

Received September 17, 2021, accepted November 8, 2021, date of publication November 16, 2021, date of current version November 29, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3128786

On the Exploitation of 5G Multi-Access Edge Computing for Spatial Audio in Cultural Heritage Applications

CLAUDIA RINALDI¹, **FABIO FRANCHI**¹, (Member, IEEE),
ANDREA MAROTTA¹, (Member, IEEE), **FABIO GRAZIOSI**¹, (Member, IEEE),
AND CARLO CENTOFANTI¹, (Graduate Student Member, IEEE)

Department of Information Engineering, Computer Science and Mathematics, Università degli Studi dell'Aquila, 67100 L'Aquila, Italy

Corresponding author: Claudia Rinaldi (claudia.rinaldi@univaq.it)

This work was supported in part by the Italian Government (December 21, 2012) through the Project INnovating City Planning through Information and Communication Technologies (INCIPICT) under Comitato interministeriale per la programmazione economica (CIPE) Resolution 135, and in part by the European Commission (EC) through the H2020-MSCA-RISE-2019 OPTIMIST Project under Project 872866.

ABSTRACT This work presents a service for the improvement of cultural heritage experiences, which exploits the advantages coming from the 5G paradigm. Indeed, in a scenario where many users need to be served by a real-time solution which is in turn required to work on different devices, the potentialities of 5G technology show their suitability. In particular, moving the computation to the edge of the network ensures the availability of resources needed for binaural spatial audio rendering in an independent fashion with reference to the client device and at the same time it guarantees real-time availability of this data since the core network, with its impairments, is not involved. This work demonstrates how 5G could be a critical enabler for delivering low latency services at guaranteed levels, data-centric services, differentiated customer experiences, improved security and reduced costs to the users.

INDEX TERMS Multi-access edge computing, cloud computing, 5G, augmented reality, virtual reality, cultural heritage, spatial audio, binaural rendering.

I. INTRODUCTION

The 5th generation of mobile network promises great changes that may lead to a different way of understanding mobile communications, moving from the need of connecting people to the need of connecting their worlds. The business ecosystem around this paradigm is involving many different vertical industries on various fields.

Three categories of services can be identified for 5G services, [1]:

- **enhanced Mobile Broadband (eMBB)** that aims to support high bandwidth demanding services;
- **ultra Reliable and Low Latency Communications (uRLLC)** has been introduced in order to cope with safety and mission critical services by guaranteeing high reliability and low latency communications;

- **massive Machine Type Communications (mMTC)** as enabler for Internet of Things (IoT) services which require high density of connected devices.

Vertical services exploited by 5G trials should include media and entertainment, public safety, e-health, automotive, transport and logistic, Cultural Heritage (CH), [2]. The latter, in particular, has been significantly affected by the progress of digital information especially conceiving its dissemination [3], offering new technological possibilities for developing, e.g. the market of tourist services [4], and CH organizations have to address new user needs by creating innovative applications [5], such as Augmented and Virtual Reality (AR/VR) based. AR technology gives a different perception of reality, as it enriches reality with a computer-generated layer containing visual, audio, and tactile information, while using a “virtual” representation of a classic museum allows access to aspects of the artifact that may otherwise be hidden [6].

The associate editor coordinating the review of this manuscript and approving it for publication was Chi-Yuan Chen¹.

During the past years, the main aim of AR/VR applications for CH changed from a mere virtual recreation of object to display, to create an entire virtual environment able to disseminate and teach culture. The idea is the opposite of a “dead museum”: users must not be exposed to an accumulation of 3D heritage objects, but feel and understand another culture through those items. An important aspect is the relation between AR/VR and education. This new way to present Culture enhances the learning process, encouraging students and researchers through stimulating methods of presentation of archival materials and historical events. Users can therefore travel through space and time without moving from their home [7]. Numerous AR/VR applications exist for CH or tourists enjoyment of places with a rich past, allowing a realistic navigation of environments that no longer exist or that may be inaccessible, [8]–[11], [12].

The most of the previously cited experiences do not take into account the advantages that may arise by a proper exploitation of sounds together with visual effects, except for audio content presentation purposes. On the contrary, an acoustic guide, properly placed in the virtual space, may drive the user toward a certain direction or the acoustic landscape of a specific historical period could be reproduced to improve the virtual experience. To the authors’ knowledge only a few experimentation have been carried out in this context.

The presented service implements spatial audio rendering through 5G in order to exploit possible advantages such as the reduction of the computational requirements for the local devices and the increase of the amount of users that may require the same service on smaller and lighter AR/VR devices.

A. MOTIVATION AND CONTRIBUTIONS

5G technology foresees impressive numbers as guaranteed latency of some milliseconds and throughput higher than 1Gb/s, [13], which may perfectly fit the requirements coming from the exploitation of AR/VR solutions for CH.

To be more specific, the need of a binaural sound rendering solution to act in real time as the user moves the head or changes the position in time as well as the addition of video streaming to the audio, would require ultra-low latency and high throughput for a proper experience.

Moreover, being aware that the computational requirements of real time adaptation of both audio and video streaming are high, the scenario under analysis may also benefit from a new born paradigm, in the context of 5G networks, known as Multi Access Edge Computing (MEC). A MEC approach in such context has also the advantage of allowing to offload the multimedia elaboration to the edge of the network, making the user devices less complex and power hungry.

MEC enables services and applications to be hosted ‘on top’ of the mobile network, i.e. above the network layer. These services and applications can benefit from being in close proximity to the users and from receiving local connectivity enabling new business opportunities. In malls, [14],

university areas, [15], or museums, [16] that are filled with high-value users, 5G MEC can provide value added services, such as local cache service, location service, and targeted advertising. At business campuses, factories, and seaports, 5G MEC can provide enterprise-level services, such as virtual private networks, service hosting, and dedicated applications. Due to the possibility to drastically reduce experienced latency and offloading computation to the edge, MEC has emerged as an enabler for a wide range novel services including Industrial IoT [17], low-latency mission critical applications [18], vehicular communications [19], and multimedia services [20].

Referring to the proposed scenario, by admitting edge computing, the limit of 40ms between head movement and spatialized sound, foreseen in [21] as the limit above which a processing/transmission delay is perceivable, can be also exploited for improving signal processing algorithms given that the communication is performed on a network responding to the uRLLC paradigm.

Advantages coming from exploitation of 5G for AR/VR for CH thus need to be deepened because 5G CH AR/VR could represent a possible killer application for this communication technology.

This paper focuses on the exploitation of audio spatialization service, supported by the 5G network architecture, for improving AR/VR experiences for CH. Here follows a list of the main topics involved in the scenario under analysis that will be discussed in the remaining of the paper:

- exploitation of *Resonance Audio* for a CH AR/VR application scenario;
- definition of a 5G based MEC architecture for AR/VR support;
- analysis of advantages achievable through the proposed solution.

Moreover, some experiences related to exploitation of 5G for AR/VR applications were already conducted by the authors within the 5G trial carried out in L’Aquila, [22].

It is worth noting that, in this phase of the research, we assume a perfect and reliable behavior of the eventual head orientation localization system, i.e. we do not assume head tracking and thus the issues that may arise such as adaptation of the binaural sound to these movements. We are anyway confident of the proper behavior of the orientation/localization system on the specific device we are using for AR. Moreover, we do not face the problem of active noise cancellation for excluding sounds from the real world since our aim is to exploit lo-fi devices.

The remaining of the paper is organized as follows: in Section II the SoA on the various topics involved is presented; Section III sketches the scenario under analysis; details on the advantages of MEC for the previously described scenario are given in Section IV; Section V discusses about the referred spatialization tools while the application is described in Section VI. Performance analysis and results are reported in Section VII while conclusions are drawn in Section VIII.

II. PREVIOUS WORKS

Real time binaural rendering solutions for AR/VR enjoyment of CH by exploiting 5G solutions is a topic involving various research aspects. In this section a brief state of the art of the most relevant topics involved in the project is presented.

A. SPATIAL AUDIO FOR AR/VR CH APPLICATIONS

As previously stated, the importance of sounds in the context of AR/VR for CH is not yet understood and the research literature is quite poor.

One of the few exceptions is given by [23], in which authors present a signal processing method for fast real-time binaural synthesis, whose main target application is the fruition of cultural heritage and in [24] where a smart headphones set is presented in order to remotely take the listener's head orientation and properly generate an audio output to attract tourists' attention toward specific points of interest in the 3D space. An interesting analysis of hardware and software requirements for this purpose is presented in [25] without references to real applications.

In [24], authors propose an interaction system for attracting the visitor toward specific cultural attractions through 3D audio. The work is focused toward the design and development of a system for gathering head orientation in real time and some interesting tests on the exploitation of spatial sounds were also presented

"*The Ghost Orchestra*" is an interesting project involving exploitation of binaural spatialization and visors for VR (i.e. Oculus) for cultural heritage as described in [26].

B. SPATIAL AUDIO FOR OTHER SIGNIFICANT SCENARIOS

The interest on audio spatialization by the research community has its origins on a paper written in the 60s by Schroeder, [27] introducing the idea of artificial reverberation based on digital signal processing, [28]. Indeed reproducing the behavior of acoustic waves spreading indoor requires ideally to spatialize each reverberation, i.e. to reproduce at the ears the sensation of a 3D space. Since then, the applied research and the market have gone toward different applications spanning from games, [29], to electroacoustic music, [30] and soundscape design, [31] up to sonification, [32]. All these scenarios have in common the initial exploitation of the stereophonic approach, that provides a multi-channel reproduction system going from the traditional two channel stereo to the modern configurations with five, seven or more loudspeakers. This can be considered a channel based approach that freezes the position of the sound to the signals relations between loudspeakers, [33] and it is the one that is currently still used for cinema, home theater and pure audio content. The need of properly preserve spatial cues of an auditory scene has brought to the separation in coding of source signal and source location and to all the variety of spatialization techniques partially listed in the paragraph below, that have in turn significantly increased the potentials of video games, audio games, music expressions etc.

Some remarkable results of applying spatial audio in these contexts are Mojang AB Minecraft for games or Karlheinz Stockhausen Cosmic Pulses for electroacoustic music.

C. BINAURAL SOUND RENDERING SOLUTIONS

The binaural sound reproduction through headphones requires full control of sound synthesis and binaural cues to be guaranteed, at the expense of the need by the users of wearing devices that may be considered intrusive, especially when noise cancellation and ear occlusion are required by the application. In order to achieve auralisation through headphones, HRTF (head related transfer function) filters are commonly used for left and right-ear because with headphones, the effect of the head and the pinna (with earplugs) is bypassed. They are then convolved with an anechoic sound signal for audio rendering. The most of the literature is focused toward the exploitation of non personalized HRTF that are recorded using a dummy head (e.g. KEMAR manikin, [34]) on a discrete spatial grid in both azimuth and elevation, [35]. Given the apparent impossibility of overcoming problems of unnatural coloration of the frequency spectrum and localization degradation, recent literature is instead focusing toward the exploitation of personalized HRTF, which still requires a long evaluation procedure (see e.g. [36]). It is worth noting that the availability of HRTFs coming from the last decade of research has brought to a standardization process known as the Spatially Oriented Format for Acoustics (SOFA), [37], a personalized version is claimed to be available here [38]. In both cases, given the discrete points along which HRTFs are measured/computed interpolation techniques have to be employed, [39].

Most of the available tools for binaural audio reproduction are based on procedures to move from surrounding systems solutions to binaural sounds. Each of them need to use static, dynamic and environmental cues, [40]. Static cues are given by HRTF or HRIR involving all the issues related to physical characteristics of each individual; dynamic cues are related to the motion of the listener and environmental cues are given by the room transfer function (RTF) or room impulse response (RIR).

For instance with the purpose of down mixing 5.1 to binaural, a solution exploiting virtual loudspeakers and HRIR, taking also into account the room response and head tracking data is proposed in [41].

The basic procedure for *Ambisonic* binaural rendering defines the virtual loudspeakers layout for which the corresponding output is computed as a linear combination of the B-format channels and finally HRTFs are introduced for each virtual loudspeaker and the obtained left channels are summed together, as it happens for right channels, and they respectively feed the left and right channels of the headphones, [42]. An optimized solution has been recently proposed in [43], where binaural decoding of Ambisonic soundfields is achieved basing on pre-computed, spherical harmonic-encoded binaural filters. Authors in [44] presented a *vector base amplitude panning (VBAP)* implementation

for 3D head-tracked binaural rendering, where the binaural implementation of VBAP is achieved in the same way as virtual Ambisonics.

With the purpose of reducing computational requirements, various solutions have also been investigated simplifying the HRTFs and RIRs, see for instance [45], [46] and references therein, most of them at the expenses of perceived quality.

Finally, referring to [43] and [44], in the light of the main topic of this work, that authors propose opposite solutions about the device in charge of making spatial sound computations. The first reference indeed suggests to implement the VBAP rendering on an embedded Linux device that is placed locally. Advantages and drawbacks of the solution are discussed in terms of proper reproduction of the virtual source and response of the system to head-tracking data. It is interesting to observe that authors suggest the exploitation of second and third order Ambisonic to increase accuracy, stating that this would require computational resources not yet available on the CPU of an embedded system. On the other hand, authors in [44] investigate the opportunity of exploiting the distributor-side. This choice was mainly justified by economic reasons and drawbacks considered only in terms of absence of information about the listener environment. Problems arising in both cases could be overcome by exploiting the architecture proposed in this paper.

D. 5G ENABLED AR/VR

According to the recent Moxel State of 5G survey, [47], AR and VR applications top the list of primary use cases for 5G technology in consumer applications.

AR and VR experiences introduce many technical challenges related to the need of combining and synchronizing the real or virtual world with the user’s motions. This requires high computational resources for rendering that can benefit from moving partially or totally the computational tasks to the edge. As a consequence, introducing 5G in this scenario would allow to satisfy users quality of experience (QoE) guaranteeing very low latency thus a realistic experience. The exploitation of the concept of private network is the solution proposed by Ericsson [48], basing on which VR services are delivered either through enterprise dedicated private networks or through a logical network slice created on the top of the existing physical public network.

Despite the large amount of informative articles available on the web about the advantages 5G could bring to AR/VR applications (e.g. [49]–[51]), scientific papers or available products exploiting these two paradigms together are quite limited.

An interesting discussion about advantages of 5G and MEC in the context of Mobile Augmented Reality (MAR) is discussed in [52] where the authors also discuss a possible application on tourism, stating the current status of infancy of these applications so far. In [53] a demonstration of the important role of 5G networking for VR game is shown by moving game servers without service interruption. An interesting project is described in [54] where public trials demonstrating

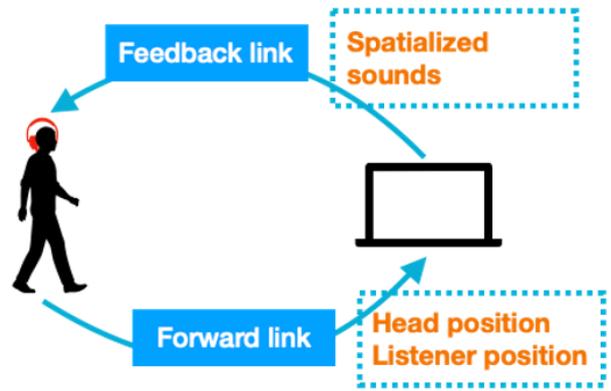


FIGURE 1. Spatial audio scenario.

the advantages of 5G for smart tourism are reported. Authors in [55] and [56] discuss in detail the potentiality of 5G and Beyond 5G (B5G) cellular networks for realizing mobile web augmented reality, presenting encouraging results.

III. SCENARIO

We assume as general scenario a museum where more users exploit AR devices for moving around and getting dedicated information as a function of what they are observing and their preferences that have been eventually previously expressed. The referred scenario for each user is sketched in Figure 1. The behavior of the service consists in a bidirectional communication between a client, that gets environmental information using different sensors (e.g., cameras, gyro, beacons, etc.), and a MEC application able to process the information in order to produce a proper spatialized stream. As a result the client application can reproduce a personalized sound built in real-time basing on users’ position into the museum space.

IV. 5G SUPPORTING SPATIAL AUDIO AT EDGE

Audio applications dealing with binaural sounds require a very fast computation and playback of sounds to fulfill users’ overall expectations. This can be reached both by computing audio at the user’s device and by delegating computation to a remote node. The former option can offer the most reliable solution but it requires energy to perform computation. More energy means more weight to be carried by the end user (e.g., batteries) or reduced battery life. The latter option allows to release power constraints but introduces information transmission delay to move data through the network. Time needed to reproduce a valid sound sample is composed by the time to compute the sound plus two times the network delay to move information forward and backward from the end user device to the computational node:

$$T_{\text{sound}} = T_{\text{processing}} + 2T_{\text{network}}$$

where

$$T_{\text{network}} \approx T_{\text{forward}} \approx T_{\text{backward}}$$

With respect to previous mobile communications systems, 5G offers an unprecedented level of flexibility to fulfill

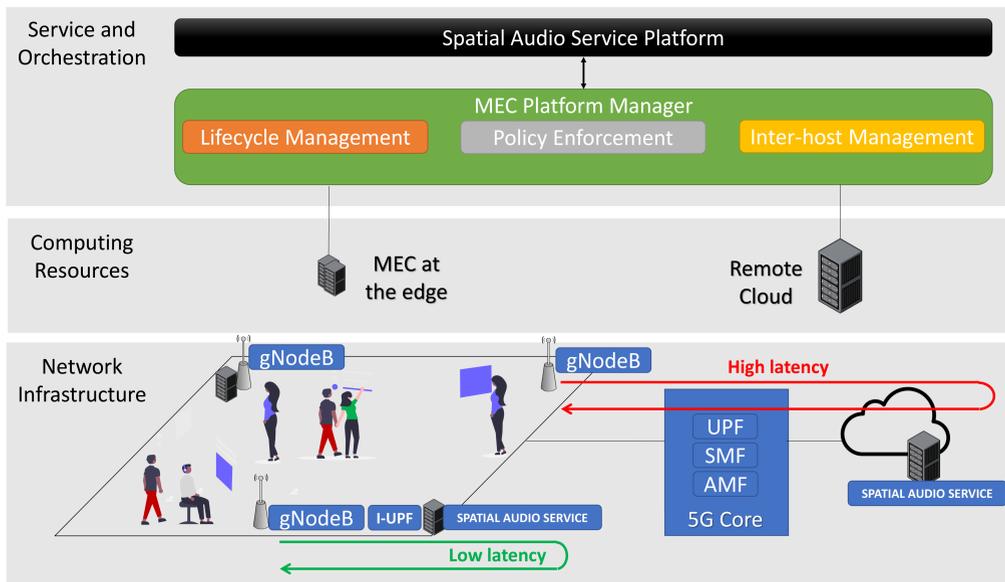


FIGURE 2. Detailed service architecture from network layers point of view.

service-specific requirements in terms of throughput, latency, and reliability. Such flexibility is achieved through novel techniques for the radio transmission and by novel architectural approaches. Two enabling technologies from the architectural viewpoint are: Software Defined Network (SDN) that allow to separate control and user plane information and to flexibly adapt the behavior of the network via software; Network Function Virtualization (NFV) that allows to implement and *orchestrate* traditional network functionalities over virtualized infrastructures.

Due to the virtualized nature of the 5G architecture, the core network is realized as a set of services among which the main ones are: (i) Access and Mobility Management Function (AMF) that is responsible for handling connection and mobility management tasks; (ii) Session Management Function (SMF) that is responsible for managing connections and session contexts; (iii) User Plane Function (UPF) that is the anchor element between the mobile and data networks.

Figure 2 shows the considered 5G architecture for the spatial audio service. Normally, the traffic generated by the users has to traverse the core network in order to leave the mobile network and reach the data network. However this may result in too high delay not compatible with the service under consideration (red line in Figure 2). In order to overcome this issue and fully enable MEC capability an instance of the UPF is deployed at the edge of the network. Since it is possible to separate control and user planes, the control plane is still redirect to the 5G Core to perform signaling and control while the intelligent UPF (I-UPF) at the edge is used as an anchor towards the data-network for the user plane. This way, the user plane traffic can leave immediately the mobile network after reaching the 5G base station (i.e., gNodeB) and flow towards the elaboration server at the edge, thus reducing latency (green line in Figure 2).

The computing infrastructure comprises the computing resources placed at the edge of the network (i.e. in proximity of the gNodeBs) allowing MEC, and computing resources available in a remote cloud. On top of the network and computing infrastructure is a Service and Orchestration layer which is responsible for the management of the deployment and life-cycle of the spatial audio service.

The Service and Orchestration Layer is composed by the MEC Platform Manager and the DASH Service Platform. The MEC Platform Manager is able to provide the upper layer with a uniform northbound interface API decoupling the Service Layer and the Computing Resources Layer. Its functions include:

- **Life-cycle Management:** this process takes care of the entire MEC application life-cycle, treating each MEC application instance as a Virtual Network Function (VNF) instance and performing health checks and auto healing for high availability
- **Policy Enforcement:** it manages networks and applications connectivity, according to the network policies
- **Inter-host Management:** this block is responsible of enabling networking and communications between different MEC hosts, exposing them to the *Service Layer*.

V. REFERRED SPATIALIZATION TOOLS

Bringing rich, dynamic audio environments into AR/VR experiences without affecting performance can be challenging. There are often few CPU resources allocated for audio, especially on mobile, which can limit the number of simultaneous high-fidelity 3D sound sources for complex environments.

Various tools implementing binaural spatialization were investigated and their advantages and disadvantages

discussed before making a choice, i.e. Resonance Audio (Google), 3D Tune-In Toolkit, [57], Audio Spatializer [Oculus (Facebook)], Steam Audio (Valve corporation), SOFALizer [58], Spatial Audio Framework [59].

The choice fell on Resonance Audio, [60]. Resonance Audio is a multi-platform spatial audio SDK, delivering high fidelity at scale. The Resonance Audio SDK uses highly optimized digital signal processing algorithms based on higher order Ambisonics to spatialize hundreds of simultaneous 3D sound sources, without compromising audio quality. The SDKs run on Android, iOS, Windows, MacOS and Linux platforms and provide integration for Unity, Unreal Engine, FMOD, Wwise and DAWs. Native APIs for C/C++, Java and Objective-C are also provided together with the full source code C++ library.

As part of the open source project, a reference implementation of YouTube’s Ambisonic-based spatial audio decoder is provided. Using this implementation, developers can easily render Ambisonic content in their VR media and other applications, while benefiting from Ambisonics open source, royalty-free model. The project also includes encoding, sound field manipulation and decoding techniques, as well as head related transfer functions (HRTFs) used to achieve rich spatial audio that scales across a wide spectrum of device types and platforms. Lastly the entire library of highly optimized DSP classes and functions is available: this includes resamplers, convolvers, filters, delay lines and other DSP capabilities.

VI. SPATIALIZATION SERVER APPLICATION

In order to demonstrate the proposed service a custom application has been developed. The application, written in C++, uses a custom Linux build of the Resonance Audio library and some utility libraries like the Boost C++ Libraries in order to handle the TCP socket communication between client and server in the easiest way, [61]. Figure 3 summarize the simple behavior of the developed service: a client streaming application (e.g., an AR mobile application) is able to connect to the *spatialization server* that is awaiting for requests from connected clients of spatialization processing.

The sample application starts with the generation of a stereo *sin* tone at 500Hz and stores the generated samples into a proper buffer. The *duration* and the *sample_rate* of the generated samples are expected as parameters. Algorithm 1 shows the samples generation process.

Algorithm 1 Algorithm for Samples Generation

Input: duration, sample_rate

Output: buffer

Initialisation:

1: buffer = ARRAY [duration*sample_rate] OF Byte

LOOP Process

2: **for** $i = 0$ to $duration * sample_rate$ **do**

3: buffer[i] $\leftarrow \sin(2\pi * frequency/sample_rate * i)$

4: **end for**

5: **return** buffer

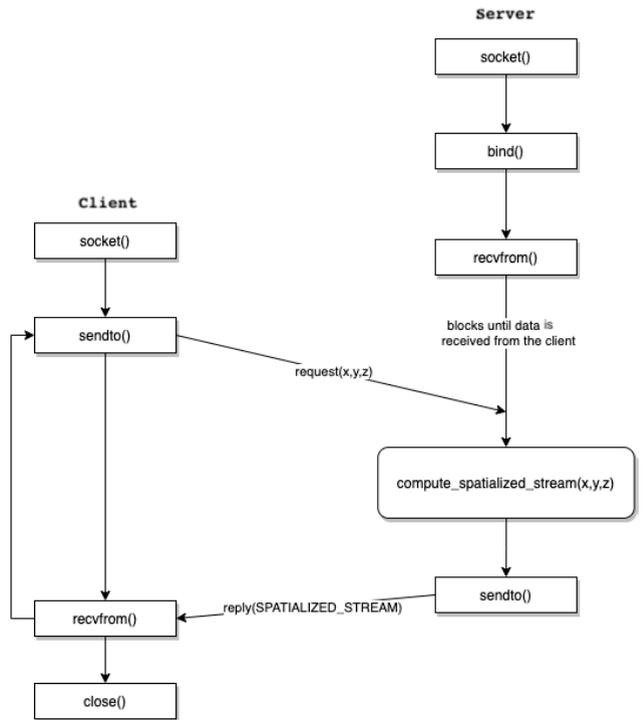


FIGURE 3. Flow diagram of the spatialization server application.

Once a socket endpoint has been set-up the main loop of the spatialization server application acts as summarized by Algorithm 2.

Algorithm 2 Spatialization Server Application Main Loop

Input: TCP_port

Initialisation:

1: samples = **Algorithm1**

2: socket = init(TCP_port)

LOOP Process

3: **while** true **do**

4: receive audio source position

5: **while** samples is not empty **do**

6: buffer = compute_spatial_stream(samples,x,y,z)

7: send buffer through socket

8: **end while**

9: **end while**

The main loop of the application is waiting for audio source position parameters sent through socket by the connected client and computes the spatialized streams. When a set of spatialized samples is ready it is transmitted to the client. The flow diagram shown in Figure 3 summarizes the spatialization server application behavior.

VII. PERFORMANCE ANALYSIS & RESULTS

In order to offer a performance analysis of the advantages introduced by MEC, in contrast to what happens with a cloud

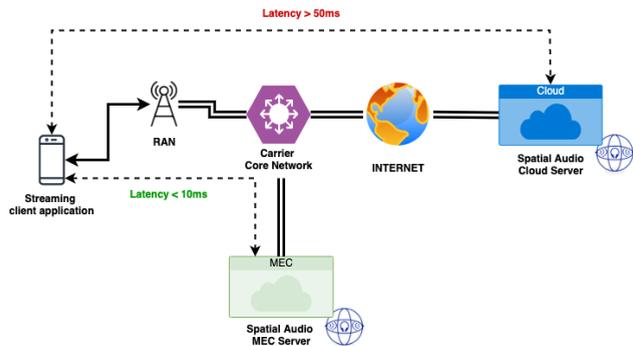


FIGURE 4. MEC vs cloud service deployment.

deployment, as shown in Figure 4, in the proposed spatial audio use case we adopt the model proposed in [62].

The main key performance indicator for the overall system performance is represented by the maximum achievable throughput to serve the spatial users which can be expressed as follow:

$$R_{max} \leq \min \left\{ \frac{c \times MSS}{RTT \times \sqrt{PL}} \times N, BR \right\} \quad (1)$$

where c is a constant depending on the specific TCP implementation, ack strategy, loss mechanism and congestion avoidance algorithm, [63], whose typical values are comprised between 0.9 and 1.2; MSS is the maximum TCP segment size; RTT is the round trip time between the users and the spatial audio elaboration node; PL is the packet loss ratio; N is the number of users supported with maximum throughput, and BR is the available data-rate, i.e. the network capacity, which in the following we assume to be $1Gbps$ per radio access point. As it can be noticed from Eq. 1, the min operator models a possible bottleneck effect of the network. However, in the considered scenario with private dedicated indoor 5G coverage and high data-rate, this is not likely to happen.

The term $RTT \times \sqrt{PL}$ shows the inverse relation between the experienced throughput and round trip time and packet loss. On this hand, the advantages deriving from MEC are twofold in fact:

- thanks to the possibility to perform the elaboration closer to the users, RTT is reduced
- packets exchanged are not sent over the internet service provider network where packet drops due to congestion can occur generating packet loss, thus MEC has also the advantage to reduce PL

Given the above, and assuming that each user generates one TCP flow, all the flows have the same size, and all the flows are terminated at the same location, the number of spatial audio users supported without any performance impairment can be expressed as:

$$N_U = \frac{R_{max}}{R_U} \quad (2)$$

where R_{max} is given by Equation 1 and R_U is the throughput per single user. R_U depends on the selected audio quality and

can be calculated as:

$$R_U = W \times S \times 2 \quad (3)$$

where W is the sample bit-width whose typical values are 16, 24 and $32bits$; S is the sample rate that we assume equal to $44100Hz$.

Figure 5 shows the number of supported spatial audio users per different audio qualities (16bit, 24bit, and 32bit, respectively) and different network conditions represented by a packet loss ratio varying from 0.01 to 0.001. As expectable, the number of supported users decreases by increasing the quality of the audio, due to the higher required throughput per user. Furthermore, it has an inverse relation with the PL. It is worth mentioning that variations of PL may be related to both the wireless channel conditions (fading, path-loss, presence of obstacles, crowdedness of users in the area) and to congestion into the transport network. In this context, MEC has the two-fold advantage of reducing the RTT and PL related to congestion in the transport network.

The results presented in Figure 5 may represent a framework to adopt some adaptive strategies for dynamic placement of audio contents or audio quality adjustment based on packet loss experienced by the users and targeted number of users (as it happens, for example, in Dynamic Adaptive video Streaming over HTTP (DASH), [64]). However, this aspect is out of the scope of this work and represents a future direction.

In order to quantify the advantage in terms of supported users for the MEC scenario with respect to the cloud case, we introduce the supported users gain metric Γ which is given by

$$\Gamma = \frac{N_{U,MEC}}{N_{U,CLOUD}} \quad (4)$$

where $N_{U,MEC}$ and $N_{U,CLOUD}$ are the number of users supported in the MEC and cloud case, respectively, and are calculated according to Eq. 2 and Eq. 1 by assuming specific values for RTT and PL for the two scenarios. Thus, Eq. 4 can be expressed as follows:

$$\Gamma = \sqrt{\frac{PL_{CLOUD}}{PL_{MEC}}} \times \frac{RTT_{CLOUD}}{RTT_{MEC}} = \sqrt{\alpha} \times \beta \quad (5)$$

where PL_{MEC} and PL_{CLOUD} are PL values for MEC and cloud respectively; RTT_{MEC} and RTT_{CLOUD} are RTT values for MEC and cloud respectively. α and β represent the ratio between cloud and MEC PL and RTT, respectively.

Figure 6 shows the performance gain for different values of the parameters α and β . Results show that gain of such a system in terms of supported spatial audio users grow more rapidly with the reduction of round trip time compared to a reduction of packet loss. In other words, to move the function responsible for the audio spatialization at the edge of the network, is more convenient when the edge elaboration reduces significantly the experienced latency. When the elaboration at the edge introduces a reduction of ten times for both the

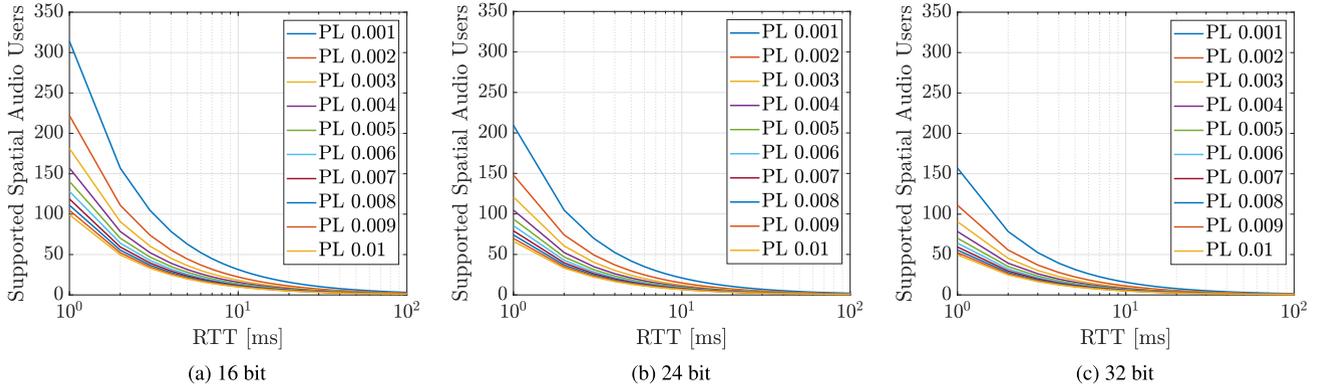


FIGURE 5. Comparison of supported users for different audio sample widths.

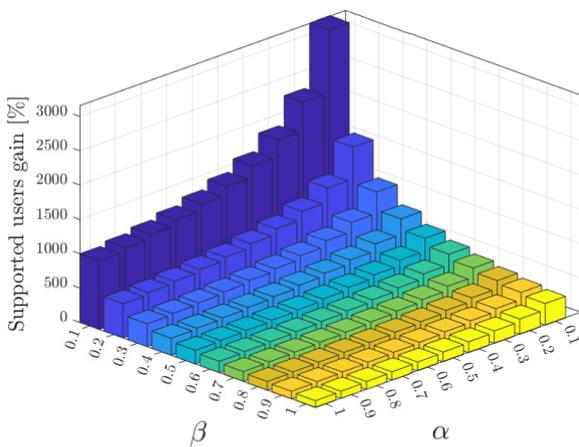


FIGURE 6. Gain in terms of supported spatial audio users.

latency and the packet loss, a MEC approach is able to support a number of users up to 30 times larger with respect to a cloud one.

As α and β can be derived by observing the network behavior and user performance, the results shown in Figure 6 can be utilized by the MEC platform manager to perform the deployment of the spatial audio service at the edge based on achievable gain and targeted performance.

VIII. CONCLUSION

This paper presented a possible sound spatialization application for cultural heritage enjoyment exploiting many advantages of the 5G architecture. The proposed solution allows to guarantee almost real-time sound spatialization as a function of users’ movements for the most generic client device. This can be achieved by exploiting the MEC paradigm, allowing to move the computational load from the client to the edge without involving the core network. Results also demonstrate that this scenario is able to serve an increased number of users with reference to a cloud based one, thus making this solution even more appealing for the CH Italian scenario, which comprises many big museums with a high daily number of visitors. From a more general point of view,

the presented application is another demonstration of the fundamental importance of the MEC paradigm as an enabler of new services and business models.

REFERENCES

- [1] *5G white paper*, NGMN Alliance, Next Gener. Mobile Netw., U.K., vol. 1, 2015.
- [2] *5G Pan-European Trials Roadmap Strategy*, 5G Infrastruct. Assoc., Brussels, Belgium, 2017.
- [3] A. Palombini, “Storytelling and telling history. Towards a grammar of narratives for cultural heritage dissemination in the digital era,” *J. Cultural Heritage*, vol. 24, pp. 134–139, Mar. 2017.
- [4] E. Madirov and S. Absalyamova, “The influence of information technologies on the availability of cultural heritage,” *Proc. Social Behav. Sci.*, vol. 188, pp. 255–258, May 2015.
- [5] A. C. Addison, “Emerging trends in virtual heritage,” *IEEE Multimedia-Mag.*, vol. 7, no. 2, pp. 22–25, Apr. 2000.
- [6] J. Malpas, “Cultural heritage in the age of new media,” in *New Heritage: New Media and Cultural Heritage*, Y. Kalay, T. Kvan, and J. Affleck, Eds. Abingdon, U.K.: Taylor & Francis, 2008.
- [7] M. Roussou, “High-end interactive media in the museum,” in *Proc. ACM SIGGRAPH Conf. Abstr. Appl. (SIGGRAPH)*, vol. 8, 1999, pp. 59–62.
- [8] F. Bernardini, H. Rushmeier, I. M. Martin, J. Mittleman, and G. Taubin, “Building a digital model of Michelangelo’s Florentine Pieta,” *IEEE Comput. Graph. Appl.*, vol. 22, no. 1, pp. 59–67, Jan. 2002.
- [9] G. Guidi, L. Micoli, M. Russo, B. Frischer, M. De Simone, A. Spinetti, and L. Carosso, “3D digitization of a large model of imperial Rome,” in *Proc. 5th Int. Conf. 3-D Digit. Imag. Model. (3DIM)*, 2005, pp. 565–572.
- [10] K. Ikeuchi, A. Nakazawa, K. Hasegawa, and T. Oishi, “The great Buddha project: Modeling cultural heritage for VR systems through observation,” in *Proc. ISMAR*, vol. 3, 2003, pp. 7–16.
- [11] E. Pietroni, A. Pagano, and C. Rufa, “The etruscanning project: Gesture-based interaction and user experience in the virtual reconstruction of the Regolini-Galassi Tomb,” in *Proc. Digit. heritage Int. Congr. (DigitalHeritage)*, vol. 2, Oct. 2013, pp. 653–660.
- [12] S. Brusaporci, F. Graziosi, F. Franchi, P. Maiezza, and A. Tata, “Mixed reality experiences for the historical storytelling of cultural heritage,” in *From Building Information Modelling to Mixed Reality*. Cham, Switzerland: Springer, 2020, pp. 33–46.
- [13] *5G Vision and Requirements*, Chin. J. Eng., IMT-2020 (5G) Promotion Group, Kazan, Russia, 2014.
- [14] D. Raj, “Enabling technologies to realise smart mall concept in 5G era,” in *Proc. IEEE Int. Conf. Comput. Intell. Comput. Res. (ICCCIC)*, Dec. 2018, pp. 1–6.
- [15] R. Jurva, M. Matinmikko-Blue, V. Niemelä, and S. Nenonen, “Architecture and operational model for smart campus digital infrastructure,” *Wireless Pers. Commun.*, vol. 2020, pp. 1–18, Mar. 2020.
- [16] S. Doukianou, D. Daylamani-Zad, and I. Paraskevopoulos, “Beyond virtual museums: Adopting serious games and extended reality,” in *Visual Computing for Cultural Heritage*. Cham, Switzerland: Springer, 2020, p. 283.
- [17] D. Borsatti, G. Davoli, W. Cerroni, and C. Raffaelli, “Enabling industrial IoT as a service with multi-access edge computing,” *IEEE Commun. Mag.*, vol. 59, no. 8, pp. 21–27, Aug. 2021.

- [18] L. D'Errico, F. Franchi, F. Graziosi, A. Marotta, C. Rinaldi, M. Boschi, and A. Colarieti, "Structural health monitoring and earthquake early warning on 5G uRLLC network," in *Proc. IEEE 5th World Forum Internet Things (WF-IoT)*, Apr. 2019, pp. 783–786.
- [19] A. Tessei, M. Luise, P. Pagano, and J. Ferreira, "Secure multi-access edge computing assisted maneuver control for autonomous vehicles," in *Proc. IEEE 93rd Veh. Technol. Conf. (VTC-Spring)*, Apr. 2021, pp. 1–6.
- [20] X. Jiang, F. R. Yu, T. Song, and V. C. M. Leung, "A survey on multi-access edge computing applied to video streaming: Some research issues and challenges," *IEEE Commun. Surveys Tuts.*, vol. 23, no. 2, pp. 871–903, Mar. 2021.
- [21] W. Fohl, J. Reichardt, and J. Kuhr, "A system-on-chip platform for HRTF-based realtime spatial audio rendering with an improved realtime filter interpolation," *Int. J. Adv. Intell. Syst.*, vol. 4, no. 3, pp. 309–317, 2011.
- [22] G. Coluccelli, V. Loffredo, L. Monti, M. R. Spada, F. Franchi, and F. Graziosi, "5G Italian MISE trial: Synergies among different actors to create a '5G road,'" in *Proc. IEEE 4th Int. Forum Res. Technol. Soc. Ind. (RTSI)*, Sep. 2018, pp. 1–4.
- [23] A. Lapini, G. Calamai, F. Argenti, and M. Carfagni, "Application of binaural audio techniques for immersive fruition of cultural heritage," *IOP Conf. Mater. Sci. Eng.*, vol. 364, Jun. 2018, Art. no. 012099.
- [24] D. D'Auria, D. Di Mauro, D. M. Calandra, and F. Cutugno, "A 3D audio augmented reality system for a cultural heritage management and fruition," *J. Digit. Inf. Manage.*, vol. 13, no. 4, p. 203, Aug. 2015.
- [25] D. Murphy and F. Neff, "Spatial sound for computer games and virtual reality," in *Game Sound Technology and Player Interaction: Concepts and Developments*. Hershey, PA, USA: IGI Global, 2011, pp. 287–312.
- [26] D. Poirier-Quinot et al. *Ghost Orchestra a Virtual Reconstruction of the Acoustics of Notre Dame Cathedral*. Accessed: Jun. 20, 2021. [Online]. Available: <https://www.lam.jussieu.fr/Projets/index.php?page=GhostOrchestra>
- [27] M. R. Schroeder and B. F. Logan, "'Colorless' artificial reverberation," *J. Audio Eng. Soc.*, vol. 9, no. 3, pp. 192–197, Jul. 1961.
- [28] V. Välimäki, J. D. Parker, L. Savioja, J. O. Smith, and J. S. Abel, "Fifty years of artificial reverberation," *IEEE Trans. Audio Speech Language Process.*, vol. 20, no. 5, pp. 1421–1448, Jul. 2012.
- [29] L. Nacke and M. Grimshaw-Aagaard, "Player-game interaction through affective sound," in *Games Computing and Creative Technologies*. Harshey, NY, USA: Inf. Sci. Reference, 2011.
- [30] M. S. Smith, "Strategies for the creation of spatial audio in electroacoustic music," Ph.D. dissertation, Division Composition Stud., Univ. North Texas, Denton, TX, USA, 2018.
- [31] J. Hong, J. He, B. Lam, R. Gupta, and W.-S. Gan, "Spatial audio for sound-scape design: Recording and reproduction," *Appl. Sci.*, vol. 7, no. 6, p. 627, Jun. 2017. [Online]. Available: <https://www.mdpi.com/2076-3417/7/6/627>
- [32] T. Kuppanda, N. Degara, D. Worrall, B. Thoshkakhna, and M. Müller, "Virtual reality platform for sonification evaluation," Georgia Inst. Technol., Atlanta, GA, USA, 2015.
- [33] M. Geier, S. Spors, and S. Weinzierl, "The future of audio reproduction," in *Proc. Int. Workshop Adapt. Multimedia Retr.* Leuven, Belgium: Springer, 2008, pp. 1–17.
- [34] C. P. Brown and R. O. Duda, "A structural model for binaural sound synthesis," *IEEE Trans. Speech Audio Process.*, vol. 6, no. 5, pp. 476–488, Sep. 1998.
- [35] V. Välimäki, J. Parker, L. Savioja, J. O. Smith, and J. Abel, "More than 50 years of artificial reverberation," in *Audio Engineering Society Conference: 60th International Conference: Dreams (Dereverberation and Reverberation of Audio, Music, and Speech)*. Berlin, Germany: Audio Eng. Soc., 2016.
- [36] W.-S. Gan, S. Peksi, J. He, R. Ranjan, N. D. Hai, and N. K. Chaudhary, "Personalized HRTF measurement and 3D audio rendering for AR/VR headsets," *J. Audio Eng. Soc.*, May 2017.
- [37] P. Majdak and M. Noisternig, *AES69-2015: AES Standard for File Exchange-Spatial Acoustic Data File Format*, Audio Eng. Soc. Conv., New York, NY, USA, Standard, 2015.
- [38] MySOFA. *Head Related Transfer Functions (HRTF) Locating Sounds in 3D With Your Headphones*. Accessed: Jun. 20, 2021. [Online]. Available: <https://mysofa.audio/index.html>
- [39] K. Hartung, J. Braasch, and S. J. Sterbing, "Comparison of different methods for the interpolation of head-related transfer functions," in *Proc. 16th Int. Conf., Spatial Sound Reproduction*, 1999, pp. 319–329.
- [40] W. Zhang, P. Samarasinghe, H. Chen, and T. Abhayapala, "Surround by sound: A review of spatial audio recording and reproduction," *Appl. Sci.*, vol. 7, no. 5, p. 532, May 2017. [Online]. Available: <https://www.mdpi.com/2076-3417/7/5/532>
- [41] C. Pike and F. Melchior, "An assessment of virtual surround sound systems for headphone listening of 5.1 multichannel audio," *J. Audio Eng. Soc.*, May 2013.
- [42] A. McKeag and D. S. McGrath, "Sound field format to binaural decoder with head tracking," in *Audio Engineering Society Convention 6R*. Audio Eng. Soc., 1996.
- [43] M. Gorzel, A. Allen, I. Kelly, J. Kammerl, A. Gungormusler, H. Yeh, and F. Boland, "Efficient encoding and decoding of binaural sound with resonance audio," in *Proc. AES Int. Conf. Immersive Interact. Audio*, 2019, pp. 1–12.
- [44] R. shukla, I. T. radu, M. sandler, and R. stewart, "Real-time binaural rendering with virtual vector base amplitude panning," *J. Audio Eng. Soc.*, vol. 45, no. 6, pp. 1–10, Mar. 2019.
- [45] B. Cowan and B. Kapralos, "Spatial sound for video games and virtual environments utilizing real-time GPU-based convolution," in *Proc. Conf. Future Play Res., Play, Share Future Play*, 2008, pp. 166–172.
- [46] B. Kapralos and N. Mekuz, "Application of dimensionality reduction techniques to HRTFS for interactive virtual environments," in *Proc. Int. Conf. Adv. Comput. Entertainment Technol. (ACE)*, 2007, pp. 256–257, doi: 10.1145/1255047.1255112.
- [47] Molex and Dimensional Research. (Mar. 2021). *Timing is Everything: Carriers and the State of 5G. A Survey of R&D, Engineering, and Product Stakeholders*. Accessed: Jul. 16, 2021.
- [48] B. Ethirajulu. *How 5G and Edge Computing can Enhance Virtual Reality*. Accessed: Jul. 16, 2021. [Online]. Available: <https://www.ericsson.com/en/blog/2020/4/how-5g-and-edge-computing-can-enhance-virtual-reality>
- [49] B. O. Donnel. (Mar. 2020). *How 5G could Make Augmented Reality and Mixed Reality More Real*. Accessed: Jul. 16, 2021. [Online]. Available: <https://eu.usatoday.com/story/tech/columnist/2020/05/29/5-g-how-could-make-augmented-reality-more-real/5275119002/>
- [50] (Mar. 2020). *Immersive Experiences With 5G, Information can Surround us in a 3D Digital World*. Accessed: Jul. 16, 2021. [Online]. Available: <https://www.nokia.com/networks/5g/use-cases/immersive-experiences/>
- [51] J. Stanislow. (Mar. 2020). *5G Impact on Virtual Reality (VR) and Augmented Reality (AR)*. Accessed: Jul. 16, 2021. [Online]. Available: <https://www.linkedin.com/pulse/5g-impact-virtual-reality-vr-augmented-ar-jeff-stanislaw/>
- [52] Y. Siritwardhana, P. Porambage, M. Liyanage, and M. Ylianttila, "A survey on mobile augmented reality with 5G mobile edge computing: Architectures, applications, and technical aspects," *IEEE Commun. Surveys Tuts.*, vol. 23, no. 2, pp. 1160–1192, 2nd Quart., 2021.
- [53] R.-S. Schmoll, S. Pandi, P. J. Braun, and F. H. Fitzek, "Demonstration of VR/AR offloading to mobile edge cloud for low latency 5G gaming application," in *Proc. 15th IEEE Annu. Consum. Commun. Netw. Conf. (CCNC)*, Jan. 2018, pp. 1–3.
- [54] K. Katsaros, D. Gkounis, D. Kaleshi, B. Thomas, J. Harris, H. Falaki, and D. Simeonidou, "Enhancing tourist experiences through 5G—The 5G smart tourism case study," in *Proc. IEEE 2nd 5G World Forum (5GWF)*, Oct. 2019, pp. 471–476.
- [55] X. Qiao, P. Ren, S. Dustdar, L. Liu, H. Ma, and J. Chen, "Web AR: A promising future for mobile augmented reality—State of the art, challenges, and insights," *Proc. IEEE*, vol. 107, no. 4, pp. 651–666, Apr. 2019.
- [56] X. Qiao, P. Ren, G. Nan, L. Liu, S. Dustdar, and J. Chen, "Mobile web augmented reality in 5G and beyond: Challenges, opportunities, and future directions," *China Commun.*, vol. 16, no. 9, pp. 141–154, Sep. 2019.
- [57] M. Cuevas-Rodríguez, L. Picinali, D. González-Toledo, C. Garre, E. de laRubia-Cuestas, L. Molina-Tanco, and A. Reyes-Lecuona, "3d tune-in toolkit: An open-source library for real-time binaural spatialization," *PLoS ONE*, vol. 14, no. 3, pp. 1–37, 2019, doi: 10.1371/journal.pone.0211899.
- [58] Various. *Sofalizer for Unity*. Accessed: Apr. 1, 2021. [Online]. Available: <https://github.com/sofocoustics/SOFALizer-for-Unity>
- [59] Various. *Spatial Audio Framework*. Accessed: Apr. 1, 2021. [Online]. Available: https://github.com/leomccormack/Spatial_Audio_Framework
- [60] Various. *Resonance Audio SDK*. Accessed: Jun. 16, 2021. [Online]. Available: <https://github.com/resonance-audio/resonance-audio>
- [61] B. Schäling, *The Boost C++ Libraries*. Denver, CO, USA: Boris Schäling, 2011.
- [62] F. Vatalaro and G. Ciccarella, "A network paradigm for very high capacity mobile and fixed telecommunications ecosystem sustainable evolution," *IEEE Access*, vol. 8, pp. 135075–135090, 2020.

- [63] M. Mathis, J. Semke, J. Mahdavi, and T. Ott, "The macroscopic behavior of the TCP congestion avoidance algorithm," *SIGCOMM Comput. Commun. Rev.*, vol. 27, no. 3, pp. 67–82, Jul. 1997, doi: 10.1145/263932.264023.
- [64] T. C. Thang, Q.-D. Ho, J. W. Kang, and A. T. Pham, "Adaptive streaming of audiovisual content using MPEG DASH," *IEEE Trans. Consum. Electron.*, vol. 58, no. 1, pp. 78–85, Feb. 2012.



CLAUDIA RINALDI received the Laurea degree (*cum laude*) in electronic engineering from the University of L'Aquila, Italy, in 2005, the master's degree in trumpet from the Conservatory of Music of L'Aquila, in 2006, the Ph.D. degree in electronic engineering from the University of L'Aquila, in 2009, and the bachelor's degree in electronic music from the Conservatory of Music of L'Aquila, in 2013.

She is currently an Assistant Professor at the Department of Information Engineering, Computer Science and Mathematics, University of L'Aquila. Her main research interests include digital signal processing algorithms and more in general on the use of technology in artistic fields. Moreover, she is concerned with design, modeling, and optimization of communication algorithms with particular emphasis on the physical layer and software defined radio for the development of transmission systems responding to cognitive radios paradigms.



FABIO FRANCHI (Member, IEEE) was born in L'Aquila, Italy, in 1985. He received the M.Sc. degree in computer engineering and the Ph.D. degree in information and communications technology from the University of L'Aquila, Italy, in 2012 and 2018, respectively.

He is currently a Research Fellow at the Department of Information Engineering, Computer Science and Mathematics, University of L'Aquila.

During his graduate studies, he collaborated with the CNR-ISTI in the computer vision research area. He performs research on next generation services based on ultra Reliable and Low Latency Communications (uRLLC), enhanced Mobile Broadband (eMBB), and massive Machine Type Communications (mMTC). He works as an Invited Lecturer in cloud architecture within the master's degree in telecommunications engineering at the University of L'Aquila.

Dr. Franchi is an AWS Academy Accredited Educator and listed in the Engineer professional register 'Ordine degli Ingegneri della provincia dell'Aquila.'



ANDREA MAROTTA (Member, IEEE) received the M.Sc. degree in computer engineering and the Ph.D. degree in information and communications technology from the University of L'Aquila, Italy, in 2015 and 2019, respectively.

During his graduate studies, he was a Visiting Student at the Group for Research on Wireless (GROW), Instituto Superior Técnico/University of Lisbon, performing research on C-RAN deployment optimization. In 2017, he spent one year as a

Visiting Researcher at the research Institute of Communication, Information and Perception Technologies (TeCIP), Scuola Superiore di Studi Universitari e di Perfezionamento Sant'Anna, Pisa, performing research on 5G radio access network design. In 2019, he spent a research period at the Computer Networks Laboratory (NetLab), University of California, Davis, performing research on slice reliability. He is currently an Assistant Professor at the Department of Information Engineering, Computer Science and Mathematics, University of L'Aquila. He has actively participated in the TPC of IEEE conferences, such as IEEE Global Communications Conference (GLOBECOM), IEEE 5G World Forum (WF-5G), and IEEE International Forum on Research and Technologies for Society and Industry (RTSI). He performs research on network reliability, SDM optical networks, multi-access edge computing, 5G software defined access, and CoMP coordinated scheduling.



FABIO GRAZIOSI (Member, IEEE) was born in L'Aquila, Italy, in 1968. He received the Laurea (*cum laude*) and Ph.D. degrees in electronic engineering from the University of L'Aquila, Italy, in 1993 and 1997, respectively.

Since February 1997, he has been with the Department of Electrical and Information Engineering, University of L'Aquila, where he is currently a Full Professor in communications systems. The research activity led to more than 200

publications on international journals and conferences and was initially focused on modeling and performance evaluation of wireless systems in complex propagation scenarios. Subsequently, the research approach has been enriched thanks to the contamination with other scientific areas, such as electronics, computer science, and control systems. Research activities on wireless sensor networks and networked embedded systems actually fall within this exciting and multidisciplinary context, which has been progressively extended to scientific areas, historically far from communications, such as structural engineering.



CARLO CENTOFANTI (Graduate Student Member, IEEE) received the bachelor's degree in ICT engineering and the M.Sc. degree in computer science engineering from the University of L'Aquila, Italy, where he is currently pursuing the Ph.D. degree in information and communication technology with the Department of Information Engineering, Computer Science and Mathematics.

He has actively contributed to ECSEL-RIA AfarCloud project designing and defining the system's software architecture. Since 2020, he has been involved in MSCA-RISE OPTIMIST Project, where he works on multi-access edge computing, adaptive video streaming, and virtualization research topics. He is an AWS Academy Accredited Educator. He won several competition notices for fixed-term research fellow positions from the University of L'Aquila.

...