating dense, geometrically accurate 3D point clouds of real-world scenes from a collection of images taken from diverse viewpoints [2]. However, photogrammetry has inherent limitations, especially when dealing with non-cooperative surfaces. Its sensitivity to object textural properties can hinder precise 3D measurements, and it may struggle to create highly detailed reconstructions. Recently, a groundbreaking approach for 3D reconstruction from image datasets, known as Neural Radiance Fields (NeRFs), has garnered significant attention in the research community [1]. This innovative method excels in producing new perspectives of intricate scenes by optimizing a continuous scene function using a set of oriented images. NeRF operates by training a fully connected network, referred to as a neural radiance field, to replicate the input scene views through a rendering loss.

While the NeRF algorithm offers substantial capabilities, acquiring a robust 2D image dataset remains a significant challenge. Without such a high-quality dataset, the NeRF algorithm might not achieve optimal results. Furthermore, various entities, such as museums and companies, maintain stringent privacy protocols for their datasets, limiting accessibility. In such instances, leveraging web-based images becomes crucial to assemble a comprehensive dataset. However, assembling such a dataset requires employing some AI-based methods like Content-based image retrieval (CBIR) [3] to distinguish between valuable and irrelevant images [4, 5]. In addition, certain fabricated images may infiltrate the relevant image set, which can be identified and removed using advanced deep learning techniques such as deepfake detection methods [6]. Once a curated dataset is established, the NeRF algorithm can then be effectively utilized for the 3D reconstruction process.

For bandwidth-demanding applications, especially for a dense user deployment, multicast traffic delivery can significantly benefit communication and computing components. This capability provides cost-effective and resource-efficient delivery mechanisms to multiple end-users requesting the same content [7]. Tailored point-to-multipoint communication strategies can provide considerable capacity gain into the beyond 5G ecosystem. In such a context, ubiquitous edge-caching and multicast support can enable a large scale of low-latency video services, as above presented. In [8], the authors analyze how applications based on omnidirectional video with ultra-large bandwidth, such as VR and holographic, heavily rely on adaptative video streaming to deliver multiple streams of high-definition content subject to the viewer's field-of-view (FoV).

Bearing the previous analysis as a benchmark, we cover essential aspects regarding 3D scenes, adversarial attacks, NeRF, and service deployment in this research. We present an image-based 3D reconstruction using Neural Radiance Fields where incremental 3D models are updated by adding images in a sequential way. The final 3D model is delivered to the end-users, combining adaptive video streaming in an edge cache-assisted network and dynamic multi-beam multicasting.

The remainder of the paper is structured as follows. Section II presents the system model and scenario description. Section III details the proposed methodology, analysis, and remarks. Finally, conclusions are drawn in.

2. Scenario Description and Proposed Method

We want to study an adversarial scenario regarding 3D scenes used in the metaverse. We consider that an Actor A possesses a valuable Target Object (TO), be it a statue, prototype, engine, or any tangible asset. Concurrently, actor B wants to develop a virtual 3D replica of TO, exploiting publicly available images of the TO, called Dataset. Actor A has the ability to engage in an adversarial attack. This involves the deliberate injection of counterfeit versions and disruptive artifacts into the Dataset. In actor *B*, we identify that needs to effectively filter out these tampered images, creating a digital twin of the TO, easily and efficiently deployable in a 3D synthetic environment.

Figure 6Picturing a possible scenario

To counter adversarial images during the 3D scene reconstruction process, we have developed a method that aims to maintain the integrity of the reconstructed model. Our proposed approach involves several key steps, as presented in Figure 2. By implementing these steps, our proposed approach defends against adversarial attacks and enables the creation of a robust and coherent 3D scene model suitable for multiresolution paradigm broadcasting in augmented reality environments. In the following subsections, we delve into the different elements of the proposed architecture.

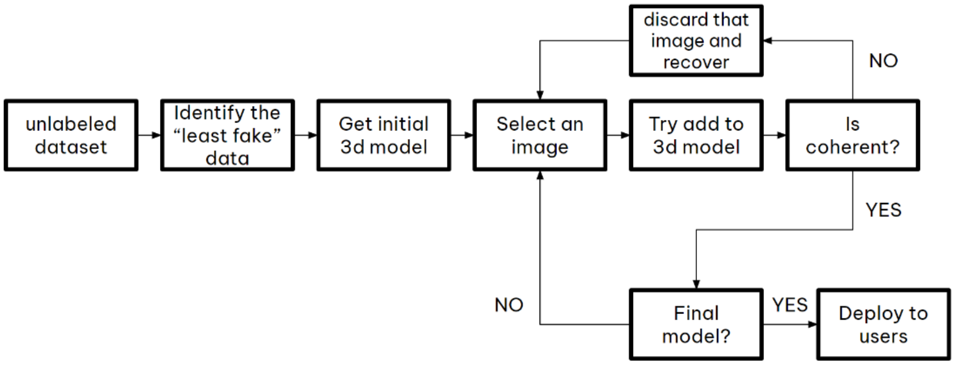


Figure 7 General overview of the proposal

2.1 Initial 3D Scene Reconstruction

Our process starts by collecting as many images as possible of the TO, in order to obtain representations of it from different points of view. Thus, this first step provides the research of the data on the web and it could lead to the possibility of finding non-compliant images. It often happens that data coming from different web sources may have been manipulated using either traditional editing tools or deep learning-based ones, or even they may have been fully generated using the more recent and sophisticated AI methods. Therefore, the first step we need to pursue is a dataset-cleaning operation that involves a filtering process by detecting all the images that seem to be fake and consequently useless.

Given the potential damage that maliciously poisoned images could produce, our system incorporates advanced image analysis techniques to detect these images and exclude them from the following steps. The Detecting operation is conducted with *D*, which represents our detecting methods. We suppose that the adversarial attack could involve either the creation of fully generated fake versions of the data or manipulated images with the injection of fake parts inside them. The *D* model is implemented as a neural network, typically these could be Convolutional Neural Network (CNN) [9, 10], or the more recent Transformer architecture [11, 12]. The techniques are usually looking for the artifacts left in the images by the manipulations. These are often invisible to the human eye and have to be searched with the help of data-driven deep-learning approaches. Indeed, during the image acquisition process each type of camera generates a specific noise that is consequently visible in real and unaltered images. However, when the images are tampered with manipulations these original camera's traces are corrupted and the alteration process becomes visible [13]. Once the dataset has been cleaned up, we can move on to the next step for the creation of the 3D scene.

Starting Robust NeRF-Based Reconstruction: Leveraging robust Neural Radiance Fields techniques, as outlined [14, 15] we perform the initial 3D scene reconstruction. This process is critical for establishing a baseline representation of the scene to which further images can be compared and aligned to.

2.2 3D Model Iterative Refinement

Incremental Image Integration: Next, we select a new image from the unprocessed ones and attempt to incorporate it into the existing 3D model. This incremental image integration allows us to continuously refine the model, evaluating each image by its own.

Coherence Evaluation and thresholding: To assess the coherence of the updated 3D model, we perform novel view synthesis to evaluate coherence and reuse the detector *D*. This step helps us determine how well the added image aligns with the existing model and whether it introduces any inconsistencies. We propose a straightforward yet effective thresholding approach to maintain low runtime complexity. The output of the detector *D* is used to decide whether to keep or discard the integrated image based on the score of the generated novel view. Images that meet a predefined coherence level are retained, while those that do not pass the threshold are discarded.

Iterative Refinement: Finally, we enter an iterative loop, returning to step 4, where we select a new image and attempt to add it to the 3D model. This iterative process safeguards the model against potential adversarial attacks.

2.3 Final Users' Deployment

The above-described reconstruction process stops when the custom objective criterion is met or all the images have been scanned. Once the final 3D model is built, it must be delivered to the end users. The multicast capability can significantly improve communication and computing components in the context of the addressed use case. Moreover, In [8], Zhong et al. present a decentralized optimization for multicast adaptative video streaming in edge-caching assisted networks. In [16], the authors present a multicast-aware optimization for resource allocation combined with edge computing and caching. As defined in [17], VR broadcasting is the most direct method that aids wireless users in getting access to the metaverse. The VR broadcast is transmitted in a heterogeneous manner in the metaverse due to the heterogeneous nature of wireless and edge devices.

From the above analysis, our service delivery strategy is based on a dynamic multi-beams multicast adaptive video streaming in an edge cache-assisted network. Figure 3, represents a simplified block diagram of the proposed delivery strategy. We avoid delivering the whole image and any unnecessary bandwidth consumption caused by the potential delivery of the full image. Integrating edge caching into the video processing chain reduces the latency and simplifies the multicast design [8, 18]. The considered adaptive video streaming management enables personalized video experience with dynamic adjustments of the video bitrate regarding the variations in the network, the users' reception conditions, and the FoV.

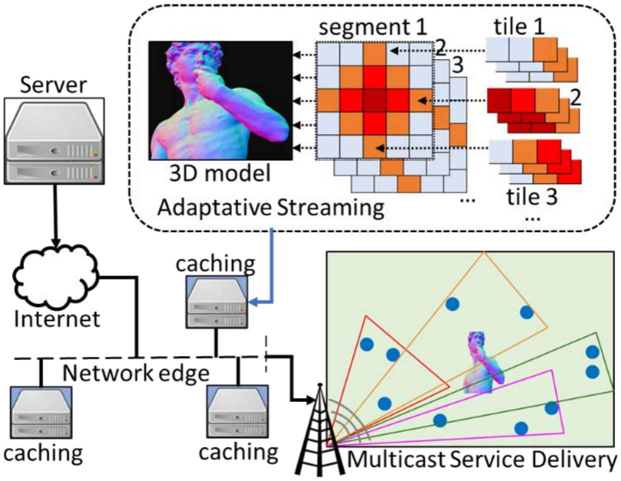


Figure 8 General illustration of a cache-assisted adaptative 3D streaming with dynamic multi-beams multicasting system

An important advantage of the presented cache-assisted network is that the core of the 3D model post-processing is executed at the nearby edge caching node instead of the far multimedia server. This characteristic is essential for our particular application, where the 3D image is dynamically tiled with multiple representations and requirements. Each tile is associated with a specific bitrate regarding the reception conditions of users with a common FoV.

Our solution includes an ML-based multicasting radio resource management (RRM) subject to the requirements of each tile and the characteristics of the users interested in the same tile. The implemented multicasting solution takes advantage of the users' spatial diversity and common FoV to execute adaptative multicast multi-beam beamforming. We dynamically adjust the number of multicast beams, the beam parameters, and the allocated resources, accomplishing the tile requirements and maximizing the users' quality of service and quality of experience.

3. Conclusions

In the context of augmented reality, the faithful representation of real objects is paramount. Our approach, designed to enhance security and fidelity in 3D scene reconstructions for multiresolution broadcasting, offers a robust solution. Through meticulous image curation and detection of maliciously altered content, we have established a defense against adversarial threats. Utilizing advanced Neural Radiance Fields (NeRF) techniques, we have strengthened our 3D model creation processes. Iterative refinement and efficient thresholding further ensure the model's integrity. As augmented reality environments advance, our method stands as a resilient guardian, securing the faithful digital twin representation of real objects in a virtual environment. The final 3D model is delivered to the user through a dynamic multi-beam multicast adaptive video streaming in an edge cache-assisted network.

References

[1] B. Mildenhall, P. P. Srinivasan, M. Tancik, J. T. Barron, R. Ramamoorthi, and R. Ng, “Nerf: Representing scenes as neural radiance fields for view synthesis,” Communications of the ACM, vol. 65, no. 1, pp. 99–106, 2021.
[2] F. Remondino and S. El-Hakim, “Image-based 3d modelling: a review,” The photogrammetric record, vol. 21, no. 115, pp. 269–291, 2006.
[3] I. M. Hameed, S. H. Abdulhussain, and B. M. Mahmmod, “Content-based image retrieval: A review of recent trends,” Cogent Engineering, vol. 8, no. 1, p. 1927469, 2021.
[4] E. Uzun, E.˙Ozhan, H. V. Agun, T. Yerlikaya, and H. N. Bulus ̧, “Automatically discovering relevant images from web pages,” IEEE Access, vol. 8, pp. 208910–208921, 2020.
[5] Y. Yao, J. Zhang, F. Shen, X. Hua, J. Xu, and Z. Tang, “Exploiting web images for dataset construction: A domain robust approach,” IEEE Transactions on Multimedia, vol. 19, no. 8, pp. 1771–1784, 2017.
[6] J. Yang, S. Xiao, A. Li, G. Lan, and H. Wang, “Detecting fake images by identifying potential texture difference,” Future Generation Computer Systems, vol. 125, pp. 127–135, 2021.
[7] V. K. Shrivastava, S. Baek, and Y. Baek, “5g evolution for multicast and broadcast services in 3gpp release 17,” IEEE Communications Standards Magazine, vol. 6, no. 3, pp. 70–76, 2022.
[8] L. Zhong, M. Wang, C. Xu, S. Yang, and G.-M. Muntean, “Decentralized optimization for multicast adaptive video streaming in edge cache-assisted networks,” IEEE Transactions on Broadcasting, 2023.
[9] D. Cozzolino and L. Verdoliva, “Noiseprint: A cnn-based camera model fingerprint,” IEEE Transactions on Information Forensics and Security, vol. 15, pp. 144–159, 2019.
[10] Q. Bammey, R. G. v. Gioi, and J.-M. Morel, “An adaptive neural network for unsupervised mosaic consistency analysis in image forensics,” in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 14194–14204, 2020.
[11] J. Hao, Z. Zhang, S. Yang, D. Xie, and S. Pu, “Transforensics: image forgery localization with dense self-attention,” in Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 15055–15064, 2021.
[12] F. Guillaro, D. Cozzolino, A. Sud, N. Dufour, and L. Verdoliva, “Trufor: Leveraging all-round clues for trustworthy image forgery detection and localization,” in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 20606–20615, 2023.
[13] L. Verdoliva, “Media forensics and deepfakes: an overview,” IEEE Journal of Selected Topics in Signal Processing, vol. 14, no. 5, pp. 910–932, 2020.
[14] C. Wang, A. Wang, J. Li, A. Yuille, and C. Xie, “Benchmarking robustness in neural radiance fields,” arXiv preprint arXiv:2301.04075, 2023.
[15] Z. Li, T. M¨uller, A. Evans, R. H. Taylor, M. Unberath, M.-Y. Liu, and C.-H. Lin, “Neuralangelo: High-fidelity neural surface reconstruction,” in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 8456–8465, 2023.
[16] H. Hao, C. Xu, S. Yang, L. Zhong, and G.-M. Muntean, “Multicast-aware optimization for resource allocation with edge computing and caching,” Journal of Network and Computer Applications, vol. 193, p. 103195, 2021.
[17] M. Ali, F. Naeem, G. Kaddoum, and E. Hossain, “Metaverse Communications, Networking, Security, and Applications: Research Issues, State-of-the-Art, and Future Directions,” arXiv preprint arXiv:2212.13993, 2022.
[18] C. Reiser, S. Peng, Y. Liao, and A. Geiger, “Kilonerf: Speeding up neural radiance fields with thousands of tiny mlps,” in Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 14335–14345, 2021

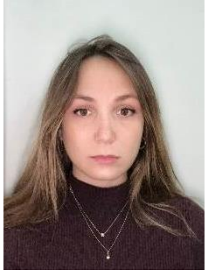
Authors Bio-data

Ernesto Fontes Pupo (e.fontespupo@studenti.unica.it) is a Ph.D. student at the Department of Electrical and Electronic Engineering (DIEE/UdR CNIT), University of Cagliari, 09123, Sardinia, Italy. His research interests include 5G/6G networks, PTP/PTM, and artificial intelligence. He is a Student Member of IEEE.

Yaser Gholizade Atani (yaser.gholizade@phd.unict.it) is a Ph.D. student at the Department of Computer Science and Mathematics, University of Catania, Catania, Italy. His research interests include Explainable Artificial Intelligence and Data Science. He is a Student Member of IEEE.

Nasrin Malekzadeh Goradel (n.malekzadehgorad@student.unisi.it) is a Ph.D. student at the Department of Information engineering and mathematics, University of Siena, Siena, Italy. Her research interests include certified deep learning, riemannian manifolds and trustworthy AI. She is a member of the Visual Information Privacy and Protection Group (VIPP) and a Student Member of IEEE.

Luca Savant Aira (luca.savant@polito.it) is a Ph.D. student at the Department of Electronics and Telecommunications, Polytechnic University of Turin, Turin, Italy. His research interests include Deep Learning and 3D Computer Vision. He is a Student Member of IEEE



Giada Zingarini (giada.zingarini@unina.it) is a Ph.D. student at the Department of Information Technology and Electrical Engineering, University of Naples Federico II, Italy. Her research interests include Multimedia Forensics and Medical Imaging. She is a Student Member of IEEE.

MMTC OFFICERS (Term 2024 — 2026)

CHAIR

Jun Wu
Fudan University
China

STEERING COMMITTEE CHAIR

Joel J. P. C. Rodrigues
Federal University of Piauí (UFPI)
Brazil

VICE CHAIRS

Shaoen Wu (North America)
Illinois State University
USA

Liang Zhou (Asia)
Nanjing University of Post and Telecommunications
China

Abderrahim Benslimane (Europe)
University of Avignon
France

Qing Yang (Letters & Member Communications)
University of North Texas
USA

SECRETARY

Han Hu
Beijing Institute of Technology
China

STANDARDS LIAISON

Weiyi Zhang
AT&T Research
USA

MMTC Communication-Frontier BOARD MEMBERS (Term 2024—2026)

Name	Role	Affiliation	Country
Danda Rawat	Director	Howard University	USA
Sudip Misra	Director	IIT Kharagpur	India
Rui Wang	Co-Director	Tongji University	China
Joy Lal Sarkar	Co-Director	Techno College of Engineering Agartala	India
Satendra Kumar	Co-Director	IIT Patna	India
Luca Foschini	Editor	University of Bologna	Italy
Mohamed Faten Zhani	Editor	l'École de Technologie Supérieure (ÉTS)	Canada
Armir Bujari	Editor	University of Padua	Italy
Kuan Zhang	Editor	University of Nebraska-Lincoln	USA
Dapeng Wu	Editor	Chongqing University of Posts and Telecommunications	China
Shuaishuai Guo	Editor	King Abdullah University of Science and Technology	Saudi Arabia
Guanyu Gao	Member	Beijing Institute of Technology	China
Alessandro Floris	Member	University of Cagliari	Italy
Simone Porcu	Member	University of Cagliari	Italy
Shiqi Wang	Member	City University of Hong Kong	Hong Kong, China
Panagiotis Sarigiannidis	Member	University of Western Macedonia	Greece