IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 11 May 2023, accepted 10 June 2023, date of publication 21 June 2023, date of current version 28 June 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3288334

TOPICAL REVIEW

5G Multi-Access Edge Computing: A Survey on Security, Dependability, and Performance

GIANFRANCO NENCIONI[1], ROSARIO G. GARROPPO[2], AND RUXANDRA F. OLIMID[3]
[1]Department of Electrical Engineering and Computer Science, University of Stavanger, 4021 Stavanger, Norway
[2]Department of Information Engineering, University of Pisa, 56126 Pisa, Italy
[3]Department of Computer Science, University of Bucharest, 030018 Bucharest, Romania

Corresponding author: Gianfranco Nencioni (gianfranco.nencioni@uis.no)

This work was supported in part by the Norwegian Research Council through the Management and Orchestration for Data and Network Integration (5G-MODaNeI) Project under Grant 308909, and in part by the Italian Ministry of Education and Research (MIUR) in the framework of the Future-Oriented Research Laboratory (FoReLab) Project (Departments of Excellence).

ABSTRACT The Fifth Generation (5G) of mobile networks offers new and advanced services with stricter requirements. Multi-access Edge Computing (MEC) is a key technology that enables these new services by deploying multiple devices with computing and storage capabilities at the edge of the network, close to end-users. MEC enhances network efficiency by reducing latency, enabling real-time awareness of the local environment, allowing cloud offloading, and reducing traffic congestion. New mission-critical applications require high security and dependability, which are rarely addressed alongside performance. This survey paper fills this gap by presenting 5G MEC's three aspects: security, dependability, and performance. The paper provides an overview of MEC, introduces taxonomy, state-of-the-art, and challenges related to each aspect. Finally, the paper presents the challenges of jointly addressing these three aspects.

INDEX TERMS 5G, MEC, security, dependability, performance.

I. INTRODUCTION

The Fifth Generation (5G) of mobile networks is currently under deployment. The main innovation is the provision of wireless connectivity for various usage scenarios [1]: *enhanced Mobile Broadband (eMBB)*, for services with very high data rate requirements (up to 20Gb/s); *massive Machine-Type Communication (mMTC)*, developed for connecting a huge number of Internet-of-Things (IoT) devices (up to one million devices/km^2); *Ultra-Reliable Low-Latency Communication (URLLC)*, for services requiring high reliability and very low latency (up to 1ms). eMBB allows improving the services provided by the Fourth Generation (4G) of mobile networks. mMTC enhances the services that are now provided by Low-Power Wide Area Networks, such as Long-Term Evolution MTC (LTE-M) and Narrowband IoT (NB-IoT). URLLC enables innovative advanced services, such as mission-critical applications, industrial automation, and enhanced Vehicular to Everything (V2X) such as platooning or remote driving. eMBB services are under deployment, and 5G smartphones have already been produced and sold by many manufacturers. Currently, mMTC is not under deployment since many network operators have deployed LTE-M and NB-IoT in relatively recent times. URRLC is a usage scenario that is more immature and challenging, and it is attracting a lot of attention from the research community.

One of the technologies that enable 5G to provide URLLC services is the *Multi-Access Edge Computing (MEC)*. MEC consists in the deployment of storage and computing platforms at the edge of the (radio) access network. In this way, MEC is enabling the delivery of services with low latency but can also enable context awareness and task offloading. Moreover, MEC is also the enabler of the edge intelligence, which is anticipated to be one of the main innovations of the Sixth Generation (6G) of mobile networks [2].

MEC is the name given by the European Telecommunications Standards Institute (ETSI) which has an Industry Specification Group (ISG) [3] that is standardizing MEC since 2014. Before 2017, MEC was standing for ''Mobile Edge Computing'', but ETSI decided to generalize the standard to other access technologies, not only 4G and 5G but also

The associate editor coordinating the review of this manuscript and approving it for publication was Filbert Juwono.

fixed-access networks and Wireless Local Area Networks (WLAN).

ETSI MEC is not the only standardization effort on edge computing. *Fog computing* and *cloudlet* are the two main alternatives. The cloudlet was proposed in 2009 [4] and can be considered the first effort on edge computing. The cloudlet consists of a micro-cloud close to the mobile device. Fog computing has been first proposed by Cisco in 2011. Since 2015, fog computing is promoted and standardized by the OpenFog consortium [5]. Fog computing has been introduced as an extension of the cloud computing paradigm from the core to the edge of the network. Fog Computing consists of a three-layer architecture where clouds, fog nodes, and IoT devices interact. There is also another technology called Mobile Cloud Computing (MCC), but it is not edge computing since it consists of offloading tasks from mobile users to the cloud. This paper uses as a reference the ETSI MEC, but many considerations can also be generalized for the other edge computing solutions, and, in our study, works on alternative edge computing are also included.

As already mentioned, URLLC is the most innovative and challenging usage scenario for 5G and the one that requires MEC. To support URLLC, MEC has to cope with high requirements of ultra-reliability, which means security and dependability, and low latency, which is a performance indicator. This is one of the reasons for which security, dependability, and performance are three critical aspects of MEC.

While MEC security has been investigated to some extent in recent years and several surveys are available, there are fewer surveys available on performance and even less on dependability. To the best knowledge of the authors, this is the first work that jointly investigates the security, dependability, and performance of 5G MEC.

This paper has the following contributions:

- State of the art and challenges on the security, dependability, and performance of 5G MEC. Each aspect is addressed *individually* by using a similar structure and content organization. This organization helps to better jointly investigate and compare the three aspects.
  - First, the *taxonomy* of the investigated aspect is introduced. In this way, experts on the other aspects can better understand the investigated aspect.
  - Second, the *state of the art* is presented. The state of the art is divided into standardization efforts and academic publications.
  - Finally, the *challenges* are presented and organized according to the ETSI MEC architecture.
- Challenges in *jointly* addressing security, dependability, and performance in 5G MEC. The joint provision of these three aspects and the related trade-offs are analyzed and discussed by including also the future perspective of 6G.

The paper flow is depicted in Figure 1. The next section introduces the necessary background concepts and definitions of 5G MEC. Sections III, IV, and V present the state of the art and the challenges of 5G MEC related to security, dependability, and performance, respectively. Section VI discusses the challenges and trade-offs of jointly addressing security, dependability, and performance aspects. Finally, Section VII presents the conclusions.

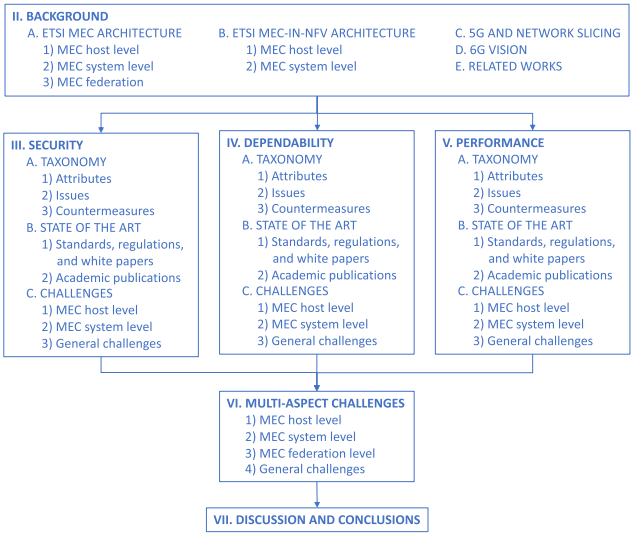

FIGURE 1. Paper flow.

II. BACKGROUND

In the ETSI specifications [6], MEC is defined as *"system which provides an IT service environment and cloud-computing capabilities at the edge of the access network which contains one or more type of access technology, and in close proximity to its users"*.

Before presenting the state of the art of MEC focusing on the three different aspects, the fundamental concepts of MEC and the related enabling technologies are presented.

Table 1 lists the main acronyms that will be used in the rest of the paper. Most of the MEC acronyms are defined in [6] and [7]. Note that what is currently defined as MEC was previously defined as Mobile Edge. For example, MEO was the acronym for Mobile Edge Orchestrator.

A. ETSI MEC ARCHITECTURE

A MEC system is defined as a collection of *MEC Hosts* (MEHs) and *MEC management* necessary to run MEC applications [6]. Figure 2 illustrates the ETSI MEC general reference architecture, divided into two levels: the *MEC host level* and the *MEC system level* [7].

1) MEC HOST LEVEL

A MEH contains a *virtualization infrastructure* that provides computation, storage, and networking resources to run *MEC applications* (MEC Apps) and a *MEC Platform* (MEP). MEC applications run as Virtual Machines (VMs) and can offer and consume *MEC services* [7]. The MEP is a collection

TABLE 1. List of acronyms.

5G	Fifth Generation of mobile networks
BSS	Business Support System
CN	Core Network
CFS	Customer-Facing Service
DNS	Domain-Name System
EM	Element Management
ETSI	European Telecommunications Standards Institute
GR	Group Report (ETSI)
GS	Group Specification (ETSI)
LCM	Life-Cycle Management
MANO	Management and Orchestration
MEAO	MEC Application Orchestrator
MEC	Multi-access Edge Computing
MEF	MEC Federator
MEFB	MEC Federation Broker
MEFM	MEC Federation Manager
MEH	MEC Host
MEO	MEC Orchestrator
MEP	MEC Platform
MEPM	MEC Platform Manager
MEPM-V	MEC Platform Manager - NFV
NFV	Network Function Virtualization
NFVI	NFV Infrastructure
NFVO	NFV Orchestrator
OSS	Operation Support System
RAN	Radio Access Network
SDN	Software-Defined Networking
SST	Slice/Service Type
VIM	Virtualization/Virtualized Infrastructure Manager
VM	Virtual Machine
VNF	Virtualized Network Function
VNFM	VNF Manager

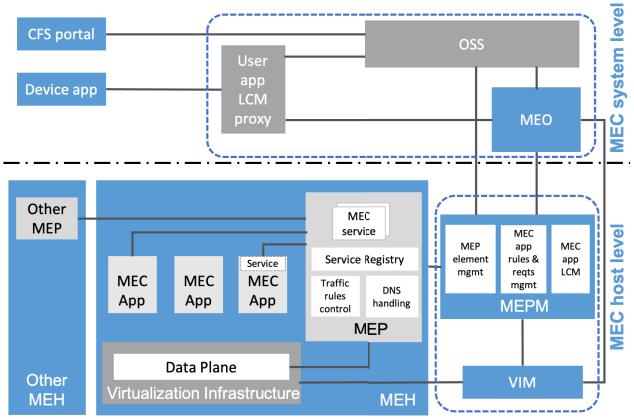

FIGURE 2. ETSI MEC reference architecture [7].

of functionalities required to run the MEC applications. The MEP can also host MEC services. More information on MEC applications can be found in [8] (development) and in [9] (enablement).

The *MEC host-level management* is composed of the *MEC Platform Manager* (MEPM) and the *Virtualization Infrastructure Manager* (VIM). The MEPM manages MEC applications' life cycle, rules, and requirements (e.g., required resources, latency), and provides element management functions to the MEP. The VIM manages the allocation and monitoring of the virtual resources and transmits faults and performance reports to the MEPM. OpenStack is commonly used to implement VIM [10].

2) MEC SYSTEM LEVEL

The *MEC system-level management* is composed of the *MEC Orchestrator* (MEO), the operator's *Operations Support System* (OSS), and the *user application Life-Cycle Management (LCM) proxy* [7].

The MEO is the core component that has an overview of the complete MEC system. MEO on-boards the application packages (performs integrity and authenticity checks, as well as compliance with the operator policies), selects the appropriate MEH(s) for MEC application instantiation based on the related rules and requirements, and triggers the MEC application instantiation, termination, and relocation (if needed).

The operator's OSS receives requests from the *Customer-Facing Service (CFS) portal* or from the *device applications*, via the *user application LCM proxy*. The CFS portal allows operators' third-party customers to select and order MEC applications or receive service level information concerning provisioned applications. A device application is an application in a device that can interact with the MEC system. In response to a user request via a device application, an *user application* is instantiated in the MEC system. The user application LCM proxy allows the device applications to request the on-boarding, instantiation, and termination of user applications in the MEC system. If the OSS decides to grant a request it transmits it to the MEO for further processing.

3) MEC FEDERATION

The MEC architecture can be extended to allow inter-MEC system communication. To this purpose, the *MEC federation* is defined as *"a federated model of MEC systems enabling shared usage of MEC services and applications"* [11].

The MEC architecture in Figure 2 can the extended by adding a *MEC Federator* (MEF), which may be composed of the functionalities of the *MEC Federation Broker* (MEFB) and the *MEC Federation Manager* (MEFM). The MEF interfaces with other MEFs enabling the information exchange. It also interfaces with at least one MEO (the one the MEF is belonging to). A MEF may serve as a single point of contact for multiple MEFs acting as a broker between MEFs [7].

Table 2 summarizes the functionalities of the main architectural elements, as explained in [7].

B. ETSI MEC-IN-NFV ARCHITECTURE

MEC uses a virtualization platform to run the MEC application in the MEH. NFV is a virtualization platform where the network functions are decoupled from the hardware by using virtual hardware abstraction. It is, therefore, beneficial to reuse the infrastructure and the infrastructure management of NFV [12].

Figure 3 illustrates the ETSI NFV reference architecture [13]. To provide network services, the NFV is composed

TABLE 2. Key elements in the MEC architecture.

Element	Functionalities
MEH	
MEP	- Offers the necessary environment for the MEC applications. - Hosts MEC services. - Instructs the data plane based on received traffic rules. - Configures the DNS proxy/server. - Provides access to persistent storage. - Provides timing information.
Virtualization Infrastructure	- Provides compute, storage, and network resources for the MEC applications. - Data plane routes traffic among MEC applications, services, DNS proxy/server, and various access networks.
MEC App	- Discovers, advertises, consumes, and offers MEC services. - Performs support procedures related to the life cycle of the application.
MEC Host-level Management	
MEPM	- Manages the life cycle, rules, and requirements of the applications. - Prepares the virtualization infrastructure to run a software image. - Provides element management functions to the MEP.
VIM	- Manages the allocation and monitoring of the virtual resources. - Prepares the virtualization infrastructure to run a software image. - Transmits faults and performance reports to the MEPM.
MEC System-level Management	
MEO	- Maintains an overview of the complete MEC system. - On-boards the application packages. - Selects the appropriate MEH(s) for MEC application instantiation. - Triggers MEC application instantiation, termination, and relocation.
OSS	- Decides on granting requests for instantiation and terminating of applications.
User app LCM proxy	- Permits the device applications to request the on-boarding, instantiation, and termination of user applications. - Informs the device applications about the status of the user applications.
MEC Federation	
MEF	- Registers the MEC system information sent by a MEO. - Discovers MEC systems. - May act as a one to many intermediary between MEFs (broker capability). - Exchanges the MEC system information. - Manages the MEC application lifecycle across different MEC systems. - Monitors the MEC application across different MEC systems.

of *Virtualized Network Functions* (VNFs), an underlying *NFV Infrastructure (NFVI)*, and a *NFV Management and Orchestration* (MANO).

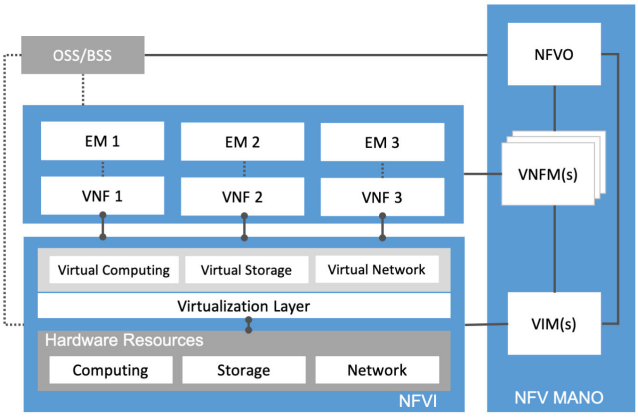

FIGURE 3. ETSI NFV reference architectural framework [13].

A VNF is a software implementation of a network function, which is decoupled from the hardware resources it uses. The VNFs rely on the NFVI, where the needed virtualized resources (computing, storage, and network) are obtained from the hardware resources through the virtualization layer. A VNF can be deployed, by case, over one or several VMs [13], where VMs are partitioned on the resources of a hardware host by software programs called *hypervisors*.

The NFV MANO is composed of three main components: the *NFV Orchestrator* (NFVO), the *VNF Manager* (VNFM), and the *Virtualized Infrastructure Manager* (VIM). The NFVO is the highest hierarchical level of the NFV MANO and is responsible for the creation and LCM of network services. The VNFMs are instead responsible for the LCM of the VNFs, which are locally managed by the *Element Management* (EM) systems. A VNFM can serve one or multiple VNFs. Finally, the VIM controls and manages the NFVI resources (e.g., it is in charge of the inventory of software, computing, storage, and network resources, increasing resources for VMs, improving energy efficiency, collection of infrastructure fault operations, capacity planning, and optimization) [13].

An NFV-based network service is composed of an ordered set of VNFs between two end-points, where the traffic is steered through them. This composition to provide an NFV-based network service is similar to the one specified by the Service Function Chaining (SFC) [14].

Figure 4 illustrates the MEC architecture deployed by using NFV [15]. For continuity and clarity, the architectural changes are explained separately for each of the two layers.

1) MEC HOST LEVEL

On the host side, both the MEC applications and the MEP are deployed as VNFs, while the virtual infrastructure is deployed as NFVI. The virtualization infrastructure, as the NFVI, can be implemented with various virtualization technologies, such as hypervisor-based or container-based solutions, but also mixing or/and nesting virtualization technologies [16]. On the host management side, the MEPM

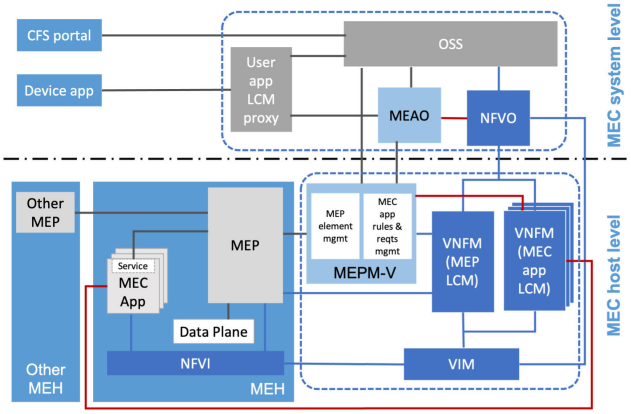

FIGURE 4. MEC-in-NFV reference architecture [15].

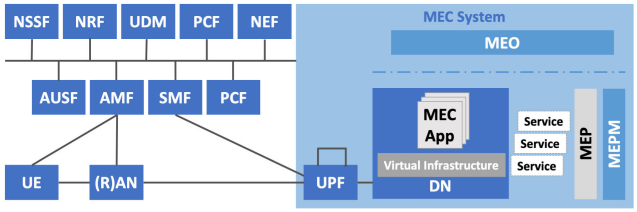

FIGURE 5. Integrated MEC deployment in 5G network [18].

is substituted by the *MEC Platform Manager - NFV* (MEPM-V) and a VNFM. The MEPM-V has the same responsibilities as the MEPM. The VNFM is delegated the management of the VNF life cycle [15]. The VIM maintains similar functionalities.

2) MEC SYSTEM LEVEL

In the MEC-in-NFV architecture, the MEO is replaced by the *MEC Application Orchestrator* (MEAO) and an NFVO. The MEAO has the same responsibilities as the MEO. However, the MEAO delegates an NFVO to perform the resource orchestration and the orchestration of the MEC applications (as VNFs). The other elements remain unaffected [15].

More details about the MEC deployment in an NFV environment can be found in [15].

C. 5G AND NETWORK SLICING

MEC is one of the key technologies of 5G, together with NFV and Software-Defined Networking (SDN) [17]. In particular, MEC provides 5G of contextual information and real-time awareness of the local environment. In [18] and [19], ETSI explains how to deploy and integrate MEC in the 5G system.

Figure 5 depicts the MEC architecture deployed in a 5G network [18]. In the left part of the figure (in dark blue), there is the 5G Service-Based Architecture (SBA), as described by 3GPP in [20], which is composed of the following network function of the 5G Core Network (CN): Network Slice Selection Function (NSSF), Network Resource Function (NRF), Unified Data Management (UDM), Policy Control

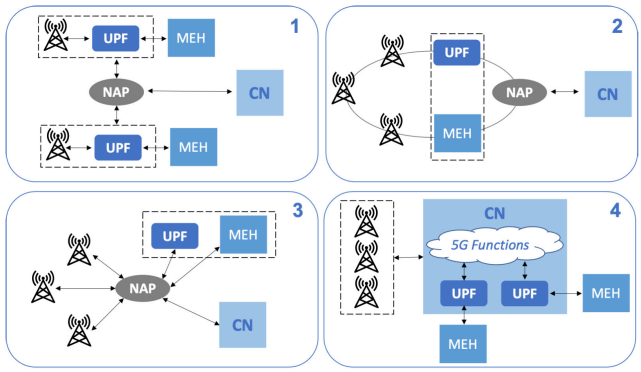

FIGURE 6. 5G-MEC deployment scenarios [18].

Function (PCF), Network Exposure Function (NEF), Application Function (AF), Authentication Server Function (AUSF), Access and Mobility Management Function (AMF), and Session Management Function (SMF). Moreover, the 5G SBA is also composed of User Equipment (UE), Radio Access Network (RAN), User Plane Function (UPF), and Data Network (DN). Each function can consume or produce services.

In 5G SBA, MEC is seen as a set of AFs. The MEO and the MEP are acting as AFs. The MEH is instead often deployed as DN.

The NRF contains the registered network functions and their provided services. The services provided by the MEC applications are instead in the service register in the MEP.

Some of the services are available only through the NEF, which acts as a centralized point for service exposure. The AUSF manages the authentication.

The PCF handles the policies and the rules. The MEP uses the PCF services to impact the traffic steering rules.

The UDM is instead responsible for services related to users and subscriptions. It generates authentication credentials, handles information related to user identification, manages access authorization, and registers users on AMF and SMF.

Finally, connected to the network slicing (which will be presented more in detail later on), the NSSF manages the selection of network slice instances and the allocation of the necessary AMFs.

A key role in the integration of 5G and MEC is performed by the UPF, which can be seen as a distributed and configurable data plane from the MEC.

In [18], four scenarios for the deployment of MEHs in a 5G network are presented (see also Figure 6):

1) MEH and the local UPF collocated with the Base Station;
2) MEH collocated with a transmission node, possibly with a local UPF;
3) MEH and the local UPF collocated with a Network Aggregation Point (NAP);
4) MEH collocated with the CN functions (i.e. in the same data centre).

ETSI has also investigated the deployment of MEC in other access technologies, such as 4G [21] and WLAN [22]. Moreover, ETSI has investigated the deployment in cloud RAN [23].

Initially, the Radiocommunication sector of the International Telecommunication Union (ITU-R) defined three usage scenarios for 5G and beyond [1]: eMMB, mMTC, and URLLC. The use scenarios have been already described in the introduction In 3rd Generation Partnership Project (3GPP), the use scenarios are called Slice/Service Types (SSTs). As ITU-R, 3GPP SSTs include eMBB and URLLC, but they do not include mMTC and there are instead Massive Internet of Things (MIoT), Vehicle-to-everything (V2X) service, and High-performance Machine-Type Communication (HMTC) [20].

The ability of 5G to provide services in very different use scenarios is enabled by network slicing. Network slicing allows the flexible and efficient creation of specialized end-to-end logical networks (network slices) on top of shared network infrastructure. In order to properly operate, the network slices have to be *isolated*. The network slice isolation needs to be valid with respect to security, dependability, and performance. More details on network slice isolation can be found in [24]. In this paper, network slicing is not further described. A more detailed description (together with a security overview) can be found in [25]. In [26], ETSI has defined how MEC supports network slicing.

D. 6G VISION

In the next generation of mobile networks, MEC will still be one of the most important technologies. The key features of 6G will be connected intelligence, programmability, deterministic end-to-end, integrated sensing and communication, sustainability, trustworthiness, scalability, and affordability [57]. As for the previous generations, 6G will introduce improvements of the performance metrics of around one order of magnitude with respect to 5G.

To obtain these enhancements, Artificial Intelligence (AI) and Machine Learning (ML) will be used to enable the system network architecture and control, the edge and ubiquitous computing [58], the radio technology and signal processing [59], the optical networks, the network and service security, the non-terrestrial network communication, and the special-purpose networks/sub-networks [57].

In conclusion, the main innovation of 6G can be summarized as *AI everywhere* to enable the *easy integration of everything*. In this context, MEC has an important role of bringing AI, enabling distributed (micro)service-based architectures, and helping 6G to reach the "zero delay".

E. RELATED WORKS

In the previous subsections, we have extensively referred to documents produced by the ETSI MEC group. Table 3 shows the ETSI specifications, reports, and white papers that have content in one of the three perspectives that are investigated.

Many surveys and reviews have been published in the recent years. Many works focus on MEC (first mobile and then multi-access), referring (at least partially) to the ETSI architecture [60], [61], [62], [63], [64], [65], [66], [67], [68], [69], [70], [71], [72], [73], [74], [75], [76], [77], [78], [79], [80], [81], [82], [83], [84], [85]. Some works focus exclusively or jointly on fog computing [63], [67], [68], [70], [71], [73], [74], [75], [79], [81], [86], [87], [88], [89]. Several works also consider cloudlets [63], [66], [68], [71], [73], [74], [75], [79], [81], [90]. Other works do not focus on any particular architecture but consider generic edge computing [91], [92], [93], [94], [95], [96], [97], [98], [99], [100], [101].

Several works are general surveys [60], [66], [68], [73], [74], [75], [78], [79], [87], [88], [91], [100]. Some works are focusing on specific perspectives: security [70], [76], [82], [94], [98], security and resilience [67], security and efficiency [95], trustworthiness [90], reliability and latency [97], dependability [86], [89], [96]. Other works focus on specific environments: vehicular networks [92], [99], IoT [71], [93], [94], [95], industrial Internet [77]. Finally, other works focus on specific tasks or parts: location trade-off [61], orchestration [62], capabilities on computing, caching, and communication [63], communication [64], computation offload [65], service adoption and provision [69], infrastructure [72], optimization [80], ML [101] tools and applications [81], integration with network slicing [83], resource allocation [85], and implementations [84].

This work focuses on MEC considering as reference the ETSI architecture, but it also considers research on other architectures. It focuses on three perspectives (security, dependability, and performance) individually and jointly. This work is up to date and does not focus on any particular environment or task. All the content related to security, dependability, and performance of the above works will be commented in the following sections. The comparative individual presentation of security, dependability, and performance in 5G MEC and the discussion of jointly addressing these three aspects in 5G MEC are new and will help the researchers to better face the challenges of future 5G MEC systems and beyond.

Several research projects focus directly or indirectly on MEC, a good summary can be found in [71] and [102]. This work is part of the 5G-MODaNeI[1] project, which focuses on dependability and security in 5G MEC.

III. SECURITY

The research community is active on the security of 5G MEC. Many works highlight the challenges and try to improve the security in 5G MEC. After a brief introduction of the security taxonomy for the readers that are not experts on the topic, we summarize the current research activity, focusing on security-oriented surveys of MEC, and discuss the security challenges.

[1] https://5g-modanei.ux.uis.no

TABLE 3. ETSI specifications, reports, and white papers.

No. & Ref.	Name	Security	Dependability	Performance
GS MEC 002 [12]	Phase 2: Use Cases and Requirements	✓	✓	✓
GS MEC 003 [7]	Framework and Reference Architecture		✓	
GS MEC-IEG 006 [27]	Market Acceleration; MEC Metrics Best Practice and Guidelines		✓	✓
GS MEC 009 [28]	General principles, patterns and common aspects of MEC Service APIs	✓		
GS MEC 010-1 [29]	Mobile Edge Management; Part 1: System, host and platform management		✓	
GS MEC 010-2 [30]	MEC Management; Part 2: Application lifecycle, rules and requirements management	✓	✓	
GS MEC 011 [9]	Edge Platform Application Enablement	✓	✓	
GS MEC 012 [31]	Radio Network Information API	✓		
GS MEC 013 [32]	Location API	✓		
GS MEC 014 [33]	UE Identity API	✓		
GS MEC 015 [34]	Traffic Management APIs	✓	✓	✓
GS MEC 016 [35]	Device application interface	✓		
GS MEC 021 [36]	Application Mobility Service API		✓	✓
GS MEC 024 [26]	Support for network slicing	✓	✓	✓
GS MEC 026 [37]	Support for regulatory requirements	✓		
GS MEC 029 [38]	Fixed Access Information API	✓		
GS MEC 030 [39]	V2X Information Service API	✓	✓	
GS MEC 031 [19]	MEC 5G Integration	✓	✓	✓
GS MEC-DEC 032-1 [40]	API Conformance Test Specification; Part 1: Test Requirements and Implementation Conformance Statement (ICS)		✓	
GS MEC-DEC 032-2 [41]	API Conformance Test Specification; Part 2: Test Purposes (TP)		✓	
GS MEC 033 [42]	IoT API	✓		
GS MEC 037 [43]	Application Package Format and Descriptor Specification	✓		✓
GS MEC 040 [44]	Federation enablement APIs	✓		
GR MEC 017 [15]	Deployment of Mobile Edge Computing in an NFV environment		✓	✓
GR MEC 018 [45]	End to End Mobility Aspects	✓	✓	✓
GR MEC 022 [46]	Study on MEC Support for V2X Use Cases		✓	✓
GR MEC-DEC 025 [47]	MEC Testing Framework		✓	
GR MEC 027 [16]	Study on MEC support for alternative virtualization technologies	✓		✓
GR MEC 035 [11]	Study on Inter-MEC systems and MEC-Cloud systems coordination	✓	✓	✓
GR MEC 038 [48]	MEC in Park Enterprises deployment scenario			✓
GR MEC 042 [49]	Guidelines on Interoperability testing			✓
White paper No. 20 [8]	Developing Software for Multi-Access Edge Computing	✓	✓	✓
White paper No. 23 [23]	Cloud RAN and MEC: A Perfect Pairing	✓		✓
White paper No. 24 [21]	MEC Deployments in 4G and Evolution Towards 5G	✓	✓	✓
White paper No. 28 [18]	MEC in 5G networks	✓	✓	✓
White paper No. 30 [50]	MEC in an Enterprise Setting: A Solution Outline		✓	✓
White paper No. 32 [51]	Network Transformation; (Orchestration, Network and Service Management Framework)		✓	
White paper No. 34 [52]	Artificial Intelligence and future directions for ETSI		✓	✓
White paper No. 36 [53]	Harmonizing standards for edge computing - A synergized architecture leveraging ETSI ISG MEC and 3GPP specifications		✓	
White paper No. 39 [54]	Enhanced DNS Support towards Distributed MEC Environment		✓	✓
White paper No. 46 [55]	MEC security: Status of standards support and future evolutions	✓		
White paper No. 49 [56]	MEC federation: deployment considerations	✓		

A. TAXONOMY

Security targets some *objectives*, for homogeneity further called *security attributes* (also known as *security requirements* or *security properties*) that are guarded against *adversaries* (or *attackers*). Adversaries are malicious parties that intentionally mount *attacks* with the aim to break one or more security attributes and thus gain illegitimate advantages. Attacks are *security issues* that are possible because of *vulnerabilities* that reside in the system and can be mitigated or defended against by enforcing a variety of *countermeasures*. Figure 7 sketches the taxonomy related to security. For more details, the National Institute of Standards and Technology (NIST) glossary containing terms and definitions related to cybersecurity is available at [103]. Similarly, for a more in-depth description of related notions, the reader may refer to [104].

1) ATTRIBUTES

Table 4 lists the commonly accepted security attributes, as defined in [103], [104], and [105]. *Confidentiality*, *Integrity*, and *Availability*, known as the *CIA Triad*, are fundamental requirements. Confidentiality is related to privacy and, in this context, guarantees the *privacy of personal data*. Privacy includes other aspects too, such as the *privacy of identity* (in relation to *anonymity* and *unlinkability*) or the *privacy of location*. Data integrity can be perceived as *data authentication* because it guarantees that the data has not been altered in any way, so it is authentic. For the purpose of this paper, both integrity (or authentication of data) and authentication of entities are valuable requirements. Moreover, *mutual authentication* is a strong form of entity authentication that requests that all the communicating parties authenticate to each other. Other traditional and generally

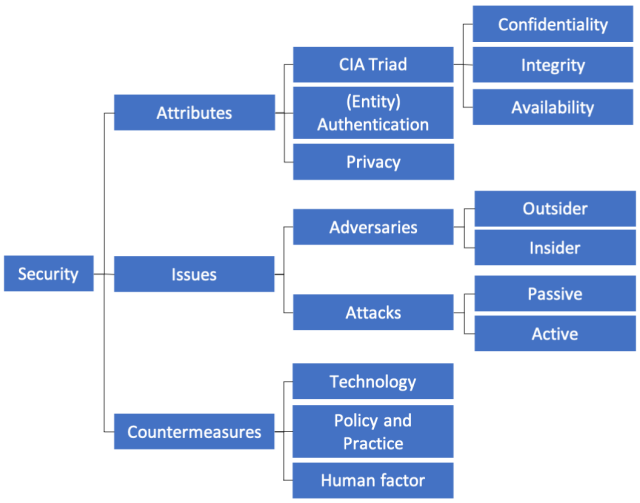

FIGURE 7. Security taxonomy.

TABLE 4. Security attributes [103], [104], [105].

Attribute	Description
Confidentiality (C)	Keep the information secret, except from the authorized parties.
Integrity (I)	Ensure that the information has not been altered in an undetected manner by unauthorized parties.
Availability (A)	Assure that the legitimate parties can access a service (or a resource) when they need to.
Authentication (Au)	Attest the identity of a party (*entity authentication*) or the source of information (*data origin authentication*).
Privacy (P)	Control how, when, and to what extent personal information is communicated to other parties. Incorporates different aspects, such as identity (e.g., name, pseudonym), related personal data (that might also become an identity in some context, e.g., phone number), location, user-created data, etc.

accepted requirements exist, such as *non-repudiation*, which prevents the denial of previous actions. However, we leave this outside of the very succinct taxonomy presented, as it is not of particular interest to our work. For more discussion on attributes, refer to [104] and [105]. These requirements are enforced by security functions, such as e.g., *access control* that protects access to resources against unauthorized parties or *authorization* that officially grants to a party the right to be or to do something. Based on previous work [106], [107], Khan et al. accept *visibility* and *centralized policy* as two additional security parameters [102]. We omit them here for several reasons, one being that they are not directly related to the security and privacy of the end beneficiaries, but they can be seen more as functions that help to achieve these.

2) ISSUES

In essence, security issues are caused by *adversaries* that mounts *attacks*. The attack surface is wide, and many classifications of adversaries and attacks exist in the literature, based on different criteria. Table 5 defines the most common types of adversaries and attacks that will be referred to in the paper.

Concerning the position of the adversaries with respect to the target, they can be classified into *outsiders* and *insiders*. An internal adversary has a clear advantage over an external adversary (e.g., has physical access to the network or resources, is authorized to access some data, resources, or services). This makes internal adversaries more powerful than outsiders and insider attacks are usually easier to mount than outsider attacks.

Further, attacks can be classified into *passive* and *active* [104], [105]. Passive attacks are simpler to mount, cheaper, and more efficient in resource consumption but less powerful than active attacks. They are more difficult to detect because the adversary does not actively interfere and thus changes nothing. Passive attacks include *eavesdropping* and *traffic analysis*. They directly threaten confidentiality but can also precede active attacks by gathering the necessary information to mount these. On the other hand, active attacks can directly target any security property by modifying, deleting, and injecting messages (or in general, data in any form: storage, computation, or transmission) or altering in any way the functionality of the target. Besides confidentiality, active attacks can also directly damage integrity and authenticity (e.g., *replay attacks*, *Man-in-the Middle Attacks*) or availability (e.g., *Denial of Service - DoS*, *Distributed DoS - DDoS*).

With respect to the implications for the target, active attacks can be roughly classified into three types: *deteriorate* - partially or fully damage the functionality of a target, *corrupt* - take control over the functionality of a target by either leaking sensitive information or behaving in a specific, desired way, or *impersonate* - fake an entity to gain an advantage of its functionality in the system. The attacks that deteriorate the good functionality usually aim to damage the availability by consuming resources in excess, so they are strongly related to performance and dependability too. Of course, disruption can also be caused by corrupting a target (by either resetting, stopping, or change its normal functionality). Moreover, an adversary can mount complex attacks that are combinations of several attacks (of different types), and the adversary can be either a single entity or a coalition of entities that collide to mount the attack together.

Of course, both adversaries and attacks can be referred to in both classifications, so we can very well refer to *active/passive adversaries* and *insider/outsider attacks* too (e.g., an active adversary mounts an active attack, an insider adversary mounts an insider attack). The reader that is unfamiliar with the attacks' exemplification can further refer to [103] and [104].

3) COUNTERMEASURES

The countermeasures, also known as *safeguards* [103], can be *reactive* or *proactive* and come in a variety of means. As traditionally categorized in the McCumber cube [108],

TABLE 5. Main types of adversaries and attacks.

	Type	Main Target	Description	Examples
Adversary	Outsider	(all)	The adversary is an outsider, unrelated to the target.	An adversary outside the organization, a malicious competitor
	Insider	(all)	The adversary is an insider, somehow related to the target (e.g., the adversary has some insight knowledge, and can access some services/resources).	A malicious employee, a compromised MEC app
Attack	Passive	(C, P)	The adversary is passive, he/she can listen but not actively interfere with the system.	Eavesdropping, data capture, (offline) (crypt)analysis on collected data.
	Active	(all)	The adversary actively interferes by modifying, deleting, injecting messages (or in general, data in any form: storage, computation, or transmission), or altering in any way the input/output or functioning of the system.	
	- Deteriorate	(A)	The active adversary damages the functionality of the victim in a black-box fashion, either partially or fully.	Service deprecation, Denial-of-Service (DoS), partial or total physical damage
	- Corrupt	(all)	The active adversary compromises the victim, controlling it either partially or fully.	Malware, credential theft, active monitoring
	- Impersonate	(all)	The adversary impersonates an entity to gain an advantage.	Man-in-the-Middle (MitM), replay attacks, rogue/fake entities

the safeguards can be enforced in *technology*, *policies and practices*, and the *human factor*. *Technology* can be implemented in either software or hardware, and can consist of cryptographic primitives and protocols (e.g., encryption to protect confidentiality), firewalls, Intrusion Detection Systems (IDS), isolation techniques, and many others. *Policies and practices* consist of rules, regulations, and best practices that specify the expected behavior of involved entities. Examples include authorization policies, incident response procedures, and recovery procedures. Finally, the *human factor* includes education, training, and awareness of users to obey the security policies and make use of technology mechanisms in correspondence to the security goals, make people aware of possible consequences, and become responsible for their acts [108].

B. STATE OF THE ART

1) STANDARDS, REGULATIONS, AND WHITE PAPERS

The ETSI specifications that refer to MEC security aspects (listed in Table 3) are used as references for presenting the state-of-the-art security requirements and regulations for MEC, presented in Table 6. These include general requirements [12], API-related aspects [9], [28], [31], [32], [33], [34], [35], [38], [39], [42], [44], network slicing [38], MEC integration within 5G [19], end-to-end mobility aspects [45]. Note that N/A here does not mean that security requirements are out of scope or importance for the given element; N/A means that no explicit requirements are listed in the specifications.

White papers by ETSI (also listed in Table 3) but even other organizations and industry companies [109], [110] refer to the security aspects of MEC within 5G too.

2) ACADEMIC PUBLICATIONS

Table 7 lists recent surveys on security and privacy for MEC. Papers that refer to general MEC security aspects are also considered. The number of publications considering MEC security and privacy is large, so the table does not intend to be exhaustive. Not many papers are fully dedicated to MEC but consider it together with other technologies such as cloud or fog (sometimes only marginally such as in e.g., [75]). As mentioned in some of these (e.g., [70], [102]), some specialized work on the security for MEC has been performed. However, the number of papers referring to particular aspects regarding general edge technologies (and, to some extent, applicable to MEC too) or general security issues that are not MEC specific is large and thus out of the goal of this paper. For example, a large number of papers are dedicated to defenses against (D)DoS in MEC (e.g., [123], [124], [125], [126]) or usage of ML for MEC (e.g., [127], [128], [129]). Security in 5G network slicing has been analyzed in [25], and some aspects are relevant for isolation in MEC too. Nevertheless, papers that survey aspects of MEC privacy and security do exist, e.g., [122], a comprehensive study performed in parallel and independent of our work, or [121] a MEC security analysis for each of the twelve considered vertical industries: (1) manufacturing industry, (2) financial sector, (3) healthcare, (4) education, (5) telecommunication, (6) authorities, (7) media and entertainment, (8) smart city, (9) agriculture and food industry, (10) logistics, (11) education, culture, and science, and critical infrastructure sectors.

C. CHALLENGES

In the following, we present security challenges for MEC by categorizing them at the MEC host level, MEC system level, and general challenges.

1) MEC HOST LEVEL

a: PHYSICAL SECURITY

The MEHs are located at the edge of the network, closer to the user and in open environments. The physical location of the MEHs becomes thus insecure [96], [102], host-level

TABLE 6. Security requirements in the MEC architecture according to standardization documentation.

Element	Security Requirements
General	- The MEC system shall provide a secure environment for running services for the involved actors, in particular the user, the network operator, the operators' third parties, and the platform vendor [12]. - The MEC system shall prevent illegal access from dishonest terminals, secured communication is necessary between the radio nodes and the MEC service [12]. - The MEC system should securely collect and store logs, including information about charging [12]. - MEC services might require end-to-end security mechanisms [12]. - APIs must be secured, including general aspects such as controlling the frequency of the API calls, unexposure of sensitive information via the API, the authentication of a client to RESTful MEC Service API is performed using OAuth 2.0, and APIs shall support HTTP over TLS 1.2 [28], [31], [33], [35]. - The MEC system discovery, including security (authentication/authorization, system topology hiding/encryption), charging, identity management, and monitoring aspects, is an essential prerequisite to form a MEC federation [11].
MEH	
MEP	- The MEP shall only provide a MEC app with the information the MEC app is authorized for [12]. - The MEP shall provide a secure environment for providing and consuming MEC services when necessary [12]. - A MEP should securely communicate to a MEP that might belong to different MEC systems [7]. - If the MEP is not dedicated to a single NSI, then the MEP shall support a different set of services and functionalities in distinct NSIs; in particular, the MEP needs to share an infrastructure that allows authentication and authorization at the NSI level and assure isolation of data and services between NSIs [26]. - The MEP discovery is provided by means of the MEC systems exchanging information about their MEPs, i.e., their identities, a list of their shared services, as well as authorization and access policies [11].
Virtualization Infrastructure	- Virtualization should not introduce any new security threat; in particular, hypervisors should not introduce new vulnerabilities, patch management, state-of-the-art protections and secure boot mechanisms should be in place [13]. - A proper isolation of VNFs [13] and VMs [16] should be realized, the interfaces between the NFV components should be secured [13].
MEC App	- The application instance must satisfy the necessary security constraints [12]. - The MEC applications should have access to a persistent storage space [12]. - The MEC application should only have access to information for which it is authorized, should manage the access control and integrity of the user content, and should be authorized to consume or provide MEC services [12]. - Upon authentication, the MEC applications should be able to communicate securely regardless if they are placed in the same MEH or different MEHs [12], even located in different MEC systems (in case of multi-operator scenario) [7], [12]. - Operator trusted MEC application (seen as an extension of the MEP functionality) have advanced privileges of secure communication with the MEP [12]. - The security should be added to the MEC app package following the requirements defined for NFV [111], where the VNF refers to MEC app [43].
MEC Host-level Management	
MEPM	- The MEC management should verify the authenticity and integrity of a MEC application [12], [30].
VIM	- N/A (see Virtualization Infrastructure)
MEC System-level Management	
MEO	- The MEO checks the integrity and authenticity of the application packages [7], [30]. - The MEO authorizes the requests of the OSS (e.g., application instantiation and termination, fetch on-boarded application package, query application package information) [30]. - MEAO adapts the orchestration operations based on the available NSIs and their requirements, including security requirements too [26].
OSS	- N/A
User app LCM proxy	- N/A
MEC Federation	
MEF	- The MEF should enable to exchange information in a secure manner among MEPs and MEC applications that belong to different MEC systems [7], [23]. - The MEF should face the security threats given by the heterogeneous scenario and edge resource sharing among operators (together with edge computing service providers and partners) [55], [56]. These threats can be related to the access network, the architecture, the core network, the MEC elements, or other [55]. - For security reasons, the information of MEP should be hidden between federated MEC systems [44].

TABLE 7. Academic surveys related to MEC security.

Ref.	Aspect	MEC only	Main contribution	Relevance to MEC security
[112]	5G security	No	Gives an overview of 5G security challenges and solutions, referring to other technologies (e.g., SDN, NFV, cloud).	Mentions briefly some MEC security challenges, considered together with mobile clouds.
[70]	Edge general	No	Analyses the security threats, challenges, and mechanisms in all edge paradigms.	Discusses MEC from the perspectives of security, dependability, and performance but keeps the discussion decoupled from the architecture.
[94]	MEC/IoT security	Yes	Considers MEC privacy issues in the context of heterogeneous IoT.	Focuses on big data privacy issues in MEC with respect to data aggregation and data mining, and considers ML privacy-preserving techniques.
[71]	MEC/IoT security	Yes	Overviews the MEC technology for the realization of IoT applications.	Reviews papers and discusses problems, challenges, and possible solutions for IoT security, privacy, and trust in relation to MEC.
[73]	MEC general	Yes	Presents a survey on different aspects related to MEC (e.g., computation, storage, energy efficiency, research infrastructure), also including security and privacy.	Presents some security mechanisms and privacy issues related to MEC keeping the discussion quite general.
[67]	Edge security and resilience	No	Reviews some security and resilience aspects in MEC and fog in relation to cloud technologies.	Generally discusses security requirements and challenges in MEC together with fog and in comparison with the cloud.
[68]	Edge general	No	Presents an overview of potential, trends, and challenges of edge computing.	Refers only briefly to security and privacy.
[113]	Edge security	No	Surveys the data security and privacy-preserving in edge computing.	Analysis the data security and privacy challenges and countermeasures.
[114]	Edge security	No	Discusses network threats in fog and edge computing.	Indicates threats, challenges and trends with respect to network security and awareness, with no focus on MEC particularities.
[93]	Edge/IoT	No	Surveys how edge computing improves the performance of IoT networks, but also considers security issues in edge computing.	Analyzes the security attributes and proposes a framework for security evaluation for edge computing-based IoT.
[96]	Edge dependability	No	Explores dependability and deployment challenges in edge computing. Considers dependability in a wider meaning that includes security.	Presents new challenges in physical security and scalable authentication, considering both centralized and decentralized security mechanisms.
[102]	5G security	No	Gives a security and privacy perspective on 5G, looking into several enabling technologies (e.g., cloud, fog, MEC, SDN, NFV, slicing).	Discusses MEC security and presents some MEC threats and recommendations, but without a special focus on MEC and mostly together with cloud-related security issues.
[76]	MEC security	Yes	Discusses MEC from a security perspective.	Identifies seven security threat vectors and discusses possible solutions and approaches to secure MEC by design.
[74]	Edge general	No	Surveys edge computing, identifies requirements, and discusses open challenges.	Gives low importance to MEC security, which is very briefly discussed.
[95]	Edge/IoT security	No	Discusses security and privacy together with efficiency in data communication and processing computation for IoT at the edge.	Discusses general security threats, secure data aggregation, and secure data deduplication at the edge, with no focus on MEC, but basing the findings on fog nodes.
[115]	Edge security	No	Reviews attacks and the corresponding defense mechanisms in edge computing systems.	Focuses on four types of attack (DDoS, side-channel attacks, malware injection attacks, and authentication and authorization attacks) and presents the related root causes, status quo, and challenges.
[116]	Edge security	No	Introduces the main technologies supporting the edge paradigm and discusses issues and solutions.	Focuses on virtualization technologies (e.g., containers and unikernels) and related security issues.
[117]	MEC/IoT security	Yes	Discusses data analytics in edge computing in the context of IoT.	Reviews the existing works on data analytics in edge computing highlighting pros and cons, proposes some requirements for secure IoT data analytics, and highlights open issues and future research directions.
[78]	MEC general	Yes	Surveys MEC fundamentals, discussing its integration within the 5G network and relation with similar UAV communication, IoT, machine learning, and others.	Discusses MEC security very briefly but it points to many related papers (e.g., in the topic of MEC for IoT, security and privacy in the context of MEC-enabled IoT in V2X, smart cities, and healthcare).
[98]	Edge security	No	Aims to provide a systematic review of security and privacy requirements in edge computing, including a taxonomy of attacks and performance evaluation.	Surveys a significant number of papers, but suffers from some clear shortcomings with respect to content and presentation.
[118]	MEC/IoT security	Yes	Surveys existing IoT security solutions at the edge.	Presents an edge-centric IoT architecture and reviews edge-based IoT research in terms of security and privacy.
[119]	Edge/IoT security	No	Surveys the security and privacy issues in the context of edge-computing-assisted IoT.	Defines security and privacy in the context of the edge-computing-assisted IoT, gives some classifications of attacks, and discusses countermeasures.
[82]	MEC security	Yes	Analyses the security and privacy of MEC.	Discusses threat vectors, vulnerabilities that lead to the identified threat vectors, and proposes solutions to overcome these. The paper can be perceived as an extension of [76].
[120]	Edge/Fog security	No	Surveys security challenges, issues, and countermeasures in edge and fog computing.	Discusses security and privacy aspects in fog and edge computing.
[121]	MEC security	Yes	Surveys security aspects in relation to twelve representative vertical industries of 5G MEC.	Presents characteristics, threats and vulnerabilities, attacks and countermeasures for each identified vertical, and further correlates the impact to the required performance of the vertical.
[122]	MEC security	Yes	Reviews the MEC architecture in relation to security and privacy aspects.	Reviews the conceptual guidelines for MEC security architecture as well as security and privacy techniques, examines and categorizes significant threats, and considers possible safeguards.

devices being even more vulnerable to physical attacks than system-level equipment that is normally placed in a more physically secured area. Moreover, the tendency to deploy many MEHs to cover an area raises problems with respect to a good physical security level [76]. This increases the risk of unauthorized physical access and hence physical

deterioration or corruption of the devices, with direct consequences against the availability (e.g., DoS attacks) and the confidentiality (e.g., data leakage by both passive and active attacks) [109]. The equipment might lack the hardware protection of commodity servers [70], but as a form of protection, the MEC devices should implement anti-theft and anti-damage measures [109]. Surely, tamper resistance is a generally good security strategy to prevent the reading of confidential data (e.g., cryptographic keys) and thus should be adopted in the case of MEC equipment too [71], [96]. In this scenario, the well-known principle of the weakest link holds: the security of the overall system is given by the security of the weakest spot (the adversaries tend to attack weak spots). The MEHs with poor security can easily become points of attacks [94].

b: LOCATION PRIVACY

Location tracking enabled by MEC can be seen as both a feature and a risk. Unauthorized access to the Location Application Programming Interface (API) for the MEC Location Service can leak sensitive information about the localization and tracking of users in time [32], similar to unauthorized access to the Radio Network Information in mobile networks (e.g., access to identifications, which might damage the privacy of the UE) [31]. MEHs (and, in consequence, end-users too) are thus directly exposed to location privacy risks [95]. To mitigate such risks, an important role is in the security of the APIs and the generated location reports or processed data. On the other hand, the MEC localization service can be beneficial when GPS coverage is unavailable or for emergencies (e.g., for healthcare applications where monitoring devices can send signals requiring assistance to the closest MEC platform in case of emergency) [130].

c: LOCAL DEFENCES

Due to their local character, attacks at the host level influence a geographically limited area, in the proximity of end users [70]. This gives MEC capabilities to enforce security mechanisms and limit attacks in the local network segment [71]. MEC is suitable to deploy a defense perimeter, one example being against (D)DoS attacks when the adversary only targets a smaller volume of traffic, and the edge can alert the core network about the source of danger, resulting in overall better availability [126]. The local character of MEC can also enhance privacy protection by preventing data to arrive at centralized servers, thus avoiding a centralized point of trust. An example from [71] consists of the processing of images with car plates at the edge and identification of the plate number only to forward to central processing (this prevents, for example, location leakage). At the same time, it is believed that local data exchange (as contrary to e.g., sending the data over the internet) reduces the exposure of data [95], but of course can raise security risks when the number of nodes is high, because of high traffic and positioning at the edge of the network [78]. Reference [131] discusses some security improvements that MEC can bring in the IoT scenario.

d: VIRTUALIZATION SECURITY

Malicious virtual machines can try to exploit their hosts [70]. Attacks such as VMs manipulation might include a malicious insider with enough privileges to access and damage a VM or a malicious VM with escalated privileges [70], [71], [132]. If a VM is running on multiple hosts, then a simple DoS attack can damage all hosts simultaneously [132]. As a protection against the DoS attacks, the VMs should be limited in resource consumption, and the resource consumption should be balanced among hosts. With respect to the privacy of data, the user data is stored at MEC host level, so it might get leaked [76]. Moreover, possible alteration of data requires adequate backup and recovery possibilities, in strong relation with dependability prevention. Virtualization attacks can damage the orchestration at the host level, and a compromised VIM can lead to the disruption of MEC services [76]. Service manipulation is another example of corruption, with important consequences such as DoS or data leakage attacks [70]. If a host is corrupted (not necessarily by virtualization attacks but in general), the adversary might interfere at several levels (e.g., apps, services, resource consumption) and run a wide set of possible attacks.

e: CONSTRAINED RESOURCES

Computational expensive security mechanisms, including the usage of heavy cryptography, can be a problem. For example, the edge devices might have limited connectivity and resources, which impose restrictions on the security protocols that can be deployed and facilitate attacks against availability [110]. This might conclude in restrictions in the deployment of highly secure protocols, for example for authentication [70]. Usage of public-key cryptography and Public-Key Infrastructure (PKI) in particular might be an issue because of high computational costs and management [96]. Lightweight cryptography should be considered. Data deduplication mechanisms at the edge (in the sense of detecting and discarding copies of data, or even preventing re-computations) would increase the performance on limited devices. But realizing this while maintaining security is normally possible via (Fully) Homomorphic Encryption ((F)HE), which by itself requires very high computational costs [95]. The European Agency for Cybersecurity (ENISA) itself identifies the complexity of the implementation of security solutions in 5G (caused by mixing technologies such as cloud, fog, and edge), as well as the efficient cryptography solutions (because of nodes constrained in resources) to be key topics in research and innovation of 5G security [133].

2) MEC SYSTEM LEVEL

a: GLOBAL DEFENCES

The management and orchestration of the diverse security mechanisms is a complex issue, and enabling security

mechanisms independently on multiple entities does not necessarily mean that the complete system is secured [70], [78]. A balance between local (decentralized) and global (centralized) defense mechanisms, between responsibility and autonomy has to be considered [70], [96]. A global monitoring that allows an overview of the MEC system should be set in place, and everything should be auditable [70]. To provide privacy, end-to-end encryption becomes a necessity whenever applicable [102].

b: MEO SECURITY

MEO is exposed to virtualization attacks [76]. A compromised MEO could have a critical impact on the functionality of the overall MEC system. Examples include the termination of MEC critical applications, on-boarding of malicious application packages, an unbalanced usage of the MEHs in terms of resources, and others. Hypervisor introspection methods need to be applied, for Linux-based platforms, Security Enhanced Linux (SELinux) might be of use [76]. More on virtualization defenses will be discussed in III-C3, the subsection dedicated to general challenges. However, in the rapidity of change in technology and attacks nowadays, it is considered suitable to approach by software programmable solutions than rigid hardware (to allow updates, moving targets, dynamic attacks, etc.) [126]. Security and privacy challenges in softwarization and virtualization are not specific to MEC, as flexible solutions are required to secure 5G in general [102]. Solutions to prevent virtualization problems and security frameworks within the MEC-in-NFV architecture have been considered [76], [134]. These include Trusted Platform Manager (TPM) to attest and validate MEC Apps and VNFs, as well as requests from the CFS portal [76]. Auto-configurable security mechanisms are proposed, as well as methods to secure VNF in NFV environments [76], [135]. Ideally, the idea of a security orchestrator based on softwarization (SDN / VNF) would replace the need for manual configuration that is no longer feasible under the current circumstances [102]. But how such a security orchestrator should be built and integrated with the architecture of MEC is still an open question.

c: INTERCONNECTION SECURITY

At the MEC system level, OSS is exposed to threats outside the MEC system through the communication with the CSF Portal and device applications via the LCM proxy. This opens up for security risks, for example, the CSF Portal is prone to (D)DoS attacks [76]. OSS can be subject to masquerading for adversaries that pretend to have legitimate access [76]. A significant number of requests from the OSS to the MEO might (in the absence of proper security mechanisms) damage the functionality of the MEO.

3) GENERAL CHALLENGES

a: TRUST MODELS AND RELATIONS

Contrary to other technologies at the edge (e.g., edge cloud, fog), in MEC there is a lower number of owners that need to cooperate [70]. However, it is of critical importance to clarify the trust models and relations between the entities involved (users, platforms, slices, apps, etc.) [109]. In particular, trust needs to be considered in relation to mobility and network functions performed in the MEC and integration to the 5G standard too [109]. Trust models have been considered for different edge technologies [70], [136]. A flexible trust manager has been proposed as a solution to incorporate within MEC [71]. Other trust schemes have been proposed for MEC (e.g., [137]). Nevertheless, general protection mechanisms such as mutual authentication and access control mechanisms at all levels should be set in place. A good prevention is to minimize the data transmitted and stored on low reputation entities [70]. The ownership of personal data must be assigned and clearly decided between different roles (e.g., stakeholders, MNO, third parties) [102]. To protect end users' data, techniques such as watermarking, visual cryptography, and biometrics were considered in the literature [102].

Nevertheless, security models (not only trust models) in accordance with the MEC requirements need to be developed and applied for security analysis.

b: NETWORK SECURITY

The heterogeneous nature of edge and its dynamic character introduces risks [74], [78], [102], [110]. Security is prone to risks at the interconnection with other technologies, mostly within the 5G context. In particular, MEC should access the internet and establish connectivity to other MEC domains via the internet [76]. Naturally, security should be considered for data in all forms (storage, computation, and transmission) and approached by appropriate cryptographical primitives and security technologies (including general approaches such as VPN communication, access control functions and policies, firewalls [76]). Lightweight encryption can be used for performance enhancements whenever convenient [76]. However, there is still a place for improvement with respect to lightweight cryptography needed for solutions such as MEC. MEC is exposed to attacks on the communication channels, mostly on the wireless channels close to the end-user [76]. A specific risk lies in the communication between the edge and the core [102], [110]. Data has to be encrypted in transit (e.g., IPSec, TLS) [109], and end-to-end encryption is a way to prevent data leakage [95]. Adaptive security protocols could be used to enhance the security of communication channels [76]. Software-Defined Virtual Private Local Area Networks (Soft-VPLS) can help in securing the communication between MEC components [76]. Soft-VPLS allows different traffic categories (e.g., MEC service requests, user data, control statistics) to be routed via distinct tunnels, aiming to enhance both end-to-end security and overall communication performance [76]. Proper isolation of network traffic, data, services, slices, etc. is required [70], [109]. The use of gateways at strategic points in the network is considered a good practice [112], and firewalls are now

implemented as NFs [130]. Techniques to provide physical layer security might turn out beneficial because of performance aspects [78], [138], [139]. More on communication security is discussed in [102].

c: MONITORING AND DETECTION

Mitigation techniques in network security include monitoring and logging, abnormal traffic analysis, malware and intrusion detection. In MEC, this should be performed at all levels. The collaboration between the edge nodes can be useful in this respect [109]. Some work on access control and Intrusion Detection Systems (IDS) for MEC has been pointed out before [70], [140]. This can prevent or mitigate attacks such as DoS, malicious actions, and rogue entities [70]. AI, in particular ML, could be successfully applied for intrusion and anomaly detection [76], [130]. Deep Learning (DL) was specifically considered for the detection of attacks, and Reinforcement Learning (RL) was proposed for edge caching security [78], [127], [128]. Federated Learning (FL) suits MEC because it allows the training data to be kept locally and privately among collaborating nodes. Thus, advances in using FL for constraint devices might be useful [78], [141], [142]. More references about ML in IDS and against DDoS attacks can be found in [102], also in correlation with SDN solutions. The study of ML (and AI in general) for MEC security is a valid research direction [71].

d: VIRTUALIZATION SECURITY

General virtualization and softwarization issues need to be considered in MEC too. This includes security issues of both SDN and NVF [70], and in particular security aspects related to network slicing [25], [102]. SDN/NFV-based frameworks or approaches to provide security in relation to MEC and IoT have been considered [71], [134]. Software-Defined Privacy (SDP), a solution currently in place for enforcing the security of Internet as a Service(IaaS) cloud customers, might be extended to provide privacy protection in MEC too [71]. Virtual Machine Introspection (VMI) and hypervisor machine introspection should monitor the activities in terms of resource utilization to prevent deprivation and DoS attacks [76]. These should be run at both host and system level [76]. Examples of VMI include LibVMI [76], [143]. Artificial intelligence can be used to achieve better VMI solutions [76], [144].

e: STANDARDIZATION AND AWARENESS

The necessity for standardized or universal security measures to ensure the security of the overall MEC system is considered to be an open problem [71]. At the same time, the human factor remains a risk, so it is important to raise awareness of the users that need to understand and apply security policies [70]. A significant category is given by the application developers [71], and special focus should be put on the integration mechanisms.

f: OTHER SECURITY CHALLENGES

Many other security challenges exist in MEC (e.g., see [102] for some specific Backhaul threats). We next refer briefly to some of these.

Privacy of identity is known to be a challenge in mobile networks (including 5G), and it remains a challenge in MEC too [102], [145]. More on Personally Identifier Information (PII) protection in the general context of *General Data Protection Regulation (GDPR)* can be found in [146].

Usage of the appropriate technologies and primitives for security is always a challenge. For example, solutions nowadays tend to adopt blockchain-based technology, including in MEC [147], [148]. However, many times the blockchain technology is in fact not needed [149] and more suitable solutions (e.g., which introduce less complexity) are available. While ETSI has a dedicated group on Permissioned Distributed Ledgers (PDL) [150], it remains open if blockchain is indeed a useful technology for MEC. Quantum security mechanisms have to be considered [102]. Cryptographical primitives currently used in 5G and MEC (mainly public-key primitives) are known to be vulnerable to quantum attacks. Although nowadays quantum attacks are still in their infancy, quantum-resistant cryptography is an active research field that aims to provide security against quantum adversaries.

Mobility-related security challenges are of importance in MEC too. More about these will be discussed together with other aspects, in Section VI.

IV. DEPENDABILITY

If security is a well-known term, dependability is a less common term, whose meaning is sometimes ignored or confused. In this paper, we define dependability as in [151]. Dependability is the *ability to deliver a service that can justifiably be trusted*.

A. TAXONOMY

Figure 8 represents the taxonomy related to dependability. The *attributes* are the various ways to evaluate the dependability of a system. The *issues* are the causes that may lead to a lack of dependability. The *countermeasures* are the methods to enhance the dependability of a system.

Alternative definitions and taxonomies can be found in literature [152], [153], where the term *threat* is used instead of *issue* and the term *mean* is used instead of *countermeasure*. Moreover, some works define a joint dependability and security taxonomy [151], [153]. Instead, performance, or performability, is sometimes seen as one of the attributes of dependability [153]. For the sake of clarity, we keep separate security, dependability, and performance, and we will jointly discuss them in Section VI.

1) ATTRIBUTES

The attributes consist in metrics that are able to characterize and measure specific properties of a system. The attributes can be applied to the MEC as a system, but also to MEC-based

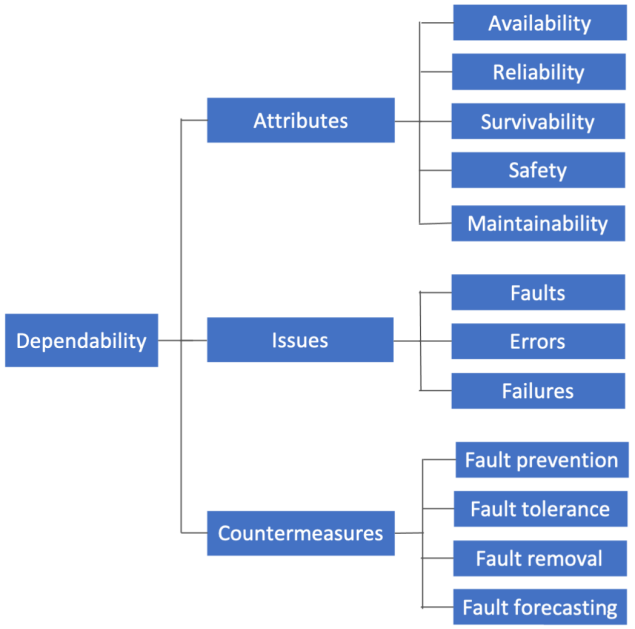

FIGURE 8. Dependability taxonomy.

TABLE 8. Dependability attributes.

Attribute	Description
Availability (A)	Readiness for a correct service.
Reliability (R)	Continuity of correct service.
Survivability (S)	Capability to continue to deliver a correct service in the presence of failures or accidents.
Safety (Sf)	Absence of catastrophic consequences on the user(s) and the environment.
Maintainability (M)	Ability to undergo modifications and repairs.

services. The main attributes are listed in Table 8. The description of the attributes is based on previous definitions [151], [153]. The most known attributes are *availability* and *reliability*. A system is available when it is ready to deliver a service that complies with the service specifications. The simplest way to compute the availability of a system is the ratio between the expected uptime of the system and the aggregate time (sum of expected values of up and down times). A system is reliable when it is able to continuously deliver a service that complies with the service specifications. The reliability can be computed based on the mean time to failure or the mean time between failures. The *survivability* and *safety* are sometimes not listed among the attributes. The *maintainability* is associated with recovery mechanisms and can be measured as the mean time to repair.

2) ISSUES

The dependability issues, fault and failure as listed in Figure 8, in the daily language are used interchangeably and mean that something that is not working. In dependability taxonomy, they are not only independent causes of lack of dependability, but they represent a cause-effect sequence and they can be defined as follows [151]:

TABLE 9. Main types of faults and failures.

Type	Description
Faults	
Development fault	Fault occurring during development.
Physical fault	Fault that affects hardware.
Interaction fault	External fault.
Failures	
Content failure	Deviation of the content of the information delivered.
Timing failure	Deviation of the arrival time or duration of the information delivered.
Erratic failure	The service is delivered (not halted) but is erratic.
Inconsistent failure	Some or all system users perceive differently incorrect service.
Catastrophic failure	The cost of harmful consequences is orders of magnitude, or even incommensurately, higher than the benefit provided by correct service delivery.

- *Fault* is the adjudged or hypothesized cause of an error;
- *Error* is part of the system state that is liable to lead to a failure;
- *Failure* is the deviation of the delivered service from the correct service (i.e., the service is no longer compliant with the specification).

An example of the cause-effect sequence is the following: an external fault flips a bit in the memory causing an error that manifests as a failure when that partition of the memory is accessed. Table 9 lists the main types of faults and failures [151].

3) COUNTERMEASURES

In order to improve the dependability of a system, various categories of methods are used [151]:

- *Fault prevention* uses development methodologies, for both software and hardware, in order to reduce the number of faults;
- *Fault tolerance* is carried out via error detection and system recovery in order to avoid failures;
- *Fault removal* can be performed during the system development (via verification, validation, and testing) and during the system use (via corrective or preventive maintenance);
- *Fault forecasting* is conducted by evaluating the system behavior (it can be both qualitative and quantitative).

B. STATE OF THE ART

1) STANDARDS, REGULATIONS, AND WHITE PAPERS

The ETSI specifications that refer to MEC dependability aspects (listed in Table 3) are used as references for presenting the state-of-the-art dependability requirements and regulations for MEC, listed in Table 10. Note that N/A here (as in Table 6 for the security) does not mean that dependability requirements are out of scope or importance for the given element; N/A means that no explicit requirements are listed in the specifications.

Availability and reliability are mentioned as non-functional metrics in [27]. Resiliency and high availability is mentioned in the white paper [8].

In many ETSI specifications, *service availability* is mentioned but not always with the same meaning in the dependability: availability tracking API in [30]; together with application availability in [9]; testing the related query in [47]; testing API query in [40] and [41]; together with 5G [19]; in connection with the coordination of inter-MEC systems and MEC-cloud systems [11].

The *service continuity* in mobility is extensively addressed in [45] and [46]. It is also mentioned in [11], [35], and [41] and in several white papers [18], [21], [50], [53].

The *fault management* is briefly addressed in [7] and more in detail in [29]. Fault management in NFV implementation is addressed in [15]. Testing of MEH fault management is presented in [47]. API test for MEH fault management is discussed in [40]. Fault management is also addressed in the white papers [51] and [52].

The *network dependability* is addressed in [34] and also in the white papers [10], [50].

Several specifications refer to *URLLC and mission-critical application*, for example in [26] and [45] and the white papers [18], [21]. The specifications highlight the importance of MEC in mission-critical low-latency applications, such as Industrial IoT and Self-Driving Cars. These applications require communication with very high reliability and availability, as well as very low end-to-end latency going down to a millisecond level.

V2X communication is maybe the most relevant use case in MEC. It is initially mentioned in [12], then more in detail in [46]. V2X communication is important because it has strict requirements in all three aspects. In particular, the ETSI specifications mention the reliability and availability (from both security and dependability perspectives) together with latency and throughput. Given the high mobility of the V2X users, one of the main topics to investigate is the handover between the MEHs, which has an impact on both service continuity and service availability [12]. Specifically, the predictive handover is investigated to meet the dependability and performance requirements [46]. In [39], service continuity is mentioned in the context of having multiple operators. The white paper [154] highlights the high reliability and security requirements in the V2X communication in 5G-MEC.

2) ACADEMIC PUBLICATIONS

Table 11 lists recent surveys that address dependability aspects in MEC. As for security, not many papers are fully dedicated to MEC but consider it together with other technologies such as cloud or fog. A survey on dependability on MEC does not exist. A few papers are mainly focused on dependability aspects [86], [89], [96] and others have parts dedicated to dependability aspects [74], [75], [77] or shared with other aspects [64], [68], [88], [97]. The other papers just mention it or use it as a requirement [64], [77], property [66], benefit [79], or challenge [74], [78].

C. CHALLENGES

In the following, we present dependability challenges for MEC by categorizing them on the MEC host level, MEC system level, and general challenges.

1) MEC HOST LEVEL

a: PHYSICAL DEPENDABILITY

The MEH is practically consisting of a computer or a small server. For this reason, traditional techniques for making dependable a computer or server can be considered. These techniques would act on both hardware and software.

Given the distributed nature of the MEC system, which is composed of multiple MEHs, the alternative is to consider a MEH expendable and provide fault tolerance by migrating to a new MEH in case of failure [86]. In this case, careful deployment of MEHs and efficient failover mechanisms are needed. Deployment and failover mechanisms will be presented more in detail later.

b: VIRTUALIZATION DEPENDABILITY

An important characteristic of a MEH is virtualization. A MEH uses virtualization technologies, such as containers or VMs, which are managed by a VIM, such as Kubernetes or Openstack [16]. Moreover, as already mentioned, the MEC architecture can be integrated with a virtualization architecture such as NFV [15].

The virtualization in the MEH needs to be considered in order to have a dependable MEC. For example, the live VM migration in Openstack while injecting network failures and increasing the system pressure can be investigated [157]. Note that the VM can be migrated to a different host or within the same host. Regarding NFV, a problem that can be addressed is the VNF placement in a MEC-NFV environment in order to maximize the availability [158]. The most critical part of this work is to identify and include in the evaluation the necessary kinds of failures, e.g., the failures of the network, VMs, or physical machines. Moreover, considering 5G network slicing, the protection of the network slices can be considered in the VNF placement [159].

2) MEC SYSTEM LEVEL

a: DEPLOYMENT AND FAILOVER MECHANISMS

As already mentioned, a manner to make a MEC system resilient to the failure of MEH is to use failover mechanisms. Failover mechanisms that need a proper deployment of the MEHs, i.e., the MEHs need to be close enough in order for a user to be able to reach multiple MEHs [86].

For this reason, the deployment of MEHs can be performed by maximizing the failover capability [160] or by considering a 1+1 protection, where the users are able to connect to two MEHs (one is active and the other one is for backup) [161].

Proactive failover mechanisms can be considered in order to reduce the impact of the failure [162]. In case a user is not able to reach another MEHs, other users can be used as a relay in order to reach active MEHs [163].

TABLE 10. Dependability requirements in the MEC architecture according to standardization documentation.

Element	Dependability Requirements
General	
	- Additional tools are needed to generate workload and challenge the service in terms of service scalability, availability, and reliability [27]. - For alarm management, the following 3GPP-defined Integration Reference Points are used: ETSI TS 132 111-2 [155] and ETSI TS 132 332 [156] [29]. - The Multi-access Traffic Steering can be used by MEC apps and MEP for seamlessly steering/splitting/duplicating application data traffic across multiple access network connections [34].
MEH	
MEP	- The MEP shall allow MEC services to announce their availability [12]. - The MEP interacts with the MEC apps via the reference point Mp1, which provides functionalities, such as service availability and session state relocation support procedures [7], [9]. - The MEP may use available radio and core network information to optimize the mobility procedures required to support service continuity [36].
Virtualization Infrastructure	- N/A
MEC App	- Some MEC applications expect to continue serving the UE after a location change of the UE in the mobile network. In order to provide continuity of service, the connectivity between the device application and the MEC application needs to be maintained [7]. - Each MEC service instance that has previously registered in MEP and is configured for heartbeat sends a heartbeat message to the MEP periodically in order to show that the MEC service instance is still operational [9].
MEC Host-level Management	
MEPM	- The MEPM also receives Virtualised resources fault reports and performance measurements from the VIM for further processing [7]. - In the NFV variant, the MEPM-V does not receive Virtualised resources fault reports and performance measurements directly from the VIM, but these are routed via the VNFM [7], [15].
VIM	- The VIM collects and reports the performance and fault information about the virtualised resources [7].
MEC System-level Management	
MEO	- The MEO maintains an overall view of the MEC system based on deployed MEHs, available resources, available MEC services, and topology [7]. - The MEO interacts with the MEPM via the reference point Mm3 to keep track of available MEC services [7].
OSS	- The OSS interacts with the MEPM via the reference point Mm2 for fault management [7], [15]. - In the NFV variant, the OSS interacts with the NFVO for fault management [15].
User app LCM proxy	- N/A
MEC Federation	
MEF	- N/A

b: RESOURCE ALLOCATION

Another manner to improve the dependability of MEC is by a proper allocation of resources, usually computing and storage, in the different MEHs. There are already several works that address this aspect [164], [165], [166], [167], [168], [169], [170], [171]. Many of these works refer to the term *task* and they are talking about *offloading*. The term task is used to refer to an application or procedure, instead offloading refers to the migration of the execution of a task from the mobile device to a MEH.

Most of the current works have as target energy efficiency and consider the dependability metrics as requirements. The main difference between current works is the dependability metrics they are considering. Some works consider the failure probability of MEHs to develop k-out-of-n allocation schemes, where the tasks are distributed among several MEH and the task is correctly executed if at least k out of n MEHs are not failed [164], [165]. One work considers as reliability the probability of the delay bound violation, which is actually more similar to the performability [166]. Another work considers an offloading failure probability, which is connected to the error probability in the transmission between the user and the access where the MEH is located [171]. Another work considers also the execution reliability [167]. Finally, some works model the MEC system with queues and define the reliability as outage probability. There is an outage when a queue length exceeds a predefined threshold [168], [169], [170].

Furthermore, there are works that address the task allocation together with the user-host association [170] or

TABLE 11. Academic surveys related to MEC dependability.

Ref.	Aspect	MEC only	Main contribution	Relevance to MEC dependability
[86]	Fog reliability	No	Addresses the reliability in fog computing.	Presents the reliability challenges by combining the reliability requirements of cloud computing and networks of sensors and actuators.
[62]	MEC general	Yes	Surveys the use cases and the enabling technologies. Focuses on orchestration and related challenges.	Mentions the five-nine availability. Briefly discusses the reliability aspect in deploying MEC services and suggests dubbed checkpoints or, for improving the scalability, the replication of MEC service instances. Discusses also resiliency as a service enhancement.
[88]	Fog general	No	Surveys taxonomy, existing works, challenges, and future directions of fog computing.	Identifies the reliability of fog computing as a poorly discussed topic and suggests better investigation of the consistency of fog nodes and availability of high-performance services.
[64]	MEC communication	No	Focuses on joint radio-and-computational resource management	Diffusely mentions reliability in various contexts (link, transmission, server). Focuses more in detail on mobility-aware fault-tolerant MEC. Mentions resiliency and high availability of MEC system as requirement.
[66]	Edge general	No	Provides an overview of the state of the art and the future research directions for edge computing.	Inserts fault tolerance as a property in comparing the existing frameworks in edge computing. Mentions the availability of resources by stating that it is mostly dependent upon server capacity and wireless access medium for ensuring constant service delivery.
[67]	Edge security and resilience	No	Reviews some security and resilience aspects in MEC and fog in relation to cloud technologies.	Generally discusses resilience requirements and challenges in MEC together with fog and in comparison with the cloud.
[93]	Edge/IoT	No	Surveys how edge computing improves the performance of IoT networks	Addresses the dependability with respect to the storage by focusing on the recovery policy.
[68]	Edge general	No	Presents an overview of potential, trends, and challenges of edge computing.	Mentions fault tolerance (together with Quality of Service) as one of the challenges of edge computing. Discusses the need for proactive fault tolerance and automatic recovery.
[71]	MEC/IoT security	Yes	Overviews the MEC technology for the realization of IoT applications.	Diffusely mentions reliability in various contexts.
[74]	Edge general	No	Surveys edge computing, identifies requirements, and discusses open challenges.	Presents the provision of low-cost fault-tolerant deployment models as an open challenge.
[75]	Edge general	No	Provides a tutorial on edge computing and related taxonomy and surveys the state-of-the-art efforts.	Considers the RAS (Reliability Availability Survivability) as objective and the resilient fog system design as a challenge.
[77]	MEC Industrial Internet	Yes	Surveys the existing works on MEc for Industrial Internet.	Discusses the high reliability as a typical industrial requirement.
[78]	MEC general	Yes	Surveys MEC fundamentals, discussing its integration within the 5G network and relation to similar UAV communication, IoT, machine learning, and others.	Considers reliability as a challenge in MEC and diffusely discusses it.
[96]	Edge dependability	No	Explores dependability and deployment challenges in edge computing. Considers dependability in a wider meaning that includes security.	Presents resiliency challenges, which include new failure modes, network impact, limited fail-over options, multi-tenancy support, and interoperability.
[89]	Fog dependability	No	Surveys the research efforts on fault tolerance and dependability in fog computing.	Presents redundancy models and fault management solutions. Discusses the dependability challenges and open problems.
[79]	Edge general	Yes	Surveys different aspects of edge computing from an architectural point of view.	Mentions the improvement of reliability as a benefit of edge computing.
[97]	Wireless edge	No	Discusses the feasibility and potential of providing edge computing services with latency and reliability guarantees.	Overviews the challenges and the enablers for realizing hugh reliability in wireless edge computing.
[80]	MEC general	Yes	Surveys the MEC and focuses on the optimization of the MEC resources	Diffusely mentions reliability, availability, and resilience.
[99]	Vehicular Edge Computing	No	Surveys the state of the art of vehicular edge computing.	Diffusely mentions reliability.

consider that different users may have different dependability requirements [172].

c: MEO DEPENDABILITY

The MEO is a critical element because it is a single point of failure of the MEC system. A dependable MEO is important because if it fails the whole MEC system became unavailable. Moreover, the MEO is also important to manage failure of the other MEC elements, in order to have a fault-tolerant MEC system. For this reason, the design of the MEO and its functionality must be addressed carefully by considering the state-of-the-art technology [51] and techniques of Artificial Intelligence [52]. The challenges for the MEO are similar to the challenges for the NFVO, the same best practice should be

followed [173]. For eliminating the single point of failure, the MEO can be designed as logically centralized but physically distributed, as for the SDN controllers [174]. In this case, the coordination among the different MEOs becomes critical.

d: CONSISTENCY

Consistency is defined as the property of multiple elements to have the same information and vision of the system. The lack of consistency is a problem for a distributed implementation of the MEO, but also for naturally distributed elements as the MEH. For example, the consistency regarding the agreement in the event of a failure needs to be addressed [175].

3) GENERAL CHALLENGES

a: DEPENDABILITY MODELLING

A first way of evaluating the dependability of a new system is to realize dependability models. For example, this approach has been used to evaluate the availability of SDN [176] and NFV [177]. A dependability model can allow to evaluate the impact of the elements composing the MEC system and identify the critical issues. Advanced models can also consider the dependability correlation bet. The more common models techniques are derived by the Petri Nets, such as the Stochastic Activity Networks and Stochastic Rewards Nets (SRN), and can be implemented by using tools such as Möbius[2] or SHARPE.[3] Some works have addressed the modeling of reliability and availability in edge computing [178], [179]. For example, SRN has been used to model the availability of an edge system [180], a semi-Markov model has been used to evaluate the impact of VNF aging in MEC [181], and a two-level model has been used to evaluate the availability of a MEC system [182].

b: NETWORK DEPENDABILITY

The network connectivity is an important element in the MEC system [34]. It includes the access network to which the users are connected, the connection between the access and the MEHs, the connections between the MEHs, and the connection between the MEHs and the MEO. Unreliable network connectivity can have a huge impact on the dependability of the MEC system [10], [50]. Reliable network connectivity can be provided via physical redundancy and dedicated protocols.

c: SERVICE CONTINUITY AND USER MOBILITY

One property that is carefully addressed in the ETSI specification is the service continuity [45], [46]. The importance is also due to the nature of MEC, or more precisely MEC with mobile access networks.

ETSI defines the service continuity as *"the perception of the out-of-service time and can be measured by the latency between the terminated instance and the resumed instance of the same service, maintaining the instance state"* [45].

[2]https://www.mobius.illinois.edu/
[3]https://sharpe.pratt.duke.edu/

The causes of a lack of continuity are application software failure, malfunction of MEC system, loss of connectivity due to network failure, and user's voluntary or involuntary action.

Depending on the scenario and the application, ETSI defines different levels of service continuity: no continuity, low continuity, soft continuity, and hard continuity [45].

d: FAILURE MANAGEMENT

MEC might introduce more complex failure modes [89], [96]. This situation is a common problem in modern complex ICT systems [183]. Moreover, faults and failures in MEC (as generally in computing platforms) are hard to detect [86].

The fault management is diffusely addressed in the ETSI specifications [29]. Anyway, there are aspects that need to be properly addressed: uncontrolled error propagation by defining error-containment regions; recovery of faulty components [89], [184]; extreme event control [97].

e: DEPENDABLE ARCHITECTURE

Beyond the ETSI architecture, alternative edge computing architectures have been proposed in order to improve dependability. One architecture aims to deliver failure resistant and efficient applications [185]. Another work proposes a dependable edge computing architecture customized for smart construction [186].

V. PERFORMANCE

The performance is usually the main target of a new technology, where security and dependability are aspects that need to be guaranteed. For this reason, the research community has focused on the performance of MEC and many works are on the topic although the number of surveys on MEC performance is limited.

A. TAXONOMY

Given the wide nature of the performance, this subsection presents the performance taxonomy specific to MEC.

Figure 9 represents the taxonomy related to the performance. The *attributes* are the various metrics to evaluate the performance of a system. The *issues* are the causes that may lead to a lack of performance. The *countermeasures* are the methods to address the performance issues.

The attributes are presented first from a general perspective, then according to what has been defined by ETSI [27]. The issues and the countermeasures are related to a networking and computing context (which MEC belongs to) because having a general perspective would have resulted in a too broad introduction.

1) ATTRIBUTES

The performance of a system can be evaluated by means of metrics. There are different well-known metrics that can be divided into two classes. A first class contains general metrics focused on only one aspect, such as data transport service. The following well-known metrics belong to this class:

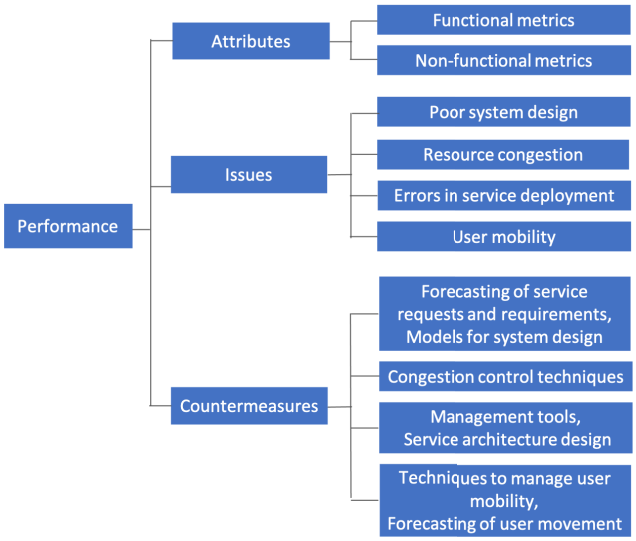

FIGURE 9. Performance taxonomy.

- *Throughput* is the number of bits or messages successfully delivered per unit of time.
- *Latency* is the delay introduced for completing a service, e.g., delivering a block of data between two points of the system.
- *Jitter* allows quantifying the latency variation.
- *Loss rate* refers to the number of bits or messages per unit of time that are lost during the service.

All these metrics can be measured at different layers of the protocol stack, depending on the particular service considered in the performance evaluation. For example, measuring the latency at the network layer allows us to quantify the delay introduced only by the network. This kind of metric is not sufficient to quantify the performance of a Voice-over-IP (VoIP) service because it does not take into account the latency added by the application during the elaboration of the received bits [187]. The second class is composed of metrics that aim to summarize in a value the performance of a complex service that requires the interaction of a set of components. Most of the metrics belonging to this class are related to the concept of Quality of Service (QoS) or Quality of Experience (QoE). These metrics can be classified as follows [187]:

- *Subjective metrics*, which require the involvement of humans for quantifying the experimented performance of the service. The quantification is performed using some reference scale, such as that defined in the Mean-Opinion-Score (MOS) procedure and its evolution [188].
- *Objective metrics*, which allow quantifying the performance by using machine-executable algorithms, such as PSNR (Peak Signal-to-Noise Ratio used for quantifying the video signal quality).

Starting from this general definition, the ETSI document [27] has defined a set of metrics specifically designed for the MEC systems. Given the high flexibility of MEC architecture, the MEC metrics have been defined with the following two goals:

- Evaluate the performance increase given by a MEC solution with respect to a non-MEC one;
- Compare the performance of different MEH locations within the network in order to select the most suitable MEH for the considered use case.

Taking into account these goals, ETSI defined two classes [27], which we will further present by summering the content of the ETSI document:

- *Functional metrics*: they quantify MEC performance impacting on user perception (often called Key Performance Indicators, KPIs). The set of KPIs is shown in Figure 10.
- *Non-functional metrics*: they are related to the performance of the service in terms of deployment and management. Figure 11 describes the set of non-functional metrics.

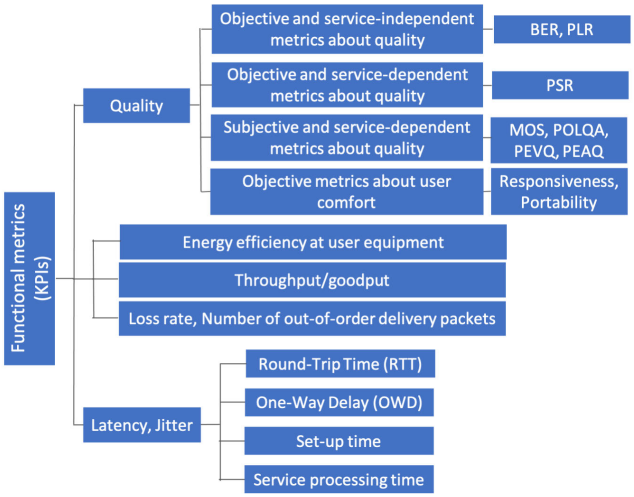

FIGURE 10. Functional metrics.

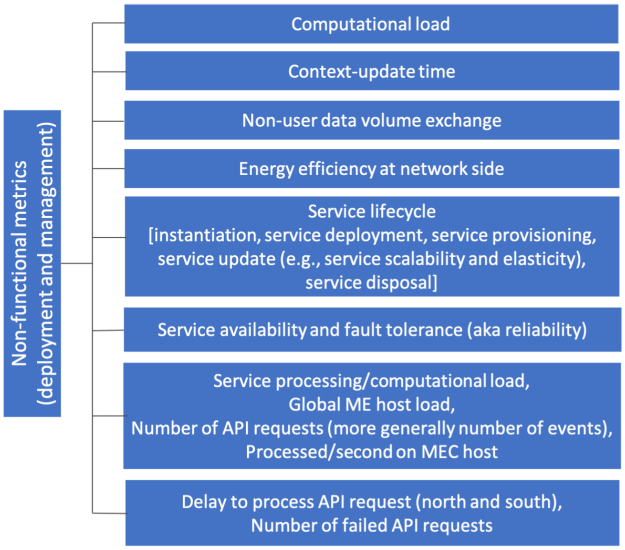

FIGURE 11. Non-Functional metrics.

As shown in Figure 10, the functional metrics consider the performance by both taking into account single aspects of the service (i.e., latency, jitter, loss rate, throughput, and energy efficiency) and referring to metrics that summarize the interaction of multiple components forming a complex service (i.e., all the metrics under the umbrella indicated as *quality*). The non-functional metrics in Figure 11 are strictly related to the aspects that are important for the network operator, i.e., deployment and management aspects. In some cases, these aspects are not related to the performance perceived by the users, such as *non-user data volume exchange* or *energy efficiency at network side*. Indeed, these performance parameters impact the costs necessary for the network operator to manage the service. On the other hand, some metrics, such as *delay to process API request* and *number of failed API requests*, negatively impact the user experience. Indeed, a high delay in processing API requests can increase the *set-up time* of a new service request and/or the *service processing time*. For both classes, the metrics need to be adapted to the particular MEC use case. In particular, the actual assessment of these metrics can depend on the particular service and/or application. For example, the latency in localization (time to fix the position) is different from the latency in content delivery. In both cases, one could measure all the statistics over the above metrics. In fact, all metrics are in principle time-variable and could be measured in a defined time interval and described by a profile over time or summarized through the following values: maximum value, mean and minimum value, standard deviation, and value of a given percentile.

2) ISSUES

Poor system design is one of the most important performance issues. The system design deals with the definition of the resources (and their location) necessary for supporting the user services. For example, if there are no (or insufficient) resources in the area where the user requires URLLC services, the service may need to be supported on a remote server, where the propagation delay is higher than the maximum accepted delay. The remoteness of resource location for video-on-demand services implies the involvement of many network resources with consequently a consumption increase of communication resources and energy.

Resources congestion can impact the different performance parameters, depending on the kind of congested resources. For example, the congestion of communication resources increases the latency for data transferring, the lack of storage resources can add data loss, whereas insufficient computation resource adds delay in processing data.

Service deployment can manifest problems at the network layer and service layer due to configuration errors or bad architecture (hardware and software) choices, which have a deep impact on performance.

User mobility adds variability to the features of the connection with the resources providing the service. The user movement can degrade the data rate available at the access network and increase the latency for achieving the service location. In some cases, in the new user location, there are no resources to support the service, leading to service interruption.

3) COUNTERMEASURES

To solve the presented issues, there are some general approaches. *Poor system design* issues can be coped with a preventive analysis of the amount of services and related features, in terms of offered load and service requirements. This information is then used as input in models developed for system design. The models differ from the main target of the design. The most common target is to reduce the capital expenditures (CAPEX) and the operating expenses (OPEX) by maximizing the number of services with satisfied requirements. Recently, some models aim to reduce energy consumption.

To cope with the *resources congestion* issues, there are two big classes of congestion control approaches: reactive and proactive. The TCP congestion control is an example of a reactive approach, whereas admission control of new service requests is an example of proactive *congestion control*. This control consists in rejecting new service requests when their admission degrades the performance of the already accepted services. The congestion is prevented by analyzing the available resources and estimating the required ones by the newly requested service.

The usage of management and monitoring tools at different layers represents an important countermeasure against service deployment problems. They help to detect and solve issues related to device, network, and service configuration. The analysis of alternative hardware and software architectures is necessary to prevent performance issues related to bad choices on deployment aspects, such as service implementation in dedicated hardware, the connection among the different components of a complex service, and virtualization technology used for service (or of one component) implementation in a central or distributed cloud.

The issues related to *user mobility* are well-known in mobile networks, where different techniques have been designed to maintain the network performance during user movements. Similar issues need to be considered at the service layer in the case of critical services, where the time necessary to achieve the resources offering the services is higher than the minimum latency requirements. In this case, the service migration is necessary. In all cases, *forecasting the user movements* can allow for proactively reserving resources and starting all operations related to the connection handover and/or service migration, reducing the issues related to user mobility.

A summary of performance issues and countermeasures is shown in Table 12.

B. STATE OF THE ART

1) STANDARDS, REGULATIONS, AND WHITE PAPERS

Table 13 shows the ETSI specifications that represent the state-of-the-art of performance requirements and regulations

TABLE 12. Summary of performance issues and countermeasures.

Issue	Description	Countermeasure
System design	- **Poor networking, memory and computation resources** negatively impact latency, loss rate, throughput and quality. - **Bad resources location** can negatively effects the performance also in the case of resource overprovisioning (but badly located).	- **Forecasting of the necessary resources** for each class of services, defined by taking into account the performance requirements and the offered load. - Usage of **models for system design** starting from the forecasted service requests and requirements.
Resources congestion	- The **resources congestion** can deeply impact the latency, throughput, loss rate, and quality at different layers.	- **Reactive congestion control** techniques aim to **detect the congestion status** of the system and then react by using strategies aimed at reducing the time for which the system is in this state, in order to reduce the negative effects of a congestion event on the user service. - **Proactive congestion control** techniques aim to prevent the congestion state of the system resources.
Service deployment	- **Service deployment errors and problems**: at the communication layer (e.g., IP address plan or routing protocols configuration errors, and radio channel selections) and at the service layer (e.g., software and hardware architecture selected for the service deployment). - The network softwarization and, in general, the service virtualization add further issues related to the used **virtualization technology**: these technologies differ in terms of required physical resource (e.g., memory required by the image and CPU utilization), and offered performance (e.g., latency and throughput).	- **Usage of management and monitoring tools** at different layers helps to detect and solve issues related to deployment errors. - These performance issues can be avoided by **proactively evaluating the performance of alternative hardware and software architectures**, e.g. service implementation in dedicated hardware, used CPU and OS, connections of different components of complex services, used virtualization technique (e.g., hypervisor-based, containers, and Unikernel) in the case of service implemented in general-purpose hardware.
User mobility	- **High variability of the edge resource utilization** can lead to temporary congestion, degradation of latency and, in some cases, loss of data. - The **user may move away** from the location of the resources supporting the requested services, degrading the performance and, in some cases, interrupting the service when there is no (or insufficient) resource in the new position.	- **Definition of techniques for managing user mobility** at network layer, such as the handover management defined in mobile networks, at service layer, such as live service migration permitted by agile virtualization technology. - In both cases, **forecasting the user's movements** can allow proactive reservation of resources and the initiation of all operations related to connection handover and/or service migration, thereby reducing the issues related to user mobility significantly.

for MEC. This list has been obtained by analyzing the documents shown in Table 3). Note that N/A here (as in Table 6 for the security and Table 10 for dependability) does not mean that performance requirements are out of scope or importance for the given element; N/A means that no explicit requirements are listed in the specifications.

Several API-focused standards [34], [36] describe some performance requirements to implement in some MEC elements. The standards that define the performance metrics is [27]. A lot of use cases with related performance requirements are described in [12]. Whereas studies for the MEC support of some important services and technical aspects, such as network slicing [26], NFV [15], mobility [45], and alternative virtualization technologies [16], indicate the performance required to the MEC elements. An important document describes the MEC 5G integration [19].

The white papers listed in Table 3 mainly discuss the problems and solutions related to the MEC deployment with 4G, 5G, and cloud RAN (CRAN) [18], [21] [23]. In these documents, some performance issues are presented, but no explicit performance requirements are clearly defined for the MEC elements. The enhanced DNS can have a key role in the performance of MEC services because it enriches the list of deployment options suitable to support the distributed MEC environment, in terms of providing the connectivity between devices and application instances [54].

2) ACADEMIC PUBLICATIONS

We can find the discussion of performance issues (and, in some cases, related solutions) of MEC deployment in many of the technical surveys dedicated to MEC. Among these, in Table 14, we summarize the most updated and important from the performance perspective.

C. CHALLENGES

In the following, we present performance challenges for MEC by categorizing them on MEC host level, MEC system level, and general challenges.

1) MEC HOST LEVEL

a: VIRTUALIZATION PERFORMANCE

The MEC architecture allows not only to improve the performance of network functions implemented in the form of

TABLE 13. Performance requirements in the MEC architecture according to standardization documentation.

Element	Performance Requirements
General	
	- A large set of use cases is described with the goal of deriving useful requirements. However, some requirements are defined by design constraints and do not originate from use cases [12]. - A number of performance metrics that can be used to demonstrate the benefits of deploying services and applications on a MEH are presented in [27]. - The MEC 5G integration implies several key issues. The issues and their related solution proposals are discussed in [18], [19]. - The current QoS-related information in RNI API is not sufficient in order to allow necessary prediction regarding the QoS performance (e.g. latency, throughput, reliability). Therefore, potential enhancements on RNI API for the prediction should be considered including both relevant measurements in RAN or processed results for the prediction [46].
MEH	
MEP	- The MEP on a MEH provides a framework for delivering MEC services and platform essential functionality to MEC applications running on the MEH [12].
Virtualization Infrastructure	- Multiple Alternative Virtualization Technologies (AVTs) need to be supported in the same MEH to satisfy the performance requirements of heterogeneous MEC services [16]. The different AVT solutions have an impact on MEC framework and on MEC management APIs [16].
MEC App	- When MEC is deployed in an NFV environment, the most simple possibility to realize the Data Plane needs the support of a physical network function or VNF or a combination thereof, and its connection to the network service that contains the MEC app VNFs [15].
MEC Host-level Management	
MEPM	- The MEPM receives Virtualized resources fault reports and performance measurements from the VIM for further processing [7]. - In the NFV variant, the MEPM-V does not receive virtualized resources fault reports and performance measurements directly from the VIM, but these are routed via the VNFM [7], [15]. - The MEPM shall be able to collect and expose performance data regarding the virtualization environment of the MEH related to a specific MEC app and a specific MEC app instance [12]. These data can be used to verify how well the Service-Level Agreements (SLAs) are met. - Different MEC applications running in parallel on the same MEH may require specific static/dynamic up/down bandwidth resources, including bandwidth size and bandwidth priority. To this aim, optional traffic management services are described in [34]. - The MEPM should be able to provide different sets of features/services in distinct Network Slice Instances (NSIs) [26]. - The MEPM should be able to provide the same feature/service differently in distinct NSIs [26]. - The MEPM collects usage and performance data per NSI [26]. - The MEPM-V can access the performance management and fault information of virtualized resources related to a particular ME app VNF instance that is lifecycle-managed by the VNFM [15].
VIM	- The VIM collects and reports the performance and fault information about the virtualized resources [7].
MEC System-level Management	
MEO	- The MEC system management should support the management of MEHs, including suspend, resume, configure, add and remove, by an authorized third-party [12]. - The functional requirements related to application mobility are reported in [36]. Examples of functional requirements are the maintaining of the connectivity between a UE and an application instance when the UE performs a handover to another cell associated/not associated with the same MEH, the support of two instances of a MEC application running on different MEHs to communicate with each other, or the use of radio/core network information for optimizing the mobility procedures required to support service continuity. - The MEO is able to trigger the application relocation owing to external relocation request, unsatisfied performance requirements, load balancing and disaster recovery [45].
MEAO	- The MEC system management shall be able to expose up to date performance data of the application to the authorized third-parties such as application developers and application providers [12]. - When MEC is deployed in an NFV environment, for performance enhancements, it can make sense to re-use the SFC functionality provided by the underlying NFVI for traffic routing. Differently from the simple solution based on the MEC App VNFs, in such a deployment, the Data Plane is obtained using the MEAO without a dedicated MEC App. The MEAO will need to translate the traffic rules into a Network Forwarding Path (NFP) and send it to the NFVO. The MEP will not control the traffic redirection directly but will pass requests to activate/deactivate/update traffic rules to the MEPM-V which will forward them to the MEAO. When receiving such a request, the MEAO will request the NFVO to update the NFP accordingly [15].
OSS	- N/A
User app LCM proxy	- N/A
MEC Federation	
MEF	- N/A

TABLE 14. Academic surveys related to MEC performance.

Ref.	Aspect	MEC only	Main contribution	Relevance to MEC performance
[62]	MEC general	Yes	Surveys the use cases and the enabling technologies. Focuses on orchestration and related challenges.	Discusses the role of MEO in orchestrating resource allocation and service placement for assuring efficient network utilization and QoE. Provides a qualitative comparison of different orchestrator deployment options.
[64]	MEC communication	No	Focuses on joint radio-and-computational resource management.	Provides a summary of future research directions for the joint radio-and-computational resource management referring to the performance issues related to the deployment of MEC systems, the Mobility management for MEC, and the Green MEC.
[66]	Edge general	No	Provides an overview of the state of the art and the future research directions for edge computing.	Provides a comprehensive review and comparison of the prevalent MEC frameworks. The comparison is based on different parameters, such as system performance, network performance, overhead of deployment, and system migration overhead.
[93]	Edge general	No	Presents an overview of potential, trends, and challenges of edge computing referring to the IoT Applications.	Discusses performance challenges related to IoT applications.
[71]	MEC for IoT	Yes	Holistic overview on the exploitation of the MEC technology for the realization of IoT applications.	Presents research topics related to MEC service level congestion control, latency-aware routing, and dynamic application routing.
[74]	Edge general	No	Surveys edge computing, identifies requirements, and discusses open challenges.	Presents a comparison of MEC solutions aimed to optimize the execution cost and deployment, reduce network latency, minimize energy consumption, and maximize throughput.
[78]	MEC general	Yes	Surveys MEC fundamentals, discussing its integration within the 5G network and relation with similar UAV communication, IoT, machine learning, and others.	Discusses the lessons learned, open challenges, and future directions on the performance of MEC integrated with the forthcoming 5G technologies.
[80]	MEC general	Yes	Surveys the MEC and focuses on the optimization of the MEC resources.	Provides a state-of-the-art study on the different approaches that optimize the MEC resources and its QoS parameters (i.e., processing, storage, memory, bandwidth, energy, and latency).
[97]	Wireless edge	No	Discusses the feasibility and potential of providing edge computing services with latency and reliability guarantees.	Overviews the challenges and the enablers for achieving low-latency and high-reliability networking; discusses network resources optimization techniques for a selection of use cases.
[99]	Vehicular Edge Computing	No	Surveys the state of the art of vehicular edge computing.	Presents a summary of techniques for caching, task offloading, and data management highlighting the considered performance parameter.
[189]	MEC and Game Theory	Yes	Discusses how Game Theory can address the emerging requirements of MEC use cases.	Overviews of Game Theory models for achieving a performance-cost balance in realistic edge network scenarios, and prospected future trends and research directions in the application of game theory in future MEC services.
[190]	Edge and IIoT	No	Discusses the opportunities and challenges of edge computing in IIoT.	Overviews of schemes for routing and task scheduling for performance improvements. Describes challenges and potential solutions of applying some new technologies to edge-based IIoT.
[191]	MEC for Industrial Verticals	Yes	Explores how the MEC is used and how it will enable industrial verticals.	Discusses MEC deployment bottlenecks and related solutions for performance improvements in a smart metropolitan area, where disparate verticals can coexist.

VNF, close to the users, but also to be prepared to host third-party services, creating a new market for the operator [192]. The virtualization technologies can deeply impact the performance of the MEC system, as shown by some works (such as the most recent [193], [194], [195] and reference therein), which presented an experimental comparison of virtualization technologies for implementing VNF and edge services. This well-known result spurs the study of new virtualization technologies and deployment paradigms in order to improve performance and resource efficiency. From container-based virtualization to micro virtual machines, new virtualization solutions claim to offer performance close to bare metal, with quick deployment and startup times.

Recent works analyzed the performance of multiple virtualization technologies, including VMs, containers, unikernels, and Kata Containers [193]. VMs have traditionally been the primary technology for VNF deployment, creating an isolated and secure environment with high associated overhead. To reduce the memory clutter of VMs, containers represent an interesting alternative solution, as they pack the application and its dependencies into a light and agile entity that can be run on any platform [196]. However, due to the

underlying shared kernel, containers face security problems in a multi-tenant environment [197]. To address the security issues related to the shared container kernel, Kata Containers have recently been proposed. Kata Containers act and perform like classic containers but provide stronger workload isolation by using hardware virtualization technology as a second layer of defense [198]. Unikernels are lightweight machine images that enclose the application and require only OS libraries and dependencies. It works in a single address space and has a much smaller attack surface due to its minimal nature. Each of these technologies faces a trade-off between isolation and agility. The studies presented in [193], [194], and [199] show the performance in terms of the following parameters:

- Image size of the service, which impacts the amount of storage required to host the application;
- CPU and memory utilization of the service, which has a great impact on the number of services that a physical server can run simultaneously;
- Throughput of service requests;
- Delay in responding to user requests, which can drastically degrade the user's satisfaction with the service (a website in the experimental analysis).

For example, the results of [193] show a boot time of the container (0.623 s) lower than that of VM (32 s), whereas the service throughput of the implementation of a HTTP server with VM (130 req./s) is higher than that of Container (143.4 req./s)

Kata Containers and unikernels are in their development phase, but they provide serious competition to the ecosystem of containers and VMs. Indeed, they offer a lightweight solution for the deployment and migration of virtualized services, improving the performance in terms of latency, throughput, and quality. These improvements are more evident in the presence of user mobility, where the agility and the lightweight of the virtualized services are key factors. However, Kata Containers are not mature yet for using them in production environments managing data-intensive workloads, as shown by the experimental results presented in [194]. Indeed, this paper confirms the results shown in [193]: Kata Containers are not really efficient in terms of memory consumption and speed, while still being a good deployment choice in security-sensitive multi-tenant environments.

b: PERFORMANCE ISOLATION

As shown in [200] performance isolation can be a very critical issue to take into account when VM shares the physical resources. With some initial experiments to understand the co-existential or neighbor-dependent behavior of VMs, the authors inferred that the problem of performance isolation can be improved by reducing the resource contention amongst the VMs on the same physical host. Performance isolation can be drawn to the lowest level abstraction of shared resources, like CPU, memory, network, and disk. For example, the disk is continuously being used by multiple processes waiting in the I/O queue. Thus, the I/O scheduler will play a vital role in resource contention, as studied in [201] and [202]. In [203], the authors present a systematic overview of existing isolation techniques in nodes and networks, especially in RAN and CN of 5G systems.

2) MEC SYSTEM LEVEL

a: MEC DEPLOYMENT

The network operator selects the number of MEHs and their location analyzing different technical and business parameters, such as available site facilities, supported applications and their requirements, measured or estimated user load, etc. As shown in Figure 6, the network operator has four different deployment options.

MEHs can be deployed in different network locations: from near the gNB to a remote data network. Although running MEHs far from the edge can be useful in scenarios in which compute power requirements are stricter than latency ones, the most interesting scenario is represented by locating MEHs close to the user (e.g., at the gNBs of a 5G system). This scenario adds two very important features for enabling new services:

- the reduction to low values of the delay between the end-user device and the MEH hosting the application, which enables low-latency services;
- the access to user context, such as the user channel quality conditions or user location, which enables context-aware services, e.g. services adaptive to network conditions.

The design of MEHs location and the instantiation of MEC services require the analysis of multiple trade-offs for efficient usage of physical and virtualized resources. Depending on the use case, the processing power demands and the latency requirements can be very heterogeneous. Other than the QoE of the users, the location of MEHs must consider the Total Cost of Ownership (TCO). On one hand, the centralized cloud decreases the TCO. On the other hand, the centralized cloud fails to address the low latency requirements. A good trade-off can be achieved by utilizing existing infrastructures such as Telco towers, central offices, and other Telco real estates. In the case of a mobile network, MEHs can be located in various parts of the network architecture, ranging from uCPEs (mainly at customer premises) to RAN-edge (co-located with base stations), Smart Central Offices (where the MEH could be hosted co-located with CRAN aggregation points), or edge data centers at the local/regional level.

The deployment problem has been analyzed by considering different optimization approaches and performance parameters. For example, using Shanghai Telecom's base station dataset, some works consider the MEH placement problem with the goal of minimizing the energy consumption [204] or balancing the workloads of MEHs while minimizing the access delay between the mobile user and MEH [205], [206]. Other recent works aim to minimize the cost of service providers while guaranteeing the completion time of services [207] or to minimize the number of MEHs while ensuring some QoS requirements [208]. In general,

the deployment of MEHs requires defining multi-objective problems, such as i) finding a Pareto front optimizing the time cost of IoT applications, load balance, and energy consumption of MEHs [205], ii) optimizing the response time taking into account heterogeneity of MEC/cloud systems and the response time fairness of base stations, which may significantly degrade the system quality of services to mobile users [209], iii) finding an optimized trade-off between response delay and energy consumption [210]. An interesting issue is the update of the MEC infrastructure. As an example, in [211] the authors consider the problem of scaling up an edge computing deployment by selecting the optimal number of new MEHs and their placement and re-allocating access points optimally to the old and new MEHs. In this case, the considered performance is the Quality of Experience of users and the QoS of the network operator. As concerning the integration MEC-5G, an interesting problem is the determination of the MEH and UPFs optimal number and locations to minimize overall costs while satisfying the service requirements [212].

b: RESOURCE ALLOCATION

Taking into account the available resources, the network operator can solve the service placement problem. Placing MEC services over a number of MEHs can prove to be critical for the user QoE and should take into account gravity points, e.g., shopping malls, which attract a plethora of users. Indeed, a bad design of the service placement can lead to the saturation of the available resources with the consequent rejection of service requests or a worsening of the user experience. The flexible availability of resources plays a crucial role in the performance of a service. For example, wireless bandwidth and computing resources can be dynamically assigned to a service for achieving optimal benefits from MEC system. Orchestrating a MEC system in terms of resource allocation and service placement is critical for assuring efficient network resource utilization, QoE, and reliability.

A set of works considers the data caching problem in the edge computing environment, proposing schemes that maximize the data caching revenue of the operator [213], or that improve content delivery speeds, network traffic congestion, cache resource utilization efficiency, and users' quality of experience in highly populated cities [214]. Referring to the VNF architecture, the SFC and VNF placement can be performed by reducing the execution time and the resource utilization [215], taking into account both service requirements and the resource capacity in the edge [216], maximizing revenue at the network level while matching demand [217], minimizing both energy consumption and resource utilization [218], or maximizing the number of user request admissions while minimizing their admission cost (i.e., computing cost on instantiations of requested VNF instances and the data packet traffic processing of requests in their VNF instances, and the communication cost of routing data packet traffic of requests between users and the MEH hosting their requested VNF instances) [219].

c: USER ASSOCIATION

The most simple strategy for the user-MEH association is to allocate the nearest MEH that offers the requested service. Indeed, this approach agrees on the proximity strategy that is desired from the performance perspective, e.g. latency and energy consumption. However, this simple strategy can lead to performance degradation when other aspects, such as the load of the MEHs, the available transmission capacity, and the users' mobility are not considered. For example, an increase in requests for MEC services in a given area can lead to an overload of MEC resources, generating performance bottlenecks. To avoid this problem, the user-MEH association strategy needs to take into account the load status of the MEC infrastructure. Different works propose optimization solutions for some metrics, such as latency and QoE, that simultaneously distribute the load between different servers, e.g. [220]. Furthermore, in the case of mobile users, the selection of the edge cloud becomes crucial due to the uncertainty of movement and wireless conditions. In this scenario, other than the MEC resources, the user mobility impacts also the available network capacity, given the uncertainty of the number of users sharing the resources and, in the case of the wireless access network, of the wireless channel that impacts the radio resource efficiency.

The user-MEH association problem can consider different aspects, such as i) the enhancement of the user allocation rate, server hiring cost, and energy consumption by means of an association scheme able to consider that edge users' resource demands arrive and depart dynamically [221], ii) the dynamic balancing of the computation load for neighboring MEHs [222]. Other approaches jointly consider the MEH placement and association problem with the goal of minimizing the deployment cost [223] or number of MEH [224] while guaranteeing certain end-to-end service latency.

3) GENERAL CHALLENGES

a: SERVICE CONTINUITY AND USER MOBILITY

Service mobility is a key aspect that can impact MAC performance in a mobile network scenario. Indeed, to maintain a high QoE it is of primary importance not only to establish a connection between the end user and MEC resources but also to maintain it throughout all the necessary stages. Optimal end-to-end session connectivity needs to be maintained for the entire course of service usage. To achieve this goal, the MEC services should be able to migrate quickly depending on user movements. The user movements can frequently change the anchor points of MEC services (e.g., from one MEH to another). In this scenario, ensuring optimum QoE for a delivered MEC service becomes challenging, especially for delay-sensitive applications. At the IP level, the Distributed Mobility Management (DMM) [225] represents a notable solution towards managing user mobility, overcoming also the scalability and reliability drawbacks of centralized mobility schemes. At the service level, the management of IP mobility is not sufficient to avoid quality degradation due

to the redirection of MEC service requests of the mobile user to a distant edge hosting the service. The MEC architecture aims for a single-hop connectivity to the service. When the user is moving away from the MEH serving its request, the MEC service should migrate from the old to a new MEH closer to the new user location for maintaining the same performance. Procedures for service migration take time, especially when these require moving large amounts of data. The user can experience a degradation of application performance, and in some cases, the service continuity cannot be guaranteed. Numerous studies have addressed the problem of VM migration, but new technical challenges have emerged from the analysis of the problem from a service point perspective [226]. In particular, the time it takes to prepare a VM for the target MEH, transfer it over the network, and finally deal with the problem of changing the IP address after relocating the VM, makes it difficult to achieve service continuity. The analysis of this problem starts with the study of the features of different virtualization technologies. Other than presenting remarkable overhead in terms of both processing and storage capabilities, hypervisor-based virtualization techniques show high latency for start-up activation and migration procedures. Container-based virtualization enables high-density deployment of services, and has limited features to support stateful service migration between different host MEHs. In both classes, the main drawbacks derive from the stateful nature of the application, which implies that the state of the application and the network stack must be preserved in the case of migration or failure. To alleviate these drawbacks, the stateful ''Follow me Cloud'' paradigm has been proposed [227]. Based on this new paradigm, some recent works propose mechanisms for fast container-based live migration, such as [228] and [229]. In [230], the authors present a survey on the techniques for service migration at the edge of the network. The development of stateless applications relies on user inputs or distributed shared storage, avoiding the storage of internal states. This feature allows a stateless application to be replicated on different MEHs. Based on the specific offloading request and current connectivity quality, the most appropriate instance could be selected. Examples of studies on the performance of stateless migration are [231] and [232].

VI. MULTI-ASPECT CHALLENGES

Dependability, security, and performance are all three important aspects in 5G MEC. These aspects are usually addressed individually but they are not independent and they can be actually conflicting (i.e., a solution for improving one aspect may impact the others).

Three examples of conflicts are shown in Figure 12. Reading the figure clockwise, the examples are the following: the usage of encryption to gain security (more precise, confidentiality) causes a delay and therefore a reduction of the performance; the reduction of the energy consumption can be achieved by consolidation (i.e., reduction of the active elements), but the consolidation can create a single point of failure and therefore a reduction of the dependability;

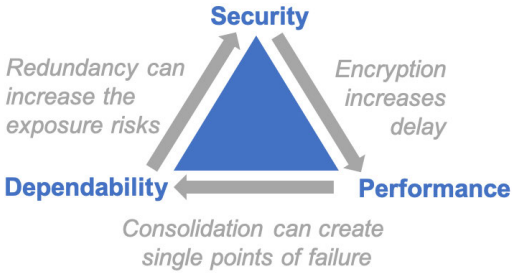

FIGURE 12. Examples of conflicts.

the dependability can be achieved by redundancy (i.e., deployment of multiple elements that perform the same functionality), the redundancy can increase the exposure risk since more elements can be attacked.

All potential conflicts will need to be studied in the design and operations of future 5G-MEC systems. In this section, we present the challenges when these three aspects are considered together. The presented challenges are summarized in Table 15.

A. MEC HOST LEVEL

1) PHYSICAL DEVICE

Section IV-C has introduced that physical dependability can be achieved by using redundant hardware and software within a MEH. Similarly, Section III-C has introduced the challenges related to physical security at the MEH level. The redundancy can have an impact on performance and security. For example, adding redundant physical devices increases the vulnerability related to direct physical access (e.g., the attacker can perform physical modifications on the device or on the communication link). The countermeasures against these threats increase costs and can impact performance. On the other hand, it was already mentioned that possible alteration or deletion of data (as a consequence of an attack) requires adequate backup and recovery techniques, which are feasible by means of dependability. Cryptographic primitives such as secret sharing schemes can be useful to allow recovery capabilities with lower duplication rates. Moreover, physical security mechanisms are normally beneficial for performance, as they usually introduce less latency [70], [78], [138].

2) VIRTUALIZATION TECHNOLOGY

In the previous sections, the importance of virtualization for each aspect has been introduced.

From a dependability perspective, the live VM migration in OpenStack [157] and VNF placement in a MEC-NFV environment [158], [159] have been introduced in Section IV-C. In both cases, dependability is investigated together with the performance. In [157], the authors analyze the impact of the system pressures and network failures on the performance of VM live migration by considering the migration time. In [158], the authors address the problem of VNF placement

TABLE 15. Summary of the multi-aspect challenges.

Topic	Challenges
MEC Host Level	
Physical device	- Evaluate the impact on performance and security when *redundant hardware and software* is used to increase the dependability of a MEH. - Investigate the use of cryptographic primitives (e.g., secret sharing schemes) to achieve dependability with a decreased rate of duplicated data and lightweight cryptography to achieve better performance.
Virtualization technology	- Develop *VM migration* and *VNF allocation* algorithms able to jointly take into the requirements in all three aspects. - Explore new methods aimed to *secure the container-based virtualization* without degrading its performance. - Develop *new lightweight and easy-to-instantiate virtualization. technologies* or *new mixing and nesting models of virtualization technologies* able to support MEC URLLC services with security constraints.
MEC System Level	
Deployment and design	- Evaluate the *deployment of MEHs* including dependability (failover), performance, and security requirements. - Develop new *failover mechanisms* able to jointly take into account security and performance constraints. - Define robust design models aimed to optimize performance while satisfying dependability constraints.
Resource allocation	- The allocation of computing and storage resources, also called *task offloading*, should not only aim to performance targets (such as energy efficiency or latency reduction), but also consider dependability and security requirements.
User association	- New multi-constrained path computation algorithms need to be developed for establishing user association methods able to jointly satisfy the constraints imposed by the three aspects. - Other than performance-related metrics, new metrics need to be defined for security and dependability to find finding Pareto frontier when the three aspects are jointly considered.
MEO	- The reliability and the performance of the MEO system can be increased by the design of *distributed MEO architecture*, which must be deeply analyzed also from the security perspective. - The *MEO functionalities* should aim to orchestrate and manage the MEC system by jointly considering the three aspects.
Consistency	- Consistency is an important requirement given the distributed and redundant nature of the system. These two features aim to improve the scalability and dependability of the system, but their impact on performance and dependability should be investigated. - Deeply analyse new methods proposed for guaranteeing consistency (even in the case of active adversaries).
MEC Federation Level	
Heterogeneous scenario	- The interconnection between the MEC systems that are geographically distributed and focused on different use cases should guarantee proper levels of security, dependability, and performance.
Multiple operators	- Having MEC systems that belong to multiple operators has several challenges: from privacy concerns to end-to-end performance guarantees; from security exposure to end-to-end dependability; from compatibility issues to waste of resources.
Multi-domain orchestration	- The MEC federation can be over multiple domains (edge and cloud). Infrastructure-as-Code can be used to implement a multi-domain orchestration.
General Challenges	
Modelling	- Develop new models for a joint evaluation of the 5G-MEC system from the three different perspectives.
Network connectivity	- The network connectivity has an important impact on dependability and performance requirements and is the source of potential threats. - New technologies are under development for increasing connectivity performance, such as Teraherz communications. This activity must be integrated with the development of new models for analyzing the impact of these technologies on the MEC system reliability and performance jointly. - Develop new solutions for mMTC and URLLC use cases, such as those based on NOMA, able to improve system capacity while satisfying the constraints on dependability and security, in scenarios characterized by a large number of users exchanging small blocks of data.
Service continuity and user mobility	- The solutions for achieving service continuity (even in the case of user mobility) may lead to conflicts between the three aspects. New solutions must be developed considering the effects on these three aspects jointly, trying to achieve Pareto optimal.
Monitoring and detection	- The monitoring and detection for security, dependability, and performance should be jointly executed to exploit the cross information. For joint monitoring and detection, AI-based techniques can provide extra capabilities.
Ubiquitous Pervasive Intelligence	- The MEC should be managed and orchestrated via secure, dependable, and performing intelligence. - The security and dependability can be provided via edge intelligence solutions.
Joint KPIs	- Availability should include both security and dependability causes, moreover a system should be considered up when only when the performance requirements are met.

by considering two conflicting objectives, namely minimizing the access latency and maximizing the service availability.

The MEC scenario can be characterized by lower computations and storage resources with respect to the cloud scenario. The hypervisor-based virtualization becomes unsuitable given that the image files of VMs are large and its overhead is non-negligible. Furthermore, when MEC is integrated with 5G, user mobility adds new requirements to the virtualization techniques. In particular, the service continuity and the fast migration of service between MEHs extend the popularity of container-based technology, which allows to easily build, run, manage, migrate, and remove containerized applications. The different container solutions (e.g. LXC, LXD, Singularity, Docker, Kata Containers, etc.) offer diverse performance in terms of CPU and memory load, security, networking bandwidth, and disk I/O, and configuration options that can further improve the performance in some specific scenarios [233]. The isolation mechanism of existing container-based technologies is weak, due to the sharing of one kernel among multiple isolated environments. This feature adds four types of problems to consider for container security: (i) protecting a container from applications inside it, (ii) inter-container protection, (iii) protecting the host from containers, and (iv) protecting containers from a malicious or semi-honest host. A big challenge is to explore new methods (e.g., using trusted images, managing container secret, securing the runtime environment, and vulnerability scanning), which can secure the container-based virtualization without degrading the performance in terms of agility, resource consumption, and computation delay. Alternatively, the unikernel technology guarantees security by the isolation provided by the underlying operating system or hypervisor. Unikernel is more secure than the container. However, from the performance perspective, unikernel presents low usability given that it does not have a shell and does not support online debugging, online upgrades and updates either. In the case of failure of a unikernel application, only the reboot solves the problem. Application and/or configuration updates require recompiling the source code to produce a new unikernel and deploy a new version. This action can be very costly and sometimes prohibitive. This weak point of unikernel technology represents a big challenge that needs to be considered in order to evaluate the most suitable virtualization technologies for MEC, able to achieve a good tradeoff between performance, dependability, and security. This evaluation can converge to a single new virtualization technology or a set of them, each one achieving the best tradeoff in some specific use cases.

B. MEC SYSTEM LEVEL

1) DEPLOYMENT AND DESIGN

The problem of the deployment of MEHs in a particular area is commonly addressed with the aim of improving performance. Many existing studies focus on different goals, for example, to minimize the overall latency of mobile users, to maximize the overall throughput of the MEH network, etc. [204], [206], [209]. Many of these studies assume that the MEHs are free of failures. However, in a dynamic and volatile network scenario, MEHs are subject to runtime errors caused by various events, such as software exceptions, hardware errors, cyber attacks, etc., similar to what occurs in cloud servers [234]. When these failure events happen, the design objectives and corresponding constraints can easily be violated. The quality of experience will decrease immediately and significantly, especially those of latency-sensitive applications. During a failure event, mobile users can be disconnected from the MEC infrastructure and the related services are unavailable. This scenario can have a disruptive impact on the perceived quality of the offered services particularly in areas with a high density of users. Given the negative effects of the MEH failures, the design of the MEC infrastructure must consider the reliability aspects of the system. The robustness of the MEC infrastructure requires coverage of a specific area with a number of MEHs greater than strictly necessary to guarantee performance in ideal situations. Making a MEC design considering only the performance, without taking into account the robustness of the infrastructure, leads to localizing the MEHs, minimizing the overlap of the coverage of each of them. Simply maximizing the collective coverage of the MEC infrastructure could lead to situations where no MEH can take control of mobile users disconnected from failed MEHs. Such a MEC infrastructure is highly vulnerable to runtime errors.

As already presented in Section IV-C, to alleviate this problem, recent studies [160], [161] investigate the dependability in the deployment of the MEHs because of its impact on the effectiveness of the failover mechanisms. Of course, considering only the dependability target is also not correct. The dependability should be jointly considered with the performance. For example, if both dependability and latency prefer a denser deployment of MEHs, energy consumption, and economic costs push for a less dense deployment. For this reason, the deployment strategies will aim to find the best trade-off in the given scenario.

Moreover, the failover mechanisms themselves should consider performance but also security aspects. For example, the failover mechanisms should be fast (for example, it can be proactive [162]) in order to reduce the delay. Moreover, using other users to rely on to reach distant MEHs [163] might introduce severe security threats.

2) RESOURCE ALLOCATION

As we have already mentioned in previous sections, another challenge addressed by current studies is the allocation of resources, usually computing and storage, in the different MEHs. As mentioned in Section IV-C, the resource allocation in MEC is often called task allocation, since an application or procedure (task) is moved (offloaded) from the local execution in the mobile device to a remote execution on a MEH.

Most of the current works have as a target the energy efficiency. As mentioned in Section IV-C, several works also consider the dependability metrics as requirements [164], [165], [235]. Other works implicitly consider reliability and latency by focusing on the queue length [168], [169], [170]. Some works consider both reliability and latency as constraints [166], [167], [236] or as targets [171].

To the best of the authors' knowledge, no work considers security aspects in depth. The importance must be better investigated. For example, the task can have different security requirements, and the MEHs have different security guarantees. A general good practice is the adding of a threshold for resource consumption and guarantee of some resource availability for security functionalities [25]. Otherwise, for example, by compromising the user apps an adversary can use too many resources and thus affect performance. More specific approaches consider for example the usage of a low resource Intrusion Detection System (IDS) [76] or lightweight cryptography, as already discussed in Section III-C. A first work that tries to address all three aspects (security, dependability, and performance) together is [237].

3) USER ASSOCIATION

As mentioned in Section IV-C, some works address the task allocation together with the user-host association [170]. For some critical applications requiring low latency and high reliability, such as in some V2X and Industrial IoT scenarios, the design of user association algorithms jointly considering the three aspects is of paramount importance. Multi-constrained optimal path computation algorithms have been proposed for solving routing problems [238]. Some of these can be extended to solve the user association problem. Other than performance-related metrics, new metrics summarizing the security and the reliability level of a link and a node need to be studied and jointly considered in the path computation towards alternative MEHs. Through these tools, path computation solutions towards different MEHs can be then compared to select the MEH satisfying the user requirements on the three different aspects. To reduce the complexity added by the path computation problem, similar metrics can be defined only for the MEH to develop a user association algorithm aimed at finding the best trade-off among the three aspects.

Concerning security and performance, whenever possible, perform user security services locally, with no need for centralization (e.g., identification and authentication [70]).

4) MEO

The MEO is an important element of the MEC architecture. It is a logically centralized element that is the main responsible for the system-level management of MEC. The ETSI MEC architecture [7] specifies that the MEC applications have a certain number of resource requirements, and these requirements are validated by the MEO.

As introduced in Section IV-C, the MEO is crucial for the availability and reliability of the MEC system because it orchestrates the creation of a MEC and the fault management of the other MEC elements. The solutions aimed to have a dependable MEO, such as physical distributed MEO, can generate challenges from a security perspective given the increment of the exposure risk. As discussed in Section III-C, special attention must be given to the virtualization security.

5) CONSISTENCY

In Section IV, consistency has been presented as a property needed when we use redundant elements for improving the dependability of a system. This is valid for the MEO but also for the MEHs (during the handover). Consistency is also important in the case of distributed systems, which aim, not only to increase dependability but also to increase scalability. For achieving consistency, the redundant and/or distributed elements need to exchange information about their states in order to have a consistent global view of the system. This information exchange can have an impact on the security and performance of the system and can be a source of risks that a potential attacker may exploit (i.e., an active attacker might aim to break the consistency of the system). Moreover, the information exchange can use resources and impact the performance of the system.

C. MEC FEDERATION LEVEL

1) HETEROGENEOUS SCENARIO

The MEC federation will operate in a heterogeneous scenario where the MEC systems are geographically distributed and focused to different use cases. This scenario brings new challenges, for example the interconnection between the MEC systems should guarantee proper levels of security, dependability and performance, which can be different from use case to use case. Therefore, the interconnectivity should provide different guarantees and isolation among the flow.

2) MULTIPLE OPERATORS

One of the possible business cases of the MEC federation considers MEC systems belonging to multiple operators [56]. This scenario implies many challenges in managing the MEC federation in all aspects including privacy concerns, end-to-end performance guarantees, security exposure, end-to-end dependability, compatibility issues, and waste of resources.

3) MULTI-DOMAIN ORCHESTRATION

Moreover, the MEC federation can be over multiple domains (edge and cloud). This scenario requires a multi-domain orchestration [56]. One of the possible orchestration methods can be the Infrastructure-as-Code, which works as a common tool that allows abstracting diverse provisioning methods (API, CLI, etc.) used in the individual domains and activate infrastructure components by using a high-level language [56]. Infrastructure-as-Code can be used to implement a combined MEO/MEPM/VIM, as deployed in the Linux

Foundation Edge (LFE) Akraino Public Cloud Edge Interface (PCEI) blueprint.[4]

D. GENERAL CHALLENGES

1) MODELLING

A joint evaluation and analysis of the security, performance, and dependability of 5G-MEC is needed. In the literature, there are several works that focus on joint modeling: performance and dependability [239]; security and dependability [153], [240]. There are also tools that aim to jointly evaluate these aspects, such as Möbius [241]. To the best of the authors' knowledge, there is no current work that jointly evaluates the three aspects of MEC. Work focuses on dependability and security in a medical IoT context [242].

2) NETWORK CONNECTIVITY

The network is a key part of a MEC system. It includes the network access to reach the user, and the edge network to interconnect the MEHs and other MEC elements. The network has an impact on performance (lack of requirements), dependability (lack or degraded connectivity), and security (the MEC system is exposed via the network).

Recently, new studies are focused on the definition of new wireless communications systems, such as mmWave and TeraHertz. This trend is spurred by the need of having a very high data rate, necessary for reducing the delay related to task offloading of applications. However, these new communication systems are vulnerable to blocking events due to obstacles or beam collisions, which can interrupt the data exchange. As a concern, this trend, one set of challenges is related to the study of new beamforming techniques based on antenna arrays both at the transmitter and at the receiver side, and to the usage of multi-link communications. Another set of challenges is related to the simultaneous study of the performance and the reliability of the network connectivity with the goal of improving some performance parameters, such as latency, energy consumption, and deployment costs while achieving the reliability requirements. As an example, the reliability can be improved allowing the UE to send information to different MEHs, via different mmWave beams and over all the available UE-MEHs links simultaneously. More effort is necessary to define strategies, such as the Parallel Redundancy Protocol (PRP), in order to find an acceptable trade-off between network resource consumption and achieved reliability.

For mMTC and some URLLC use cases, the Non-Orthogonal Multiple Access (NOMA) represents a key technology of 5G for improving the system capacity and the connection density [243]. The key idea of NOMA is to share a given resource slot (e.g., time/frequency) among multiple users, and apply Successive Interference Cancellation (SIC) to decode the transmitted information. NOMA allows grant-free transmission. This feature is particularly important for the uplink transmission of small data blocks, such as in the IoT scenarios. In this case, the control signaling overhead is reduced, as well as the transmission latency and the power consumption of the terminal device. However, NOMA adds some security issues that need to be carefully studied. In the case the NOMA connection is used by two users for offloading tasks to a MEH at the same time when SIC is performed, one user can decode the other user's message. During this period, an eavesdropper or an attacker may attempt to decode the mobile user's message. To cope with this problem, the key challenges are to combine physical-layer security and NOMA-MEC in order to find solutions that can achieve a good trade-off between security and performance. Given the NOMA features, the performance is related to the latency, system capacity, and energy consumption on the user side.

Concerning both security and performance, Soft-VPLS (discussed in Section III-C) allows different traffic categories (e.g., MEC service requests, user data, control statistics) to be routed via distinct tunnels, aiming to enhance both end-to-end security and overall communication performance [76].

3) SERVICE CONTINUITY AND USER MOBILITY

The service continuity has been defined in Section IV and is critical even because the user(s) might be mobile. A lack of service continuity has a huge impact on both performance and dependability. However, it is often not easy to implement securely while satisfying performance needs.

Solutions to improve the service continuity can lead to a conflict between performance and dependability. In MEC-host-assisted proactive state relocation for UE in connected mode [45], the improved latency is accomplished at the price of relocation success, because failed handover operation may nullify the state transition preparation made by the MEHs. Therefore, a trade-off between latency and reliability is needed.

Similarly, solutions to achieve secure user mobility can lead to a conflict between security and performance. Special security protocols need to be set in place to allow secure mobility of users (e.g., from moving from one MEH to another) so that this process does not expose sensible data (e.g., identifiers, keys, sensible data). Forward and backward security play an important role in these scenarios (e.g., compromising a key shared between the user and a MEH should not expose previous or further keys shared by the user with past or future MEHs). Protocols such as key update or setting up a new security context (if needed) introduce some latency that, if not properly adopted, might affect performance.

4) MONITORING AND DETECTION

As presented in Section III-C, monitoring and detection are key elements for the security in 5G MEC, but they are also important for fault and performance management. The monitoring and detection for security, dependability, and performance should be addressed jointly in order to exploit the cross information and provide advanced features. As already mentioned for security, AI-based techniques can

[4]https://www.lfedge.org/projects/akraino/release-4-2/public-cloud-edge-interface-pcei/

be developed to provide advanced monitoring and detection functionalities [52]. Moreover, Virtual Machine Introspection (VMI), and hypervisor machine introspection should monitor the activities in terms of resource utilization to prevent performance issues and DoS attacks [76]. As already discussed in Section III-C, a good tradeoff between local and global monitoring techniques would be beneficial for security in relation to performance.

5) UBIQUITOUS PERVASIVE INTELLIGENCE

As introduced in Section II, the main innovation of 6G is the ubiquitous and pervasive use of intelligence. MEC will help the achievement of such intelligence [2]. MEC will provide the computing resources close to the end users that will be used to provide the intelligence. MEC will use intelligence to manage and orchestrate its resources. In using and proving intelligence, there is the need to take care of security, dependability, and performance. The 6G requirements show extreme performances and 7-nine values of availability and reliability. These performance and dependability targets need to be jointly addressed in order to provide new specific use cases that are the ''extreme'' combination of the 5G use cases [244]. Security will need to be also considered because trustworthiness is one of the main 6G features. The 6G security will be enabled by trust foundations, privacy-enhancing technologies, AI/ML assurance and defense, distributed ledger technologies, quantum security, and physical-layer security [245].

6) JOINT KPIs

New KPIs can be defined to jointly consider all the three aspects: security, dependability, and performance. Given the Tables 4 and 8, both security and dependability have availability as an attribute, even if its definition in the two contexts is different. Availability can be jointly defined as the readiness to access a correct service by legitimate parties. A correct service must also meet the performance requirements (which can be related to the metrics defined in Figure 10), and the causes of failing the performance requirements can be both failures and attacks, not only load dynamics.

VII. DISCUSSION AND CONCLUSION

In this survey, the state of the art and the challenges of the 5G MEC have been studied with respect to three aspects: security, performance, and dependability.

First, the ETSI MEC architecture has been introduced including the NFV-based version and the integration with 5G.

Second, for each aspect, the taxonomy, the state of the art, and the challenges have been presented.

The taxonomy has given a general introduction to the aspects to the readers that are only experts in the other aspects. Comparing the taxonomy of the three aspects, we can notice differences - especially for the issues and the countermeasures, but also similarities - such as the presence of availability as an attribute for both security and dependability.

The state of the art has introduced the standards and the current surveys that address the investigated aspect of 5G MEC. It resulted that security is the most studied aspect of 5G-MEC systems, while dependability is the least studied. The requirements for each element of the ETSI MEC architecture have also been highlighted, as indicated in the specifications. It was discovered that ETSI does not specify yet the requirements for important elements, such as the virtualization infrastructure, the user app LCM proxy, and the MEF.

The challenges of the investigated aspect of 5G MEC have been presented by categorizing them in MEC host level, MEC system level, and general challenges. This study has shown that, although the main target of 5G-MEC systems is to improve the performance of network services, many works have addressed the security and strict security requirements have been specified. Fewer works have addressed the dependability of 5G MEC, even though URRLC services and mission-critical applications have strict requirements on dependability. Moreover, several challenges, such as deployment, resource allocation, virtualization, service continuity, and MEO, are common to multiple aspects, but they are not jointly addressed yet.

Finally, the challenges of jointly addressing security, dependability, and performance have been investigated by using the same categorization (MEC host level, MEC system level, and general challenges) plus the MEC federation level. The investigation has shown the importance of jointly addressing the three aspects and how focusing on only one aspect can cause the failure of meeting the requirements on the other aspects. The new concept of MEC federation makes the integration of the three aspects even more important. The orchestration of heterogeneous resources and services at all levels becomes enormously complex, which may be efficiently managed by advanced AI techniques. In this context, the MEO becomes a critical element, especially from the security and dependability points of view. In the future perspective, the ubiquitous pervasive intelligence will help to manage the complexity of the 5G-MEC systems towards the 6G.

REFERENCES

[1] *IMT Vision—Framework and Overall Objectives of the Future Development of IMT for 2020 and Beyond*, document ITU-R M.2083-0, ITU-R, Sep. 2015.

[2] 6G Flagship. (2020). *6G White Paper Edge Intelligence-6G Research Visions*. [Online]. Available: https://www.6gchannel.com/items/6g-white-paper-edgeintelligence/

[3] *ETSI Industry Specification Group (ISG) on Multi-Access Edge Computing (MEC)*. Accessed: Apr. 13, 2021. [Online]. Available: https://www.etsi.org/committee/mec

[4] M. Satyanarayanan, P. Bahl, R. Caceres, and N. Davies, "The case for VM-based cloudlets in mobile computing," *IEEE Pervasive Comput.*, vol. 8, no. 4, pp. 14–23, Oct. 2009.

[5] *OpenFog Consortium*. Accessed: Apr. 13, 2021. [Online]. Available: http://www.openfogconsortium.org/

[6] *Multi-Access Edge Computing (MEC); Terminology*, Standard GS MEC 001, Version 3.1.1, ETSI, Jan. 2022.

[7] *Multi-Access Edge Computing (MEC); Framework and Reference Architecture*, Standard GS MEC 003, Version 3.1.1, Mar. 2022.

[8] D. Sabella, V. Sukhomlinov, L. Trang, S. Kekki, P. Paglierani, R. Rossbach, X. Li, Y. Fang, D. Druta, F. Giust, L. Cominardi, W. Featherstone, B. Pike, and S. Hadad, "Developing software for multi-access edge computing," ETSI, Sophia Antipolis, France, White Paper 20, 2019.
[9] *Multi-Access Edge Computing (MEC); Edge Platform Application Enablement*, Standard GS MEC 011, Version 3.1.1, Sep. 2022.
[10] *Cloud Edge Computing: Beyond the Data Center*, OpenStack, Austin, TX, USA, 2018.
[11] *Multi-Access Edge Computing (MEC); Study on Inter-MEC Systems and MEC-Cloud Systems Coordination*, Standard GR MEC 035, Version 3.1.1, ETSI, Jun. 2021.
[12] *Multi-Access Edge Computing (MEC); Use Cases and Requirements*, Standard GS MEC 002, Version 3.2.1, Apr. 2023.
[13] *Network Functions Virtualisation (NFV); Architectural Framework*, Standard GS NFV 002, Version 1.2.1, Dec. 2014.
[14] J. M. Halpern and C. Pignataro, *Service Function Chaining (SFC) Architecture*, document RFC 7665, Internet Engineering Task Force, Fremont, CA, USA, Oct. 2015.
[15] *Mobile Edge Computing (MEC); Deployment of Mobile Edge Computing in an NFV Environment*, Standard GR MEC 017, Version 1.1.1, ETSI, Feb. 2018.
[16] *Multi-Access Edge Computing (MEC); Study on MEC Support for Alternative Virtualization Technologies*, Standard GR MEC 027, Version 2.1.1, ETSI, Nov. 2019.
[17] G. Nencioni, R. G. Garroppo, A. J. Gonzalez, B. E. Helvik, and G. Procissi, "Orchestration and control in software-defined 5G networks: Research challenges," *Wireless Commun. Mobile Comput.*, vol. 2018, pp. 1–18, Aug. 2018.
[18] S. Kekki, W. Featherstone, Y. Fang, P. Kuure, A. Li, A. Ranjan, D. Purkayastha, F. Jiangping, D. Frydman, G. Verin, K.-W. Wen, K. Kim, R. Arora, A. Odgers, L. M. Contreras, and S. Scarpina, "MEC in 5G networks," ETSI, Sophia Antipolis, France, White Paper 28, Jun. 2018.
[19] *Multi-Access Edge Computing (MEC); MEC 5G Integration*, Standard GR MEC 031, Version 2.1.1, Oct. 2020.
[20] *Technical Specification Group Services and System Aspects; System Architecture for the 5G System (5GS); Stage 2 (Release 17)*, document TS 23.501, Version 17.0.0, 3GPP, Mar. 2021.
[21] F. Giust, G. Verin, K. Antevski, J. Chou, Y. Fang, W. Featherstone, F. Fontes, D. Frydman, A. Li, A. Manzalini, D. Purkayastha, D. Sabella, C. Wehner, K.-W. Wen, and Z. Zhou, "MEC deployments in 4G and evolution towards 5G," ETSI, Sophia Antipolis, France, White Paper 24, Feb. 2018.
[22] *Multi-Access Edge Computing (MEC); WLAN Information API*, Standard GS MEC 028, Version 2.3.1, ETSI, Jul. 2022.
[23] A. Reznik, L. M. C. Murillo, Y. Fang, W. Featherstone, M. Filippou, F. Fontes, F. Giust, Q. Huang, A. Li, C. Turyagyenda, C. Wehner, and Z. Zheng, "Cloud RAN and MEC: A perfect pairing," ETSI, Sophia Antipolis, France, White Paper 23, Feb. 2018.
[24] A. J. Gonzalez, J. Ordonez-Lucena, B. E. Helvik, G. Nencioni, M. Xie, D. R. Lopez, and P. Grønsund, "The isolation concept in the 5G network slicing," in *Proc. Eur. Conf. Netw. Commun. (EuCNC)*, Dubrovnik, Croatia, Jun. 2020, pp. 12–16.
[25] R. F. Olimid and G. Nencioni, "5G network slicing: A security overview," *IEEE Access*, vol. 8, pp. 99999–100009, 2020.
[26] *Multi-Access Edge Computing (MEC); Support for Network Slicing*, Standard GR MEC 024, Version 2.1.1, Nov. 2019.
[27] *Mobile Edge Computing; Market Acceleration; MEC Metrics Best Practice and Guidelines*, document GS MEC-IEG 006, Version 1.1.1, Jan. 2017.
[28] *Multi-Access Edge Computing (MEC); General Principles, Patterns and Common Aspects of MEC Service APIs*, Standard GS MEC 009, Version 3.2.1, Jul. 2022.
[29] *Mobile Edge Computing (MEC); Mobile Edge Management; Part 1: System, Host and Platform Management Management*, Standard GS MEC 010-1, Version 1.1.1, Oct. 2017.
[30] *Multi-Access Edge Computing (MEC); MEC Management; Part 2: Application Lifecycle, Rules and Requirements Management*, Standard GS MEC 010-2, Version 2.2.1, Feb. 2022.
[31] *Multi-Access Edge Computing (MEC); Radio Network Information API*, Standard GS MEC 012, Version 2.2.1, Feb. 2022.
[32] *Multi-Access Edge Computing (MEC); Location API*, Standard GS MEC 013, Version 2.2.1, Jan. 2022.
[33] *Mobile Edge Computing (MEC); UE Identity API*, Standard GS MEC 014, Version 2.1.1, Mar. 2021.
[34] *Multi-Access Edge Computing (MEC); Traffic Management APIs*, Standard GS MEC 015, Version 2.2.1, Dec. 2022.
[35] *Multi-Access Edge Computing (MEC); Device Application Interface*, Standard GS MEC 016, Version 2.2.1, Apr. 2020.
[36] *Multi-Access Edge Computing (MEC); Application Mobility Service API*, Standard GS MEC 021, Version 2.2.1, Feb. 2022.
[37] *Multi-Access Edge Computing (MEC); Support for Regulatory Requirements*, Standard GS MEC 026, Version 2.1.1, ETSI, Jan 2019.
[38] *Multi-Access Edge Computing (MEC); Fixed Access Information API*, Standard GS MEC 029, Version 2.2.1, Jan. 2022.
[39] *Multi-Access Edge Computing (MEC); V2X Information Service API*, Standard GS MEC 030, Version 3.1.1, Mar. 2023.
[40] *Multi-Access Edge Computing (MEC); API Conformance Test Specification; Part 1: Test Requirements and Implementation Conformance Statement (ICS)*, Standard GS MEC-DEC 032-1, Version 3.1.1, Apr. 2022.
[41] *Multi-Access Edge Computing (MEC); API Conformance Test Specification; Part 2: Test Purposes (TP)*, Standard GS MEC-DEC 032-2, Version 3.1.1, Apr. 2022.
[42] *Multi-Access Edge Computing (MEC); IoT API*, Standard GS MEC 033, Version 3.1.1, Dec. 2022.
[43] *Multi-Access Edge Computing (MEC); Application Package Format and Descriptor Specification*, Standard GS MEC 037, Version 3.1.1, Mar. 2023.
[44] *Multi-Access Edge Computing (MEC); Federation Enablement APIs*, Standard GS MEC 040, Version 3.1.1, Feb. 2023.
[45] *Mobile Edge Computing (MEC); End to End Mobility Aspects*, Standard GR MEC 018, Version 1.1.1, Oct. 2017.
[46] *Multi-Access Edge Computing (MEC); Study on MEC Support for V2X Use Cases*, Standard GR MEC 022, Version 2.1.1, Sep. 2018.
[47] *Multi-Access Edge Computing (MEC); MEC Testing Framework*, Standard GR MEC-DEC 025, Version 2.1.1, Jun. 2019.
[48] *Multi-Access Edge Computing (MEC); MEC in Park Enterprises Deployment Scenario*, Standard GR MEC 038, Version 3.1.1, Nov. 2022.
[49] *Multi-Access Edge Computing (MEC); Guidelines on Interoperability Testing*, Standard GR MEC-DEC 042, Version 3.1.1, Nov. 2022.
[50] A. Reznik, A. Sulistio, A. Artemenko, Y. Fang, D. Frydman, F. Giust, H. Lv, S. U. Sheikh, Y. Yu, and Z. Zheng, "MEC in an enterprise setting: A solution outline," ETSI, Sophia Antipolis, France, White Paper 30, Sep. 2018.
[51] Chairmen of ISG ENI, MEC, NFV and ZSM, "Network transformation; (orchestration, network and service management framework)," ETSI, Sophia Antipolis, France, White Paper 32, Oct. 2019.
[52] L. Frost, T. B. Meriem, J. M. Bonifacio, S. Cadzow, F. da Silva, M. Essa, R. Forbes, P. Marchese, M.-P. Odini, N. Sprecher, C. Toche, and S. Wood, "Artificial intelligence and future directions for ETSI," ETSI, Sophia Antipolis, France, White Paper 34, Jun. 2020.
[53] N. Sprecher et al., "Harmonizing standards for edge computing—A synergized architecture leveraging ETSI ISG MEC and 3GPP specifications," ETSI, Sophia Antipolis, France, White Paper 36, Jul. 2020.
[54] M. Suzuki, T. Miyasaka, D. Purkayastha, Y. Fang, Q. Huang, J. Zhu, B. Burla, X. Tong, D. Druta, J. Shen, H. Ding, G. Song, M. Angaroni, and V. Costa, "Enhanced DNS support towards distributed MEC environment," ETSI, Sophia Antipolis, France, White Paper 39, Sep. 2020.
[55] D. Sabella, A. Reznik, K. R. Nayak, D. Lopez, F. Li, U. Kleber, A. Leadbeater, K. Maloor, S. B. M. Baskaran, L. Cominardi, C. Costa, F. Granelli, V. Gazis, F. Ennesser, and X. Gu, "MEC security: Status of standards support and future evolutions," ETSI, Sophia Antipolis, France, White Paper 46, Sep. 2022.
[56] M. Suzuki, T. Joh, H. Lee, W. Featherstone, N. Sprecher, D. Sabella, N. Oliver, S. Shailendra, F. Granelli, C. Costa, L. Chen, H. Nieminen, O. Berzin, and F. Naim, "MEC federation: deployment considerations," ETSI, Sophia Antipolis, France, White Paper 49, Jun. 2022.
[57] The 5G Infrastructure Association (5G-IA). (Jun. 2021). *European Vision for the 6G Network Ecosystem*. [Online]. Available: https://5g-ppp.eu/wp-content/uploads/2021/06/WhitePaper-6G-Europe.pdf
[58] K. B. Letaief, Y. Shi, J. Lu, and J. Lu, "Edge artificial intelligence for 6G: Vision, enabling technologies, and applications," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 1, pp. 5–36, Jan. 2022.
[59] J. Hoydis, F. A. Aoudia, A. Valcarce, and H. Viswanathan, "Toward a 6G AI-native air interface," *IEEE Commun. Mag.*, vol. 59, no. 5, pp. 76–81, May 2021.

[60] Y. Yu, "Mobile edge computing towards 5G: Vision, recent progress, and open challenges," *China Commun.*, vol. 13, no. 2, pp. 89–99, 2016. [Online]. Available: https://ieeexplore.ieee.org/document/7405725/

[61] B. Hibat Allah and I. Abdellah, "MEC towards 5G: A survey of concepts, use cases, location tradeoffs," *Trans. Mach. Learn. Artif. Intell.*, vol. 5, no. 4, pp. 1–10, Aug. 2017. [Online]. Available: http://www.scholarpublishing.org/index.php/TMLAI/article/view/3215

[62] T. Taleb, K. Samdanis, B. Mada, H. Flinck, S. Dutta, and D. Sabella, "On multi-access edge computing: A survey of the emerging 5G network edge cloud architecture and orchestration," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 3, pp. 1657–1681, 3rd Quart., 2017. [Online]. Available: http://ieeexplore.ieee.org/document/7931566/

[63] S. Wang, X. Zhang, Y. Zhang, L. Wang, J. Yang, and W. Wang, "A survey on mobile edge networks: Convergence of computing, caching and communications," *IEEE Access*, vol. 5, pp. 6757–6779, 2017. [Online]. Available: http://ieeexplore.ieee.org/document/7883826/

[64] Y. Mao, C. You, J. Zhang, K. Huang, and K. B. Letaief, "A survey on mobile edge computing: The communication perspective," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 4, pp. 2322–2358, 4th Quart., 2017. [Online]. Available: http://ieeexplore.ieee.org/document/8016573/

[65] P. Mach and Z. Becvar, "Mobile edge computing: A survey on architecture and computation offloading," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 3, pp. 1628–1656, 3rd Quart., 2017. [Online]. Available: http://ieeexplore.ieee.org/document/7879258/

[66] S. Shahzadi, M. Iqbal, T. Dagiuklas, and Z. U. Qayyum, "Multi-access edge computing: Open issues, challenges and future perspectives," *J. Cloud Comput.*, vol. 6, no. 1, p. 30, Dec. 2017.

[67] S. N. Shirazi, A. Gouglidis, A. Farshad, and D. Hutchison, "The extended cloud: Review and analysis of mobile edge computing and fog from a security and resilience perspective," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 11, pp. 2586–2595, Nov. 2017.

[68] K. Bilal, O. Khalid, A. Erbad, and S. U. Khan, "Potentials, trends, and prospects in edge technologies: Fog, cloudlet, mobile edge, and micro data centers," *Comput. Netw.*, vol. 130, pp. 94–120, Jan. 2018.

[69] K. Peng, V. C. M. Leung, X. Xu, L. Zheng, J. Wang, and Q. Huang, "A survey on mobile edge computing: Focusing on service adoption and provision," *Wireless Commun. Mobile Comput.*, vol. 2018, pp. 1–16, Oct. 2018. [Online]. Available: https://www.hindawi.com/journals/wcmc/2018/8267838/

[70] R. Roman, J. Lopez, and M. Mambo, "Mobile edge computing, Fog et al.: A survey and analysis of security threats and challenges," *Future Gener. Comput. Syst.*, vol 78, pp. 680–698, Jan. 2018.

[71] P. Porambage, J. Okwuibe, M. Liyanage, M. Ylianttila, and T. Taleb, "Survey on multi-access edge computing for Internet of Things realization," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 4, pp. 2961–2991, 4th Quart., 2018. [Online]. Available: https://ieeexplore.ieee.org/document/8391395/

[72] H. Tanaka, M. Yoshida, K. Mori, and N. Takahashi, "Multi-access edge computing: A survey," *J. Inf. Process.*, vol. 26, pp. 87–97, Feb. 2018. [Online]. Available: https://www.jstage.jst.go.jp/article/ipsjjip/26/0/2687/article

[73] N. Abbas, Y. Zhang, A. Taherkordi, and T. Skeie, "Mobile edge computing: A survey," *IEEE Internet Things J.*, vol. 5, no. 1, pp. 450–465, Feb. 2018. [Online]. Available: http://ieeexplore.ieee.org/document/8030322/

[74] W. Z. Khan, E. Ahmed, S. Hakak, I. Yaqoob, and A. Ahmed, "Edge computing: A survey," *Future Gener. Comput. Syst.*, vol. 97, pp. 219–235, Aug. 2019.

[75] A. Yousefpour, C. Fung, T. Nguyen, K. Kadiyala, F. Jalali, A. Niakanlahiji, J. Kong, and J. P. Jue, "All one needs to know about fog computing and related edge computing paradigms: A complete survey," *J. Syst. Archit.*, vol. 98, pp. 289–330, Sep. 2019.

[76] P. Ranaweera, A. D. Jurcut, and M. Liyanage, "Realizing multi-access edge computing feasibility: Security perspective," in *Proc. IEEE Conf. Standards Commun. Netw. (CSCN)*, Oct. 2019, pp. 1–7.

[77] Z. Li, X. Zhou, and Y. Qin, "A survey of mobile edge computing in the industrial internet," in *Proc. 7th Int. Conf. Inf., Commun. Netw. (ICICN)*, Macao, China, Apr. 2019, pp. 94–98. [Online]. Available: https://ieeexplore.ieee.org/document/8834959/

[78] Q. Pham, F. Fang, V. N. Ha, M. J. Piran, M. Le, L. B. Le, W. Hwang, and Z. Ding, "A survey of multi-access edge computing in 5G and beyond: Fundamentals, technology integration, and state-of-the-art," *IEEE Access*, vol. 8, pp. 116974–117017, 2020.

[79] P. Habibi, M. Farhoudi, S. Kazemian, S. Khorsandi, and A. Leon-Garcia, "Fog computing: A comprehensive architectural survey," *IEEE Access*, vol. 8, pp. 69105–69133, 2020. [Online]. Available: https://ieeexplore.ieee.org/document/9046806/

[80] A. Filali, A. Abouaomar, S. Cherkaoui, A. Kobbane, and M. Guizani, "Multi-access edge computing: A survey," *IEEE Access*, vol. 8, pp. 197017–197046, 2020.

[81] F. Vhora and J. Gandhi, "A comprehensive survey on mobile edge computing: Challenges, tools, applications," in *Proc. 4th Int. Conf. Comput. Methodol. Commun. (ICCMC)*, Erode, India, Mar. 2020, pp. 49–55. [Online]. Available: https://ieeexplore.ieee.org/document/9076528/

[82] P. Ranaweera, A. D. Jurcut, and M. Liyanage, "Survey on multi-access edge computing security and privacy," *IEEE Commun. Surveys Tuts.*, vol. 23, no. 2, pp. 1078–1124, 2nd Quart., 2021.

[83] S. D. A. Shah, M. A. Gregory, and S. Li, "Cloud-native network slicing using software defined networking based multi-access edge computing: A survey," *IEEE Access*, vol. 9, pp. 10903–10924, 2021.

[84] P. Cruz, N. Achir, and A. C. Viana, "On the edge of the deployment: A survey on multi-access edge computing," *ACM Comput. Surv.*, vol. 55, no. 5, pp. 1–34, May 2023, doi: 10.1145/3529758.

[85] A. Sarah, G. Nencioni, and M. M. I. Khan, "Resource allocation in multi-access edge computing for 5G-and-beyond networks," *Comput. Netw.*, vol. 227, May 2023, Art. no. 109720. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1389128623001652

[86] H. Madsen, B. Burtschy, G. Albeanu, and F. Popentiu-Vladicescu, "Reliability in the utility computing era: Towards reliable fog computing," in *Proc. 20th Int. Conf. Syst., Signals Image Process. (IWSSIP)*, Jul. 2013, pp. 43–46.

[87] P. Hu, S. Dhelim, H. Ning, and T. Qiu, "Survey on fog computing: Architecture, key technologies, applications and open issues," *J. Netw. Comput. Appl.*, vol. 98, pp. 27–42, Nov. 2017. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S1084804517302953

[88] R. Mahmud, R. Kotagiri, and R. Buyya, "Fog computing: A taxonomy, survey and future directions," 2016, *arXiv:1611.05539*.

[89] Z. Bakhshi and G. Rodriguez-Navas, "A preliminary roadmap for dependability research in fog computing," *ACM SIGBED Rev.*, vol. 16, no. 4, pp. 14–19, Jan. 2020, doi: 10.1145/3378408.3378410.

[90] J. M. G. Sánchez, N. Jörgensen, M. Törngren, R. Inam, A. Berezovskyi, L. Feng, E. Fersman, M. R. Ramli, and K. Tan, "Edge computing for cyber-physical systems: A systematic mapping study emphasizing trustworthiness," *ACM Trans. Cyber-Phys. Syst.*, vol. 6, no. 3, pp. 1–28, Sep. 2022, doi: 10.1145/3539662.

[91] P. G. Lopez, A. Montresor, D. Epema, A. Datta, T. Higashino, A. Iamnitchi, M. Barcellos, P. Felber, and E. Riviere, "Edge-centric computing: Vision and challenges," *ACM SIGCOMM Comput. Commun. Rev.*, vol. 45, no. 5, pp. 37–42, Sep. 2015.

[92] X. Huang, R. Yu, J. Kang, Y. He, and Y. Zhang, "Exploring mobile edge computing for 5G-enabled software defined vehicular networks," *IEEE Wireless Commun.*, vol. 24, no. 6, pp. 55–63, Dec. 2017. [Online]. Available: http://ieeexplore.ieee.org/document/8246847/

[93] W. Yu, F. Liang, X. He, W. G. Hatcher, C. Lu, J. Lin, and X. Yang, "A survey on the edge computing for the Internet of Things," *IEEE Access*, vol. 6, pp. 6900–6919, 2018. [Online]. Available: http://ieeexplore.ieee.org/document/8123913/

[94] M. Du, K. Wang, Y. Chen, X. Wang, and Y. Sun, "Big data privacy preserving in multi-access edge computing for heterogeneous Internet of Things," *IEEE Commun. Mag.*, vol. 56, no. 8, pp. 62–67, Aug. 2018.

[95] J. Ni, X. Lin, and X. S. Shen, "Toward edge-assisted Internet of Things: From security and efficiency perspectives," *IEEE Netw.*, vol. 33, no. 2, pp. 50–57, Mar. 2019.

[96] S. Bagchi, M.-B. Siddiqui, P. Wood, and H. Zhang, "Dependability in edge computing," *Commun. ACM*, vol. 63, no. 1, pp. 58–66, Dec. 2019.

[97] M. S. Elbamby, C. Perfecto, C. Liu, J. Park, S. Samarakoon, X. Chen, and M. Bennis, "Wireless edge computing with latency and reliability guarantees," *Proc. IEEE*, vol. 107, no. 8, pp. 1717–1737, Aug. 2019.

[98] M. Yahuza, M. Y. I. B. Idris, A. W. B. A. Wahab, A. T. S. Ho, S. Khan, S. N. B. Musa, and A. Z. B. Taha, "Systematic review on security and privacy requirements in edge computing: State of the art and future research opportunities," *IEEE Access*, vol. 8, pp. 76541–76567, 2020.

[99] L. Liu, C. Chen, Q. Pei, S. Maharjan, and Y. Zhang, "Vehicular edge computing and networking: A survey," *Mobile Netw. Appl.*, vol. 26, no. 3, pp. 1145–1168, Jul. 2020, doi: 10.1007/s11036-020-01624-1.

[100] H. Xue, F. Dai, G. Liu, P. Cao, and B. Huang, "Edge computing: A systematic mapping study," in *Proc. IEEE Int. Conf. Dependable, Autonomic Secure Comput., Int. Conf Pervasive Intell. Comput., Int. Conf. Cloud Big Data Comput., Int. Conf. Cyber Sci. Technol. Congr. (DASC/PiCom/CBDCom/CyberSciTech)*, Oct. 2021, pp. 507–514.
[101] H. Hua, Y. Li, T. Wang, N. Dong, W. Li, and J. Cao, "Edge computing with artificial intelligence: A machine learning perspective," *ACM Comput. Surv.*, vol. 55, no. 9, pp. 1–35, Sep. 2023, doi: 10.1145/3555802.
[102] R. Khan, P. Kumar, D. N. K. Jayakody, and M. Liyanage, "A survey on security and privacy of 5G technologies: Potential solutions, recent advancements, and future directions," *IEEE Commun. Survey Tuts.*, vol. 22, no. 1, pp. 196–248, 1st Quart., 2020.
[103] *Glossary*, Comput. Secur. Res. Center (CSRC), Nat. Inst. Standards Technol. (NIST), Gaithersburg, MD, USA, 2021.
[104] H. C. Van Tilborg and S. Jajodia, *Encyclopedia of Cryptography and Security*. New York, NY, USA: Springer, 2014.
[105] A. J. Menezes, P. C. Van Oorschot, and S. A. Vanstone, *Handbook of Applied Cryptography*. Boca Raton, FL, USA: CRC Press, 2018.
[106] M. Liyanage, A. Gurtov, and M. Ylianttila, *Software Defined Mobile Networks (SDMN): Beyond LTE Network Architecture*. Hoboken, NJ, USA: Wiley, 2015.
[107] M. Liyanage, A. Braeken, A. D. Jurcut, M. Ylianttila, and A. Gurtov, "Secure communication channel architecture for software defined mobile networks," *Comput. Netw.*, vol. 114, pp. 32–50, Feb. 2017.
[108] J. McCumber, "Information systems security: A comprehensive model," in *Proc. 14th Nat. Comput. Secur. Conf.*, 1991, pp. 328–337.
[109] *5G Security White Paper—Security Makes 5G Go Further*, ZTE, Shenzhen, China, May 2019.
[110] *5G Security—Package 3: Mobile Edge Computing/Low Latency/Consistent User Experience*, NGMN, Frankfurt, Germany, Feb. 2018.
[111] *Network Functions Virtualisation (NFV) Release 4; Protocols and Data Models; VNF Package and PNFD Archive Specification*, Standard GS NFV-SOL 004, Version 4.4.1, ETSI, Mar. 2023.
[112] I. Ahmad, T. Kumar, M. Liyanage, J. Okwuibe, M. Ylianttila, and A. Gurtov, "Overview of 5G security challenges and solutions," *IEEE Commun. Standards Mag.*, vol. 2, no. 1, pp. 36–43, Mar. 2018.
[113] J. Zhang, B. Chen, Y. Zhao, X. Cheng, and F. Hu, "Data security and privacy-preserving in edge computing paradigm: Survey and open issues," *IEEE Access*, vol. 6, pp. 18209–18237, 2018.
[114] R. Rapuzzi and M. Repetto, "Building situational awareness for network threats in fog/edge computing: Emerging paradigms beyond the security perimeter model," *Future Gener. Comput. Syst.*, vol. 85, pp. 235–249, Aug. 2018.
[115] Y. Xiao, Y. Jia, C. Liu, X. Cheng, J. Yu, and W. Lv, "Edge computing security: State of the art and challenges," *Proc. IEEE*, vol. 107, no. 8, pp. 1608–1631, Aug. 2019.
[116] M. Caprolu, R. Di Pietro, F. Lombardi, and S. Raponi, "Edge computing perspectives: Architectures, technologies, and open security issues," in *Proc. IEEE Int. Conf. Edge Comput. (EDGE)*, Jul. 2019, pp. 116–123.
[117] D. Liu, Z. Yan, W. Ding, and M. Atiquzzaman, "A survey on secure data analytics in edge computing," *IEEE Internet Things J.*, vol. 6, no. 3, pp. 4946–4967, Jun. 2019.
[118] K. Sha, T. A. Yang, W. Wei, and S. Davari, "A survey of edge computing-based designs for IoT security," *Digit. Commun. Netw.*, vol. 6, no. 2, pp. 195–202, May 2020.
[119] A. Alwarafy, K. A. Al-Thelaya, M. Abdallah, J. Schneider, and M. Hamdi, "A survey on security and privacy issues in edge-computing-assisted Internet of Things," *IEEE Internet Things J.*, vol. 8, no. 6, pp. 4004–4022, Mar. 2021.
[120] A. M. Alwakeel, "An overview of fog computing and edge computing security and privacy issues," *Sensors*, vol. 21, no. 24, p. 8226, Dec. 2021.
[121] T. W. Nowak, M. Sepczuk, Z. Kotulski, W. Niewolski, R. Artych, K. Bocianiak, T. Osko, and J. Wary, "Verticals in 5G MEC-use cases and security challenges," *IEEE Access*, vol. 9, pp. 87251–87298, 2021.
[122] B. Ali, M. A. Gregory, and S. Li, "Multi-access edge computing architecture, data security and privacy: A review," *IEEE Access*, vol. 9, pp. 18706–18721, 2021.
[123] X. Tan, H. Li, L. Wang, and Z. Xu, "Global orchestration of cooperative defense against DDoS attacks for MEC," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Apr. 2019, pp. 1–6.
[124] H. Li and L. Wang, "Online orchestration of cooperative defense against DDoS attacks for 5G MEC," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Apr. 2018, pp. 1–6.
[125] A. Serrano Mamolar, Z. Pervez, J. M. Alcaraz Calero, and A. M. Khattak, "Towards the transversal detection of DDoS network attacks in 5G multi-tenant overlay networks," *Comput. Secur.*, vol. 79, pp. 132–147, Nov. 2018.
[126] P. Krishnan, S. Duttagupta, and K. Achuthan, "SDNFV based threat monitoring and security framework for multi-access edge computing infrastructure," *Mobile Netw. Appl.*, vol. 24, no. 6, pp. 1896–1923, Dec. 2019.
[127] L. Xiao, X. Wan, C. Dai, X. Du, X. Chen, and M. Guizani, "Security in mobile edge caching with reinforcement learning," *IEEE Wireless Commun.*, vol. 25, no. 3, pp. 116–122, Jun. 2018.
[128] A. Abeshu and N. Chilamkurti, "Deep learning: The frontier for distributed attack detection in Fog-to-Things computing," *IEEE Commun. Mag.*, vol. 56, no. 2, pp. 169–175, Feb. 2018.
[129] S. Singh, R. Sulthana, T. Shewale, V. Chamola, A. Benslimane, and B. Sikdar, "Machine-learning-assisted security and privacy provisioning for edge computing: A survey," *IEEE Internet Things J.*, vol. 9, no. 1, pp. 236–260, Jan. 2022.
[130] D. Thembelihle, M. Rossi, and D. Munaretto, "Softwarization of mobile network functions towards agile and energy efficient 5G architectures: A survey," *Wireless Commun. Mobile Comput.*, vol. 2017, pp. 1–21, Nov. 2017.
[131] D. He, S. Chan, and M. Guizani, "Security in the Internet of Things supported by mobile edge computing," *IEEE Commun. Mag.*, vol. 56, no. 8, pp. 56–61, Aug. 2018.
[132] J. Okwuibe, M. Liyanage, I. Ahmad, and M. Ylianttila, "Cloud and MEC security," in *A Comprehensive Guide to 5G Security*. Hoboken, NJ, USA: Wiley, 2018, p. 373.
[133] *Research Topics-ENISA Threat Landscape*, ENISA, Athens, Greece, Oct. 2020.
[134] I. Farris, J. B. Bernabé, N. Toumi, D. Garcia-Carrillo, T. Taleb, A. Skarmeta, and B. Sahlin, "Towards provisioning of SDN/NFV-based security enablers for integrated protection of IoT systems," in *Proc. IEEE Conf. Standards Commun. Netw. (CSCN)*. Sep. 2017, pp. 169–174.
[135] A. M. Zarca, J. B. Bernabe, R. Trapero, D. Rivera, J. Villalobos, A. Skarmeta, S. Bianchi, A. Zafeiropoulos, and P. Gouvas, "Security management architecture for NFV/SDN-aware IoT systems," *IEEE Internet Things J.*, vol. 6, no. 5, pp. 8005–8020, Oct. 2019.
[136] I. Petri, O. F. Rana, Y. Rezgui, and G. C. Silaghi, "Trust modelling and analysis in peer-to-peer clouds," *Int. J. Cloud Comput.*, vol. 1, nos. 2–3, pp. 221–239, 2012.
[137] Y. He, F. R. Yu, N. Zhao, and H. Yin, "Secure social networks in 5G systems with mobile edge computing, caching, and device-to-device communications," *IEEE Wireless Commun.*, vol. 25, no. 3, pp. 103–109, Jun. 2018.
[138] T. Bai, J. Wang, Y. Ren, and L. Hanzo, "Energy-efficient computation offloading for secure UAV-edge-computing systems," *IEEE Trans. Veh. Technol.*, vol. 68, no. 6, pp. 6074–6087, Jun. 2019.
[139] J. Xu and J. Yao, "Exploiting physical-layer security for multiuser multicarrier computation offloading," *IEEE Wireless Commun. Lett.*, vol. 8, no. 1, pp. 9–12, Feb. 2019.
[140] A. Mtibaa, K. Harras, and H. Alnuweiri, "Friend or foe? Detecting and isolating malicious nodes in mobile edge computing platforms," in *Proc. IEEE 7th Int. Conf. Cloud Comput. Technol. Sci. (CloudCom)*, Nov. 2015, pp. 42–49.
[141] W. Y. B. Lim, N. C. Luong, D. T. Hoang, Y. Jiao, Y. Liang, Q. Yang, D. Niyato, and C. Miao, "Federated learning in mobile edge networks: A comprehensive survey," *IEEE Commun. Surveys Tuts.*, vol. 22, no. 3, pp. 2031–2063, 3rd Quart., 2020.
[142] T. Li, A. K. Sahu, A. Talwalkar, and V. Smith, "Federated learning: Challenges, methods, and future directions," *IEEE Signal Process. Mag.*, vol. 37, no. 3, pp. 50–60, May 2020.
[143] *Virtual Machine Introspection*, LibVMI, Los Altos, CA, USA, 2021.
[144] M. A. Ajay Kumara and C. D. Jaidhar, "Leveraging virtual machine introspection with memory forensics to detect and characterize unknown malware using machine learning techniques at hypervisor," *Digit. Invest.*, vol. 23, pp. 99–123, Dec. 2017.
[145] S. F. Mjolsnes and R. F. Olimid, "Private identification of subscribers in mobile networks: Status and challenges," *IEEE Commun. Mag.*, vol. 57, no. 9, pp. 138–144, Sep. 2019.
[146] *CYBER: Application of Attribute Based Encryption (ABE) for PII and Personal Data Protection on IoT Devices, WLAN, Cloud and Mobile Services-High Level Requirements*, Standard TS 103 458, Version 1.1.1, ETSI, Jun. 2018.

[147] Z. Liao, X. Pang, J. Zhang, B. Xiong, and J. Wang, "Blockchain on security and forensics management in edge computing for IoT: A comprehensive survey," *IEEE Trans. Netw. Service Manage.*, vol. 19, no. 2, pp. 1159–1175, Jun. 2022.

[148] A. V. Rivera, A. Refaey, and E. Hossain, "A blockchain framework for secure task sharing in multi-access edge computing," *IEEE Netw.*, vol. 35, no. 3, pp. 176–183, May 2021.

[149] K. Wüst and A. Gervais, "Do you need a blockchain?" in *Proc. Crypto Valley Conf. Blockchain Technol. (CVCBT)*, Jun. 2018, pp. 45–54.

[150] *Industry Specification Group (ISG) Permissioned Distributed Ledger (PDL)*, ETSI, Sophia Antipolis, France, 2022.

[151] A. Avizienis, J.-C. Laprie, B. Randell, and C. Landwehr, "Basic concepts and taxonomy of dependable and secure computing," *IEEE Trans. Depend. Sec. Comput.*, vol. 1, no. 1, pp. 11–33, Mar. 2004.

[152] *ResliNets Wiki*. Accessed: Jan. 22, 2021. [Online]. Available: https://resilinets.org/

[153] K. S. Trivedi, D. Seong Kim, A. Roy, and D. Medhi, "Dependability and security models," in *Proc. 7th Int. Workshop Design Reliable Commun. Netw.*, Oct. 2009, pp. 11–20.

[154] *Toward Fully Connected Vehicles: Edge Computing for Advanced Automotive Communications*, 5GAA Automot. Assoc., Munich, Germany, Dec 2017.

[155] *Digital Cellular Telecommunications System (Phase 2+) (GSM); Universal Mobile Telecommunications System (UMTS); LTE; Telecommunication management; Fault Management; Part 2: Alarm Integration Reference Point (IRP): (IS) (Release 16)*, document TS 132 111-2, Version 16.0.0, 3GPP, Aug. 2020.

[156] *Digital Cellular Telecommunications System (Phase 2+) (GSM); Universal Mobile Telecommunications System (UMTS); LTE; Telecommunication management;(NL) Integration Reference Point (IRP); Information Service (IS), (Release 16)*, document TS 132 332, Version 16.0.0, 3GPP, Aug. 2020.

[157] J. Hao, K. Ye, and C.-Z. Xu, "Live migration of virtual machines in openstack: A perspective from reliability evaluation," in *Proc. Int. Conf. Cloud Comput.* New York, NY, USA: Springer, 2019, pp. 99–113.

[158] L. Yala, P. A. Frangoudis, and A. Ksentini, "Latency and availability driven VNF placement in a MEC-NFV environment," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2018, pp. 1–7.

[159] H. D. Chantre and N. L. S. D. Fonseca, "The location problem for the provisioning of protected slices in NFV-based MEC infrastructure," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 7, pp. 1505–1514, Jul. 2020.

[160] B. Li, Q. He, G. Cui, X. Xia, F. Chen, H. Jin, and Y. Yang, "READ: Robustness-oriented edge application deployment in edge computing environment," *IEEE Trans. Services Comput.*, vol. 15, no. 3, pp. 1746–1759, May 2022.

[161] F. Tonini, B. Khorsandi, E. Amato, and C. Raffaelli, "Scalable edge computing deployment for reliable service provisioning in vehicular networks," *J. Sens. Actuator Netw.*, vol. 8, no. 4, p. 51, Oct. 2019. [Online]. Available: https://www.mdpi.com/2224-2708/8/4/51

[162] H. Huang and S. Guo, "Proactive failure recovery for NFV in distributed edge computing," *IEEE Commun. Mag.*, vol. 57, no. 5, pp. 131–137, May 2019.

[163] D. Satria, D. Park, and M. Jo, "Recovery for overloaded mobile edge computing," *Future Gener. Comput. Syst.*, vol. 70, pp. 138–147, May 2017. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0167739X16302096

[164] C. Chen, M. Won, R. Stoleru, and G. G. Xie, "Energy-efficient fault-tolerant data storage and processing in mobile cloud," *IEEE Trans. Cloud Comput.*, vol. 3, no. 1, pp. 28–41, Jan. 2015.

[165] C. Chen, R. Stoleru, and G. G. Xie, "Energy-efficient and fault-tolerant mobile cloud storage," in *Proc. 5th IEEE Int. Conf. Cloud Netw. (Cloudnet)*, Oct. 2016, pp. 51–57.

[166] C.-F. Liu, M. Bennis, and H. V. Poor, "Latency and reliability-aware task offloading and resource allocation for mobile edge computing," in *Proc. IEEE Globecom Workshops (GC Wkshps)*, Dec. 2017, pp. 1–7.

[167] H. Liu, L. Cao, T. Pei, Q. Deng, and J. Zhu, "A fast algorithm for energy-saving offloading with reliability and latency requirements in multi-access edge computing," *IEEE Access*, vol. 8, pp. 151–161, 2020.

[168] M. Merluzzi, N. di Pietro, P. Di Lorenzo, E. C. Strinati, and S. Barbarossa, "Network energy efficient mobile edge computing with reliability guarantees," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2019, pp. 1–6.

[169] M. Merluzzi, P. D. Lorenzo, S. Barbarossa, and V. Frascolla, "Dynamic computation offloading in multi-access edge computing via ultra-reliable and low-latency communications," *IEEE Trans. Signal Inf. Process. Netw.*, vol. 6, pp. 342–356, 2020.

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

[171] J. Liu and Q. Zhang, "Offloading schemes in mobile edge computing for ultra-reliable low latency communications," *IEEE Access*, vol. 6, pp. 12825–12837, 2018.

[172] P. A. Apostolopoulos, E. E. Tsiropoulou, and S. Papavassiliou, "Risk-aware data offloading in multi-server multi-access edge computing environment," *IEEE/ACM Trans. Netw.*, vol. 28, no. 3, pp. 1405–1418, Jun. 2020.

[173] A. J. Gonzalez, G. Nencioni, A. Kamisinski, B. E. Helvik, and P. E. Heegaard, "Dependability of the NFV orchestrator: State of the art and research challenges," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 4, pp. 3307–3329, 4th Quart., 2018.

[174] A. J. Gonzalez, G. Nencioni, B. E. Helvik, and A. Kamisinski, "A fault-tolerant and consistent SDN controller," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2016, pp. 1–6.

[175] S.-C. Wang, W.-S. Hsiung, C.-F. Hsieh, and Y.-T. Tsai, "Reliability enhancement of edge computing paradigm using agreement," *Symmetry*, vol. 11, no. 2, p. 167, Feb. 2019. [Online]. Available: https://www.mdpi.com/2073-8994/11/2/167

[176] G. Nencioni, B. E. Helvik, and P. E. Heegaard, "Including failure correlation in availability modeling of a software-defined backbone network," *IEEE Trans. Netw. Service Manage.*, vol. 14, no. 4, pp. 1032–1045, Dec. 2017.

[177] B. Tola, G. Nencioni, and B. E. Helvik, "Network-aware availability modeling of an end-to-end NFV-enabled service," *IEEE Trans. Netw. Service Manage.*, vol. 16, no. 4, pp. 1389–1403, Dec. 2019.

[178] J. Liang, B. Ma, S. Ali, and J. Huang, "Model-based evaluation and optimization of dependability for edge computing systems," in *Collaborative Computing: Networking, Applications and Worksharing*, H. Gao and X. Wang, Eds. Cham, Switzerland: Springer, 2021, pp. 728–747.

[179] P. Maciel, J. Dantas, C. Melo, P. Pereira, F. Oliveira, J. Araujo, and R. Matos, "A survey on reliability and availability modeling of edge, fog, and cloud computing," *J. Reliable Intell. Environ.*, vol. 8, no. 3, pp. 227–245, Sep. 2022.

[180] H. Raei and N. Yazdani, "Performability analysis of cloudlet in mobile cloud computing," *Inf. Sci.*, vols. 388–389, pp. 99–117, May 2017. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0020025517301330

[181] J. Bai, X. Chang, F. Machida, L. Jiang, Z. Han, and K. S. Trivedi, "Impact of service function aging on the dependability for MEC service function chain," *IEEE Trans. Depend. Sec. Comput.*, early access, Feb. 14, 2022, doi: 10.1109/TDSC.2022.3150782.

[182] T. Pathirana and G. Nencioni, "Availability model of a 5G-MEC system," in *Proc. Int. Conf. Comput. Commun. Netw. (ICCCN)*, 2023, pp. 1–10.

[183] P. E. Heegaard, B. E. Helvik, G. Nencioni, and J. Wäfler, "Managed dependability in interacting systems," in *Principles of Performance and Reliability Modeling and Evaluation*. New York, NY, USA: Springer, 2016, pp. 197–226.

[184] K. Ray and A. Banerjee, "Prioritized fault recovery strategies for multi-access edge computing using probabilistic model checking," *IEEE Trans. Depend. Sec. Comput.*, vol. 20, no. 1, pp. 797–812, Jan. 2023.

[185] M. Le, Z. Song, Y. Kwon, and E. Tilevich, "Reliable and efficient mobile edge computing in highly dynamic and volatile environments," in *Proc. 2nd Int. Conf. Fog Mobile Edge Comput. (FMEC)*, May 2017, pp. 113–120.

[186] P. Kochovski and V. Stankovski, "Supporting smart construction with dependable edge computing infrastructures and applications," *Autom. Construct.*, vol. 85, pp. 182–192, Jan. 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0926580517304776

[187] D. Malas and A. Morton, *Basic Telephony SIP End-to-End Performance Metrics*, document RFC 6076, IETF, 2011.

[188] T. Hoßfeld, P. E. Heegaard, M. Varela, and S. Möller, "QoE beyond the MOS: An in-depth look at QoE via better metrics and their relation to MOS," *Qual. User Exp.*, vol. 1, no. 1, pp. 1–23, Dec. 2016.

[189] J. Moura and D. Hutchison, "Game theory for multi-access edge computing: Survey, use cases, and future trends," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 1, pp. 260–288, 1st Quart., 2019.

[190] T. Qiu, J. Chi, X. Zhou, Z. Ning, M. Atiquzzaman, and D. O. Wu, "Edge computing in industrial Internet of Things: Architecture, advances and challenges," *IEEE Commun. Surveys Tuts.*, vol. 22, no. 4, pp. 2462–2488, 4th Quart., 2020.

[191] F. Spinelli and V. Mancuso, "Toward enabled industrial verticals in 5G: A survey on MEC-based approaches to provisioning and flexibility," *IEEE Commun. Surveys Tuts.*, vol. 23, no. 1, pp. 596–630, 1st Quart., 2021.

[192] R. Mijumbi, J. Serrat, J. Gorricho, N. Bouten, F. De Turck, and R. Boutaba, "Network function virtualization: State-of-the-art and research challenges," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 1, pp. 236–262, 1st Quart., 2016.

[193] V. Aggarwal and B. Thangaraju, "Performance analysis of virtualisation technologies in NFV and edge deployments," in *Proc. IEEE Int. Conf. Electron., Comput. Commun. Technol. (CONECCT)*, Jul. 2020, pp. 1–5.

[194] R. Perez, P. Benedetti, M. Pergolesi, J. Garcia-Reinoso, A. Zabala, P. Serrano, M. Femminella, G. Reali, and A. Banchs, "A performance comparison of virtualization techniques to deploy a 5G monitoring platform," in *Proc. EUCNC, 6G Summit*, 2021, pp. 472–477.

[195] P. Valsamas, L. Mamatas, and L. M. Contreras, "A comparative evaluation of edge cloud virtualization technologies," *IEEE Trans. Netw. Service Manage.*, vol. 19, no. 2, pp. 1351–1365, Jun. 2022.

[196] O. Bentaleb, A. S. Z. Belloum, A. Sebaa, and A. El-Maouhab, "Containerization technologies: Taxonomies, applications and challenges," *J. Supercomput.*, vol. 78, no. 1, pp. 1144–1181, Jan. 2022.

[197] F. Manco, C. Lupu, F. Schmidt, J. Mendes, S. Kuenzer, S. Sati, K. Yasukata, C. Raiciu, and F. Huici, "My VM is lighter (and safer) than your container," in *Proc. 26th Symp. Operating Syst. Princ.* New York, NY, USA: Association for Computing Machinery, 2017, pp. 218–233, doi: 10.1145/3132747.3132763.

[198] Z. Yu, "The application of kata containers in Baidu AI cloud," Baidu, Beijing, China, Tech. Rep., 2019.

[199] R. Behravesh, E. Coronado, and R. Riggio, "Performance evaluation on virtualization technologies for NFV deployment in 5G networks," in *Proc. IEEE Conf. Netw. Softw. (NetSoft)*, Jun. 2019, pp. 24–29.

[200] D. Gupta, L. Cherkasova, R. Gardner, and A. Vahdat, "Enforcing performance isolation across virtual machines in Xen," in *Proc. ACM/IFIP/USENIX Int. Conf. Middleware*. Berlin, Germany: Springer-Verlag, 2006, pp. 342–362.

[201] L. Cherkasova, D. Gupta, and A. Vahdat, "Comparison of the three CPU schedulers in Xen," *ACM SIGMETRICS Perform. Eval. Rev.*, vol. 35, no. 2, pp. 42–51, Sep. 2007, doi: 10.1145/1330555.1330556.

[202] J. Li, S. Xue, W. Zhang, R. Ma, Z. Qi, and H. Guan, "When I/O interrupt becomes system bottleneck: Efficiency and scalability enhancement for SR-IOV network virtualization," *IEEE Trans. Cloud Comput.*, vol. 7, no. 4, pp. 1183–1196, Oct. 2019.

[203] Z. Kotulski, T. W. Nowak, M. Sepczuk, and M. A. Tunia, "5G networks: Types of isolation and their parameters in RAN and CN slices," *Comput. Netw.*, vol. 171, Apr. 2020, Art. no. 107135. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1389128619304797

[204] Y. Li and S. Wang, "An energy-aware edge server placement algorithm in mobile edge computing," in *Proc. IEEE Int. Conf. Edge Comput. (EDGE)*, Jul. 2018, pp. 66–73.

[205] S. Wang, Y. Zhao, J. Xu, J. Yuan, and C.-H. Hsu, "Edge server placement in mobile edge computing," *J. Parallel Distrib. Comput.*, vol. 127, pp. 160–168, May 2019. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0743731518304398

[206] S. K. Kasi, M. K. Kasi, K. Ali, M. Raza, H. Afzal, A. Lasebae, B. Naeem, S. U. Islam, and J. J. P. C. Rodrigues, "Heuristic edge server placement in industrial Internet of Things and cellular networks," *IEEE Internet Things J.*, vol. 8, no. 13, pp. 10308–10317, Jul. 2021.

[207] B. Li, P. Hou, H. Wu, and F. Hou, "Optimal edge server deployment and allocation strategy in 5G ultra-dense networking environments," *Pervas. Mobile Comput.*, vol. 72, Apr. 2021, Art. no. 101312. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1574119220301401

[208] F. Zeng, Y. Ren, X. Deng, and W. Li, "Cost-effective edge server placement in wireless metropolitan area networks," *Sensors*, vol. 19, no. 1, p. 32, Dec. 2018. [Online]. Available: https://www.mdpi.com/1424-8220/19/1/32

[209] K. Cao, L. Li, Y. Cui, T. Wei, and S. Hu, "Exploring placement of heterogeneous edge servers for response time minimization in mobile edge-cloud computing," *IEEE Trans. Ind. Informat.*, vol. 17, no. 1, pp. 494–503, Jan. 2021.

[210] B. Li, P. Hou, H. Wu, R. Qian, and H. Ding, "Placement of edge server based on task overhead in mobile edge computing environment," *Trans. Emerg. Telecommun. Technol.*, vol. 32, no. 9, Sep. 2021, Art. no. e4196. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/ett.4196

[211] L. Lovén, T. Lähderanta, L. Ruha, T. Leppänen, E. Peltonen, J. Riekki, and M. J. Sillanpää, "Scaling up an edge server deployment," in *Proc. IEEE Int. Conf. Pervasive Comput. Commun. Workshops (PerCom Workshops)*, Mar. 2020, pp. 1–7.

[212] I. Leyva-Pupo, A. Santoyo-González, and C. Cervelló-Pastor, "A framework for the joint placement of edge service infrastructure and user plane functions for 5G," *Sensors*, vol. 19, no. 18, p. 3975, 2019. [Online]. Available: https://www.mdpi.com/1424-8220/19/18/3975

[213] Y. Liu, Q. He, D. Zheng, X. Xia, F. Chen, and B. Zhang, "Data caching optimization in the edge computing environment," *IEEE Trans. Services Comput.*, vol. 15, no. 4, pp. 2074–2085, Jul. 2022.

[214] H. Sinky, B. Khalfi, B. Hamdaoui, and A. Rayes, "Adaptive edge-centric cloud content placement for responsive smart cities," *IEEE Netw.*, vol. 33, no. 3, pp. 177–183, May 2019.

[215] M. Wang, B. Cheng, and J. Chen, "An efficient service function chaining placement algorithm in mobile edge computing," *ACM Trans. Internet Technol.*, vol. 20, no. 4, pp. 1–21, Oct. 2020, doi: 10.1145/3388241.

[216] L. Gu, J. Hu, D. Zeng, S. Guo, and H. Jin, "Service function chain deployment and network flow scheduling in geo-distributed data centers," *IEEE Trans. Netw. Sci. Eng.*, vol. 7, no. 4, pp. 2587–2597, Oct. 2020.

[217] M. Siew, K. Guo, D. Cai, L. Li, and T. Q. S. Quek, "Let's share VMs: Optimal placement and pricing across base stations in MEC systems," in *Proc. IEEE Conf. Comput. Commun. (INFOCOM)*, May 2021, pp. 1–10.

[218] M. Chen, Y. Sun, H. Hu, L. Tang, and B. Fan, "Energy-saving and resource-efficient algorithm for virtual network function placement with network scaling," *IEEE Trans. Green Commun. Netw.*, vol. 5, no. 1, pp. 29–40, Mar. 2021.

[219] Y. Ma, W. Liang, M. Huang, W. Xu, and S. Guo, "Virtual network function service provisioning in MEC via trading off the usages between computing and communication resources," *IEEE Trans. Cloud Comput.*, vol. 10, no. 4, pp. 2949–2963, Oct. 2022.

[220] Q. Liu, R. Mo, X. Xu, and X. Ma, "Multi-objective resource allocation in mobile edge computing using PAES for Internet of Things," *Wireless Netw.*, pp. 1–13, Jul. 2020.

[221] C. Wu, Q. Peng, Y. Xia, Y. Ma, W. Zheng, H. Xie, S. Pang, F. Li, X. Fu, X. Li, and W. Liu, "Online user allocation in mobile edge computing environments: A decentralized reactive approach," *J. Syst. Archit.*, vol. 113, Feb. 2021, Art. no. 101904. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1383762120301739

[222] L. Wang, Y. Zhang, and S. Chen, "Computation offloading via Sinkhorn's matrix scaling for edge services," *IEEE Internet Things J.*, vol. 8, no. 10, pp. 8097–8106, May 2021.

[223] Z. Liu, J. Zhang, Y. Li, and Y. Ji, "Hierarchical MEC servers deployment and user-MEC server association in C-RANs over WDM ring networks," *Sensors*, vol. 20, no. 5, p. 1282, 2020. [Online]. Available: https://www.mdpi.com/1424-8220/20/5/1282

[224] S. Lee, S. Lee, and M. Shin, "Low cost MEC server placement and association in 5G networks," in *Proc. Int. Conf. Inf. Commun. Technol. Converg. (ICTC)*, Oct. 2019, pp. 879–882.

[225] F. Giust, L. Cominardi, and C. J. Bernardos, "Distributed mobility management for future 5G networks: Overview and analysis of existing approaches," *IEEE Commun. Mag.*, vol. 53, no. 1, pp. 142–149, Jan. 2015.

[226] L. F. Bittencourt, M. M. Lopes, I. Petri, and O. F. Rana, "Towards virtual machine migration in fog computing," in *Proc. 10th Int. Conf. P2P, Parallel, Grid, Cloud Internet Comput. (3PGCIC)*, Nov. 2015, pp. 1–8.

[227] T. Taleb and A. Ksentini, "Follow me cloud: Interworking federated clouds and distributed mobile networks," *IEEE Netw.*, vol. 27, no. 5, pp. 12–19, Sep. 2013.

[228] R. A. Addad, D. L. C. Dutra, M. Bagaa, T. Taleb, and H. Flinck, "Fast service migration in 5G trends and scenarios," *IEEE Netw.*, vol. 34, no. 2, pp. 92–98, Mar. 2020.

[229] L. Ma, S. Yi, N. Carter, and Q. Li, "Efficient live migration of edge services leveraging container layered storage," *IEEE Trans. Mobile Comput.*, vol. 18, no. 9, pp. 2020–2033, Sep. 2019.

[230] H. Abdah, J. P. Barraca, and R. L. Aguiar, "QoS-aware service continuity in the virtualized edge," *IEEE Access*, vol. 7, pp. 51570–51588, 2019.

[231] G. Avino, M. Malinverno, F. Malandrino, C. Casetti, and C. F. Chiasserini, "Characterizing Docker overhead in mobile edge computing scenarios," in *Proc. Workshop Hot Topics Container Netw. Netw. Syst.* New York, NY, USA: Association for Computing Machinery, 2017, pp. 30–35, doi: 10.1145/3094405.3094411.

[232] T. V. Doan, G. T. Nguyen, H. Salah, S. Pandi, M. Jarschel, R. Pries, and F. H. P. Fitzek, "Containers vs virtual machines: Choosing the right virtualization technology for mobile edge cloud," in *Proc. IEEE 2nd 5G World Forum (5GWF)*, Sep./Oct. 2019, pp. 46–52.

[233] E. Casalicchio and S. Iannucci, "The state-of-the-art in container technologies: Application, orchestration and security," *Concurrency Comput., Pract. Exp.*, vol. 32, no. 17, Sep. 2020, Art. no. e5668. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/cpe.5668

[234] I. Benkacem, T. Taleb, M. Bagaa, and H. Flinck, "Optimal VNFs placement in CDN slicing over multi-cloud environment," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 3, pp. 616–627, Mar. 2018.

[235] X. Li, X. Lan, A. Mirzaei, and M. J. A. Bonab, "Reliability and robust resource allocation for cache-enabled HetNets: QoS-aware mobile edge computing," *Rel. Eng. Syst. Saf.*, vol. 220, Apr. 2022, Art. no. 108272. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0951832021007468

[236] Y. Zhu, Y. Hu, T. Yang, T. Yang, J. Vogt, and A. Schmeink, "Reliability-optimal offloading in low-latency edge computing networks: Analytical and reinforcement learning based designs," *IEEE Trans. Veh. Technol.*, vol. 70, no. 6, pp. 6058–6072, Jun. 2021.

[237] G. H. S. Carvalho, I. Woungang, A. Anpalagan, I. Traore, and P. Chatzimisios, "Edge-assisted secure and dependable optimal policies for the 5G cloudified infrastructure," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2022, pp. 847–852.

[238] R. G. Garroppo, S. Giordano, and L. Tavanti, "A survey on multi-constrained optimal path computation: Exact and approximate algorithms," *Comput. Netw.*, vol. 54, no. 17, pp. 3081–3107, Dec. 2010, doi: 10.1016/j.comnet.2010.05.017.

[239] K. S. Trivedi, J. K. Muppala, S. P. Woolet, and B. R. Haverkort, "Composite performance and dependability analysis," *Perform. Eval.*, vol. 14, nos. 3–4, pp. 197–215, Feb. 1992. [Online]. Available: https://www.sciencedirect.com/science/article/pii/016653169290004Z

[240] D. M. Nicol, W. H. Sanders, and K. S. Trivedi, "Model-based evaluation: From dependability to security," *IEEE Trans. Depend. Sec. Comput.*, vol. 1, no. 1, pp. 48–65, Jan. 2004.

[241] T. Courtney, S. Gaonkar, K. Keefe, E. W. D. Rozier, and W. H. Sanders, "Möbius 2.3: An extensible tool for dependability, security, and performance evaluation of large and complex system models," in *Proc. IEEE/IFIP Int. Conf. Dependable Syst. Netw.*, Jun. 2009, pp. 353–358.

[242] T. A. Nguyen, D. Min, E. Choi, and J. Lee, "Dependability and security quantification of an Internet of Medical Things infrastructure based on cloud-fog-edge continuum for healthcare monitoring using hierarchical models," *IEEE Internet Things J.*, vol. 8, no. 21, pp. 15704–15748, Nov. 2021.

[243] Y. Yuan, Z. Yuan, and L. Tian, "5G non-orthogonal multiple access study in 3GPP," *IEEE Commun. Mag.*, vol. 58, no. 7, pp. 90–96, Jul. 2020.

[244] NTT DOCOMO. (Jan. 2020). *White Paper-5G Evolution and 6G*. [Online]. Available: https://www.docomo.ne.jp/binary/pdf/corporate/technology/rd/docomo5g/20200122_01/DOCOMO_6G_White_PaperEN_20200124.pdf,

[245] Hexa-X. (Feb. 2022). *WP1-Deliverable D1.3: Targets and Requirements for 6G-Initial E2E Architecture*. [Online]. Available: https://hexa-x.eu/wp-content/uploads/2022/03/Hexa-XWP1D1.3Summary-slides.pdf

GIANFRANCO NENCIONI received the M.Sc. degree in telecommunication engineering and the Ph.D. degree in information engineering from the University of Pisa, Italy, in 2008 and 2012, respectively. In 2011, he was a Visiting Ph.D. Student with the Computer Laboratory, University of Cambridge, U.K. He was a Postdoctoral Fellow with the University of Pisa, from 2012 to 2015; and the Norwegian University of Science and Technology, Norway, from 2015 to 2018. He has been an Associate Professor with the University of Stavanger, Norway, since 2018. He is currently the Head of the Computer Networks (ComNet) Research Group and a Leader of the 5G-MODaNeI Project, funded by the Norwegian Research Council. His research interest includes modeling and optimization in emerging networking technologies (e.g., SDN, NFV, 5G, network slicing, and multi-access edge computing). His past research activity was focused on energy-aware routing and design in both wired and wireless networks and dependability of SDN and NFV.

ROSARIO G. GARROPPO is an Associate Professor with Dipartimento di Ingegneria dell'Informazione, University of Pisa. He has expertise in networking. His main research activities are focused on experimental measurements and traffic modeling in broadband and wireless networks, MoIP systems, traffic control techniques for multimedia services in wireless networks, network optimization, and green networking. On these topics, he has published more than 100 peer-reviewed articles in international journals and conference proceedings. He has won the Best Paper Award at the fourth International Workshop on Green Communications, in 2011. He served as a technical program committee member for several international conferences on wireless networks and a referee for several international journals and conferences. He was a co-creator and a co-organizer of the International IEEE Workshop on Advanced EXPerimental activities ON WIRELESS Networks and Systems (EXPONWIRELESS), held in conjunction with IEEE WoWMoM, from 2006 to 2009.

RUXANDRA F. OLIMID received the first B.Sc. degree in computer science from the University of Bucharest, in 2008, the second B.Sc. degree in telecommunications from the Politehnica University of Bucharest, in 2009, and the M.Sc. and Ph.D. degrees in computer science from the University of Bucharest, in 2010 and 2013, respectively. She is an Associate Professor with the Department of Computer Science, University of Bucharest. She was a Postdoctoral Researcher and an Adjunct Associate Professor with the Department of Information Security and Communication Technology, Norwegian University of Science and Technology, Norway, from 2016 to 2018, and from 2018 to 2022, respectively. Her past experience includes Cisco certifications (CCNA and WLAN/FE) and almost ten years with Orange, Romania. Her research interests include cryptography and privacy, with a current focus on privacy and security in communication networks.