The Security and Privacy of Mobile-Edge Computing: An Artificial Intelligence Perspective

Cheng Wang, Zenghui Yuan, Pan Zhou, *Senior Member, IEEE*, Zichuan Xu, *Member, IEEE*, Ruixuan Li, *Member, IEEE*, and Dapeng Oliver Wu, *Fellow, IEEE*

***Abstract*—Mobile-edge computing (MEC) is a new computing paradigm that enables cloud computing and information technology (IT) services to be delivered at the network's edge. By shifting the load of cloud computing to individual local servers, MEC helps meet the requirements of ultralow latency, localized data processing, and extends the potential of the Internet of Things (IoT) for end-users. However, the crosscutting nature of MEC and the multidisciplinary components necessary for its deployment have presented additional security and privacy concerns. Fortunately, artificial intelligence (AI) algorithms can cope with excessively unpredictable and complex data, which offers a distinct advantage in dealing with sophisticated and developing adversaries in the security industry. Hence, in this article, we comprehensively provide a survey of security and privacy in MEC from the perspective of AI. On the one hand, we use European Telecommunications Standards Institute (ETSI) MEC reference architecture as our-based framework while merging the software-defined network (SDN) and network function virtualization (NFV) to better illustrate a serviceable platform of MEC. On the other hand, we focus on new security and privacy issues, as well as potential solutions from the viewpoints of AI. Finally, we comprehensively discuss the opportunities and challenges associated with applying AI to MEC security and privacy as possible future research directions.**

***Index Terms*—Artificial intelligence (AI), fifth generation (5G), Internet of Things (IoT), machine learning (ML), mobile-edge computing (MEC), security and privacy, software-defined network (SDN) security, virtual machine security.**

I. Introduction

THE NUMBER of end devices, which include smartphones, wearable gadgets, tablets, Internet of Things (IoT) devices, etc., has exploded in recent years. Simultaneously, a growing number of mobile and IoT applications, such as online gaming, virtual reality, and self-driving vehicles, have become more resource intensive and latency sensitive. These applications render the conventional cloud computing paradigm obsolete for its long propagation delays [1]. To meet the exponentially growing resource and latency requirements of these applications, the mobile-edge computing (MEC) paradigm was proposed [2]. The purpose of MEC is to push the powerful cloud computing capacities onto the edge servers that are in close proximity with end-users. MEC enables the provision of information technology (IT) services and cloud computing capabilities at the mobile network's edge, within the radio access network (RAN), and in close proximity to mobile subscribers. End devices can access the applications, services, and data on these edge servers with ultralow latency provided by the application vendors [3].

As a new distributed computing paradigm, MEC has brought many research topics to academic and industrial areas, such as computation offloading, data caching, service placement, etc. [3], [4], [5]. The security and privacy issues of the MEC environment have gradually attracted the attention of researchers due to the complexity of the MEC service model, multisource heterogeneous data, and resource-constrained end devices [6]. For example, a malware named "Mirai" manages as many as 400 000 compromised smart devices into a controlled "zombies" network to launch a Distributed Denial-of-Service (DDoS) attack against edge servers, shutting down over 178 000 domains. The number of cyberattacks on such facilities doubled again between November 2020 and January 2021 [7]. These fragile IoT devices directly or indirectly lead to this despondent result.

Compared with security and privacy issues in traditional cloud computing, MEC possesses several unique characteristics. First, in order to compete, a growing number of IoT devices deployed in the MEC environment are produced with economically manufactured circuitry that employs weak, guessable, or hardcoded passwords and other brittle security measures. Most such devices are easily preempted, contaminated, and destroyed by malicious users. Therefore, the MEC system is vulnerable to security and privacy threats directly brought by IoT devices. In this article, we purposefully incorporate the IoT system's security and privacy issues into the MEC in order to thoroughly investigate the MEC's security and privacy issues. Second, all end devices requesting edge

Manuscript received 12 December 2022; accepted 31 July 2023. Date of publication 11 August 2023; date of current version 7 December 2023. This work was supported in part by the National Natural Science Foundation of China under Grant 61972448, Grant 62172068, and Grant 61802048; and in part by the Hong Kong Research Grants Council, General Research Fund under Grant 11203523. *(Cheng Wang and Zenghui Yuan contributed equally to this work.) (Corresponding authors: Pan Zhou; Ruixuan Li.)*

Cheng Wang, Zenghui Yuan, and Pan Zhou are with the Hubei Engineering Research Center on Big Data Security, Key Laboratory of Distributed System Security of Hubei Province, School of Cyber Science and Engineering, Huazhong University of Science and Technology, Wuhan 430074, China (e-mail: wangcheng20@hust.edu.cn; zenghuiyuan@hust.edu.cn; panzhou@hust.edu.cn).

Zichuan Xu is with the School of Software and the Key Laboratory for Ubiquitous Network and Service Software of Liaoning Province, Dalian University of Technology, Dalian 116024, China (e-mail: z.xu@dlut.edu.cn).

Ruixuan Li is with the Intelligent and Distributed Computing Laboratory, School of Computer Science and Technology, Huazhong University of Science and Technology, Wuhan 430074, China (e-mail: rxli@hust.edu.cn).

Dapeng Oliver Wu is with the Department of Computer Science, City University of Hong Kong, Hong Kong (e-mail: dpwu@ieee.org).

Digital Object Identifier 10.1109/JIOT.2023.3304318

servers' services must through RAN, and this critical juncture is one of the weakest points in the entire network, resulting in serious communication link security issues, such as eavesdropping, hijacking, and DDoS attacks[1] [8]. Third, in order to achieve a serviceable platform with dynamic resource allocation capability and ease the burden of the network management, software-defined network (SDN)[2] [9], network function virtualization (NFV)[3] [10], and virtualization technologies are vital for realizing the MEC paradigm. These technologies are viewed as the potential solution to MEC's cost and efficiency problems. However, each of these technologies has its own security and privacy challenges which are easy targets for attackers. Finally, in the MEC environment, user data, such as user identity information, location information, and sensitive data is typically stored and processed by an honest-but-curious authorized entity, and the user has no way of knowing whether these semitrusted authorized entities will secretly obtain the user's private information for the purpose of illegal profit.

To decrease the security and privacy burden associated with MEC, numerous research directions have been explored, such as context-aware security, microservices, blockchain, etc. [6]. However, the unique characteristics of artificial-intelligence (AI)-based approaches have attracted much more attention than other approaches since it has the advantage of handling a large amount of unpredicted and complex data automatically [11]. AI technologies have been developed rapidly in the past few decades, from initial laboratory research to various commercial applications. As an important subset of AI, machine learning (ML) refers to the concept that computer programs can automatically learn from and adapt to new data without being assisted by humans. With the prosperity of graph processing units (GPUs) and big data, ML has completed remarkable achievements in many fields, such as computer version (CV), robotic, social media marketing, and gaming [12], and the simplicity and functionality of deploying learning algorithms in these fields significantly surpass almost all traditional rule-based algorithms. These leading advantages are also affecting the development of the security and privacy field.

According to the AV-TEST report, more than 450 000 new malware are registered every day [13]. However, most of the instances are just minor variants of the existing malware. Nonetheless, the correct identification of these specific malware needs to be based on many complex classification methods, such as hashes, simple rules, or heuristic fingerprints [11]. Fortunately, AI algorithms can manage vast amounts of unpredictable and complex data, which offers a distinct advantage in dealing with sophisticated and developing adversaries in the security industry. In order to meet the challenges of security and privacy, many AI algorithms have been used to protect data privacy and address security issues, such as spoofing attacks [14], DoS attacks [15], DDoS attacks [8], intrusions [16], jamming [17], eavesdropping [18], and malware [19]. A suitable learning algorithm can particularly use the trained model generated from the labeled data to recognize new security and privacy threats. Due to the advantages of learning algorithms, an increasing number of researchers are focusing on how to use them to address security and privacy concerns associated with MEC. The MEC environment integrates a variety of devices, such as IoT devices, mobile devices, and third-party servers, so that MEC naturally has a complicated network architecture, communication links, and various network protocols. In addition, the subscription nodes and the service nodes in the MEC have the defect of resource constraints. Therefore, in such a resource-scarce and heterogeneous distributed environment, traditional AI-based methods can no longer cope with the security and privacy challenges brought by MEC. However, the prosperity of various distributed, lightweight and green learning algorithms has paved the way for deploying such algorithms in the MEC environment.

In this article, we comprehensively consider security and privacy issues in the ETSI reference architecture from the viewpoints of AI.

The key contributions of this survey are listed as follows.

1) We purposefully incorporate the security and privacy of IoT systems as well as the SDN/NFV into the MEC environment to thoroughly investigate the MEC's security and privacy issues.
2) We provide an in-depth review of recent security and privacy issues in the MEC environment and from the layer viewpoints of ETSI reference architecture to thoroughly discuss the solutions for implementing AI technologies.
3) At the end of this article, we meticulously highlight the possible future directions about AI approaches for MEC security and privacy issues.

Fig. 1 depicts the overall architecture of this article. In the following content of this section, we investigate the current survey on discussing MEC's security and privacy and compare the papers that have been published so far and summarize the uniqueness and importance of our work. Then, we systematically overview the reference architecture of ETSI MEC in Section II. In Section III, we start to comprehensively study the MEC-specific security and privacy issues in the MEC environment. Based on the presented security and privacy issues, in Section IV, we thoroughly discuss the most promising AI algorithms, their advantages, disadvantages, and applications in the MEC security and privacy domains. Then, in Sections V and VI, we systematically introduce the AI approaches for layer-based MEC's security and privacy issues. In Section VII, we summarize the overall work of this article and propose research problems and future directions. Finally, we draw conclusions in Section VIII.

A. Related Work

In order to meet the increasing demand for sensitive applications and the proliferation of IoT devices, some papers have proposed to add an edge side fall between the cloud and users

[1]It is a malicious attempt to interrupt normal traffic to a targeted server, service, or network by flooding the target or its surrounding infrastructure with Internet traffic.

[2]It is a networking method that employs software-based controllers or APIs to interface with underlying hardware infrastructure and direct network traffic.

[3]It is the replacement of network appliance hardware with virtual machines.

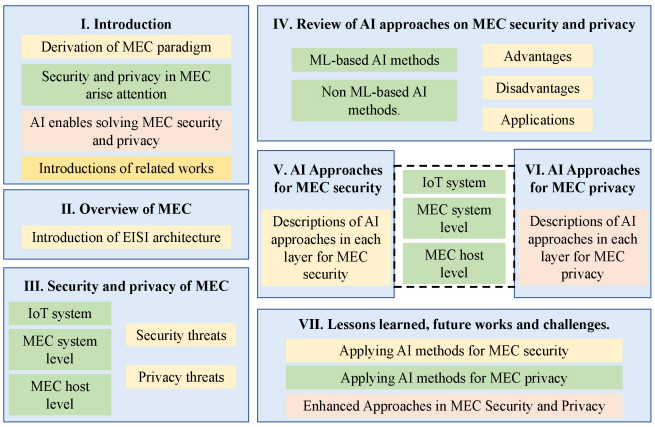

Fig. 1. Overall architecture of this article.

to reduce the pressure on the bandwidth and network traffic of cloud servers and increase the Quality of Services (QoS) of users. This article [20], [21], [22], [23], [24] proposed an initial three-tier framework about EC paradigms. Shi et al. [20] defined "edge" as any computing and network resources located between data sources and cloud data centers, and they provided several case studies involving computing offloading, data caching, data processing, and service delivery.

The development of fifth-generation (5G) wireless networks is gaining momentum, with the goal of connecting almost every aspect of life via the network at a significantly faster speed, with extremely low latency and ubiquitous connectivity. Additional to this, 5G networks, based on The 3rd Generation Partnership Project (3GPP) 5G specifications [25] will be a key future target environment for MEC deployments in the coming years. Mao et al. [1] provided an in-depth overview and future research directions for MEC from a 5G communication perspective. Peng et al. [26] presented a comprehensive survey of MEC from the standpoint of service adoption and provision. To meet the stringent latency requirements of applications (e.g., real-time applications) and to reduce energy consumption at user equipment (UE), Mach and Becvar [27] proposed a survey about computing offloading issues from different aspects.

The MEC's ability to ensure security and privacy has been widely questioned due to the use of multiple communication technologies, complex network structures, and multisource heterogeneous data types [6], [8], [28]. This article [26], [27], [29], [30] made no mention of security or privacy as critical aspects of MEC. Although certain surveys, such as [8], [23], [29], [31], and [32], mainly focused on security and privacy, the contexts are not concurring to the ETSI-standardized MEC architecture and its components. When it comes to security and privacy, the MEC paradigm is closely aligned with ETSI standards, according to a survey published in [6]. This survey aimed to guide the research communities on their way toward a feasible MEC deployment. Ali et al. [28] proposed a survey about the data security and privacy based on the ETSI-standardized MEC architecture. At the same time, some surveys about MEC security and privacy around ML began to appear [8], [33], [34]. Al-Garadi et al. [35] provided a comprehensive survey of ML methods and recent advances in deep learning (DL) methods that can be used to develop enhanced security methods for IoT systems. Shambour and Gutub [36] specifically summarize the application of some AI technologies in IoT-based Hajj and Umrah scenarios. Singh et al. [37] presented a review of ML for assisting security and privacy issues of EC, however, they only discussed the naive three-tier architecture and did not consider the security and privacy issues that two critical auxiliary technologies (i.e., SDN/NFV) brought to the EC environment.

To the best of our knowledge, we are the first survey that synthetically investigates ETSI-standardized MEC, IoT system, as well as their assistive technologies SDN/NFV's security and privacy issues from the perspective of AI. On the one hand, this survey identifies and compares the opportunities, benefits, and drawbacks of various AI approaches for MEC security and privacy. On the other hand, we comprehensively consider various possible AI solutions from the layer viewpoints of ETSI MEC reference architecture. Based on reviewing potential AI applications in the MEC security and privacy context, we discuss and present the identified challenges and future directions.

II. Overview of MEC

The European 5G Infrastructure Public Private Partnership (5G PPP) research body recognizes the MEC as one of the key technologies in the development of 5G and beyond 5G technologies [2]. MEC opens up services to mobile users and enterprise entities as well as to adjacent manufactories, these entities now can deliver their resource-intensive and latency-sensitive applications over the mobile network. ETSI and information services group (ISG) proposed the primitive standards and architecture of MEC which illustrated in Fig. 2. As an extension of cloud computing, the purpose of MEC is to push the powerful cloud computing capacities onto the MEC servers that are in close proximity with end-users. MEC has the following unique characteristics which differ from the traditional computing paradigms.

1) *On-Premises:* MEC can run independently of other networks, this is an essential attribution in machine-to-machine (M2M) scenarios.
2) *Proximity:* The MEC servers are often attached to the base stations or access points close to the end-users, improving real-time response ability.
3) *Lower Latency:* End-users data can be directly processed on the nearby MEC servers without delivering to the remote cloud. Hence, the end-users QoS and Quality of Experience (QoE) will get a considerable improvement.
4) *Location Awareness:* The MEC servers can only serve the specific geographic location's users covered by the base station to which it is connected.

In this section, we divide the edge server system into two layers according to the ETSI-published MEC framework and the reference architecture [6], [38]. Each layer has several submodules, and there are three groups of reference points to link different submodules, where Mx, Mm, and Mp denote reference points connecting to external entities, management

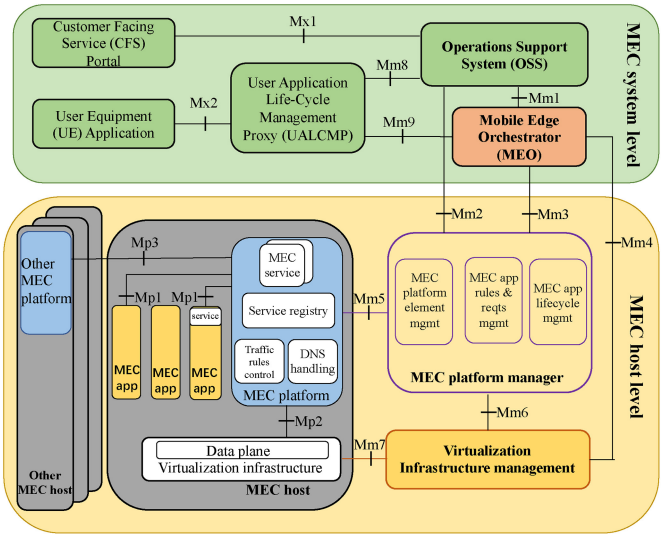

Fig. 2. MEC reference architecture.

reference points, and reference points regarding the MEC platform functionality, respectively.

1) MEC System Level: UE and Customer Facing Service (CFS) Portal are all external devices subscribed to the edge server services. Users interact with the ME system through the UE application that is instantiated in the UE. Specifically, the CFS Portal enables third-party customers to select and order a variety of edge applications to meet their unique requirements. Both UE application and CFS Portal connect the edge server through ANs. User app life-cycle management proxy (UALCMP) determines the UE applications can initialize, terminate or relocate ME applications and decides on the granting of these requests to be forwarded to the operation support system (OSS) and ME orchestrator (MEO). Then, it can send the status information of ME applications to UE applications. In particular, UALCMP can only receive requests which are within the geographic area that MEC servers covered, which means that it needs to fulfill the proximity constraint. OSS receives requests for initiation or termination that forwarded from the UE applications and CFS Portal, then decides whether to authorize these requests for further processing on MEO. MEO, as the backbone functional block of the mobile system layer, manages both MEC system level and MEC host level. Specifically, it can grant requests forwarded from the MEC system level to initiate, terminate or relocate ME applications, and dominate the resources of mobile host layer, such as selecting the appropriate ME host (MEH) to tackle low-latency or resource-incentive requests.

2) MEC Host Level: MEC host level can be divided into MEC server management level and MEC server level. ME platform manager (MEPM), as the backbone of the MEC host level, is responsible for the management of various functional blocks in MEH, including MEC platform element management, ME application rules and requirements management (i.e., service authorizations, traffic rules formulation, domain name system (DNS) configuration, and resource conflict resolution). MEPM is also responsible for managing the life cycle of applications and forwarding information to related applications through MEO. Virtualization infrastructure manager (VIM) is mainly responsible for allocating and releasing virtual resources (storage, compute, and networking resources) of VI, and using VI resources to create software images for serving CFS Portal and UE applications. VIM also collects the current status information of VI resources to MEO and MEPM. ME applications are software entities built on the top of VI. In a MEH, the connectivity among different ME applications is established through the local-area data network (LADN) [39]. After receiving traffic rules and DNS records from MEPM, the MEC Platform uses the traffic rule controller and DNS handling to configure the traffic rules of MEC applications and DNS proxy/server. MEC Service contains different services to facilitate MEC applications and MEC platform.

III. Security and Privacy Challenges of MEC

EC scenarios are centred on time-sensitive services, such as industrial IoT, autonomous driving, smart cities, etc. As a result, when designing network protocols and topologies, security and privacy are frequently sacrificed in favor of real-time and effective communication. Edge servers with on-demand and close proximity attributions are exposed to the network edge, shortening the distance between the attacker and the MEC physical devices. At the same time, the widely open application programming interfaces (APIs) make it easy for attackers to initiate security threats, such as data theft, information tampering, and node intrusion [40]. Finally, end-users' mobile nature allows them to dynamically join or exit the edge servers that cover them, and frequent topology changes between mobile devices and edge servers will have an impact on network resource management, allowing attackers to launch adaptive attacks by exploiting the calculation transforming between different edge servers. Fig. 3 [41], [42], [43] shows the overall architecture of a typical MEC deployment with its security and privacy issues.

A. Security Threats

In this section, we comprehensively consider the security issues in MEC and IoT systems with their serviceable platforms SDN and NFV.

1) IoT System Security: From the perspective of the ETSI MEC architecture, MEC-enabled IoT systems consists of two parts, one is IoT devices deployed in the actual environment, and the other is MEC IoT platform hosts as software instance which is migrated to the MEC facilities. IoT devices easily suffer various security risks due to current manufacturing and service vendors' lack of security awareness, lagging security standards, limited software and hardware resources, and the weakness of security protection capabilities. Large-scale cyber attacks targeting or originating from IoT have sparked widespread concern. For example, a malware named Mirai can exploit as many as 400 000 compromised smart devices into a controlled "zombies" network to launch a DDoS attack [44]. The MEC IoT platform appears as a service provider for MEC applications running on MEC hosts and enabled via the Mp1 interface as depicted in Fig. 2 [41]. In other words, MEC

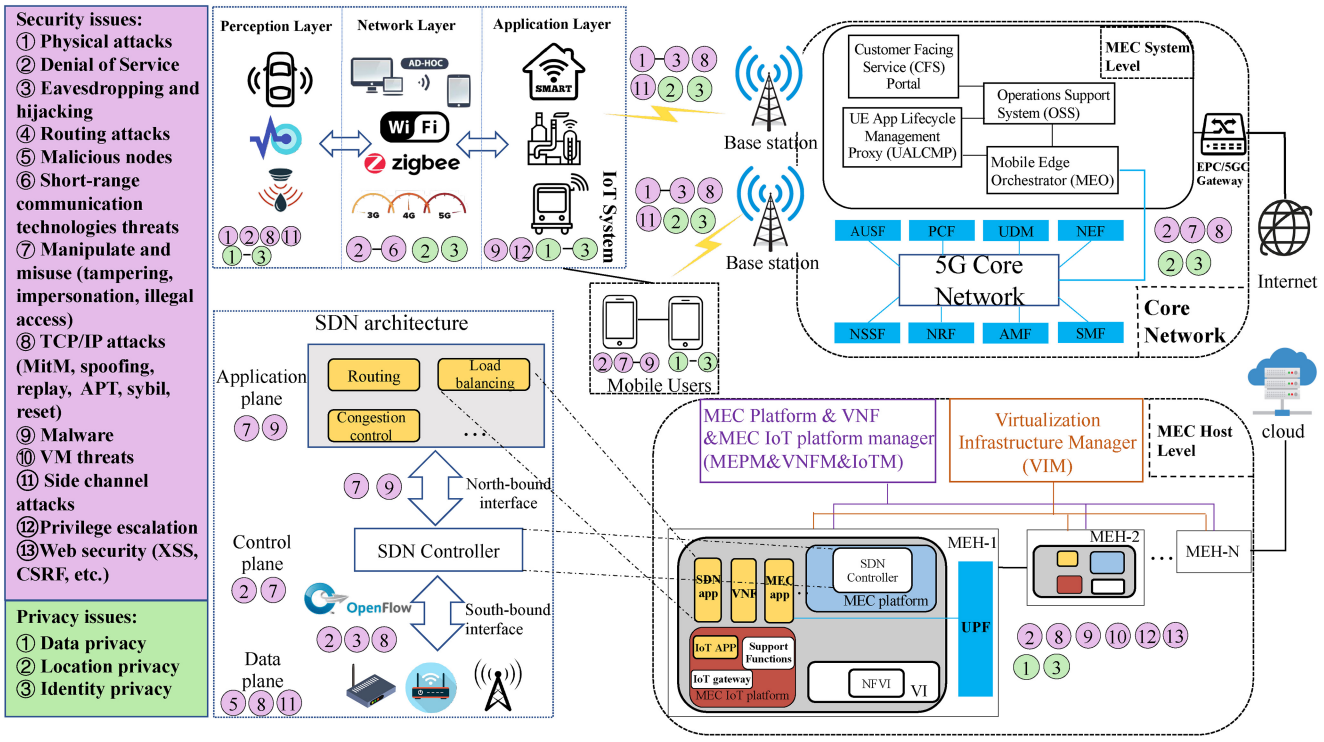

Fig. 3. Security and privacy threats of a typical MEC deployment.

applications can find the MEC IoT platform by querying its service registry, and interact with it using a defined IoT API exposed by the MEC IoT platform.

a) Perception layer: The main goals of the perception layer are to interact with the complicated real world and cooperatively connect heterogeneous sensors to provide diversely intelligent services. The perception layer can also be called the sensors layer. Considering the economic pressure brought by the massive deployment of IoT devices, they are generally designed as resource-constrained devices with low power and limited computing resources [45]. Therefore, it is a challenge to identify malicious data and authenticate benign equipment through traditional intrusion detection and authentication methods [14].

DoS/DDoS Attacks: The notorious DDoS attack, which has been intensively studied in the traditional cloud computing environment, is a new challenge and an open research topic in the MEC environment. Unlike traditional DDoS attacks in the cloud environment that mainly utilize computers as bots, a DDoS attack in the EC environment is often coordinated via control of mobile and IoT devices [46]. There are many applications with different functions installed on mobile user devices, but the security consciousness of different application developers is uneven, and some malicious applications often try to require permissions that are out of scope. Especially for some android devices, although the open-source android framework brings convenience for developers to require APIs to implement various functions, the fragile supervision capabilities of android communities make users easily threatened by malware. Simultaneously, some applications may hijack the user device's microphone in order to collect private information from daily conversations.

TCP/IP Attacks: If a fake base station can easily transmit a spoofing synchronizing signal with sufficient high power during the cell selection stage, wireless mobile UE may be drawn to and attempt to camp on the fake base station rather than any legitimate base station [47].

b) Network layer: The network layer is the backbone of the IoT system and has three main functions. First, it is responsible for establishing the network topology among IoT devices, then receiving and processing the raw information of the perception layer, and finally, the refined information is transmitted to the application layer to provide services for real-world customers [35].

Ad Hoc Network Security: Some short-range communication technologies can establish connections between UE equipments, such as Bluetooth, WiFi, IrDA, ZigBee, etc. This type of connection is device-to-device which establishes a direct communication link among UE equipments in IoT systems without requiring any edge servers for connection, and we call this type of network as the mobile ad hoc network (MANET) [40]. These communication technologies allow billions of heterogeneous devices to connect to the backbone and communicate with each other to complete complex and diverse intelligent functions. However, most end devices that adopt these technologies are resource-constrained so that the adversary can violate the route structure, congest wireless channels, and inject malicious nodes with ease.

Routing Attacks: While the MANET's nodes roam in a hostile environment, some statical security solutions are unable to adapt to the dynamic change in topology. The routing protocol of MANET such as dynamic source routing (DSR) establishes route paths from a source node to a destination node by exchanging the network topology information between these

two nodes. Since all messages are transmitted through wireless channels, so that any intruder can forge a valid routing updates and maliciously furnish incorrect routing states information. Intruders can also modify the route request (RREQ) or route replay (RREP) packets to delete a node, switch the order of the nodes, and append a node.

Attack on Short-Range Communication Technologies: In an ad hoc network, all communication signals must pass through a bandwidth-limited wireless channel, making the network vulnerable to physical-layer threats. The attacker can also eavesdrop and modify the information within the wireless communication channels or impersonate a legitimate participant to intentionally inconsistent wireless channels. Bandwidth-limited wireless channels are also vulnerable to DoS attacks via network-layer packet blasting. These packets exhaust a significant proportion of the computing resources, which introduce the wireless channel contention and network congestion.

Malicious Nodes: In a hostile environment, it is impractical for an ad hoc network to rely on the cooperation of all nodes to complete decentralized control, while malicious nodes can suspend the cooperation algorithm by refusing to respond. This situation also makes some centralized intrusion detection mechanisms incapable of detecting a node failure or a malicious intrusion.

Side Channel Attacks: Some IoT devices, such as smart medical devices, wearable devices, and smart home devices, collect raw data about the users' behavior and status [48]. These data are much more detailed and accurate, which contain much more sensitive information. As a result, an adversary can exploit these raw data to obtain additional privacy information. For example, Liu et al. [49] analyzed the data collected by the accelerometer sensors in the smartwatch via the new and practical side-channel to successfully infer the user's keystroke behavior [49].

c) Application layer: The application layer serves application subscribers by providing intelligent services. For instance, the application layer can provide surveillance video data and humidity measurements to subscribers. The importance of the application layer is to provide users with high-quality intelligent services.

Malware: IoT system applications are most closely connected with users, and they are required to defend against security and privacy threats while providing services for users. However, malware will deliberately obtain extra information beyond its requirements, resulting in the leakage of user's sensitive data and the preempt of IoT devices. For example, Fernandes et al. [50] analyzed the source code of SmartThings apps and found that more than 50% of the apps on the Samsung smart home platform have significant overprivileged.

2) MEC System Level With 5G Core Network: Mobile core network needs to ensure that end devices are securely connected to edge servers through an access interface and run a series of authentication protocols to prevent unauthorized devices. It also needs to provide some security measures, such as integrity protection and encryption to protect the communication information from manipulating and eavesdropping on the wireless communication channel. At the same time, the mobile core network is eager for the necessary differentiated security mechanisms to serve various personal businesses and vertical services.

a) AN security: As part of a mobile telecommunication system, AN resides between end devices and provides wide-area wireless connectivity to edge servers. It traverses the emanated service requests located at the different geographic locations to the edge servers due to the resource-constrained UE. In order to enable high network throughput, ubiquitous connections, and low latency, some novel technologies, such as massive multiple-input–multiple-output (MIMO), interference-aware receivers, and advanced coding/modulations are proposed to improve the spectral efficiency [51]. The connectivity between the UE and edge servers through these heterogeneous technologies raises several security concerns that could be exploited by the attacker [6].

Denial of Service (DoS): Jamming a wireless channel or compromising a service via DoS is to destroy the availability of AN connected to edge servers. The attacker raises a DoS attack in the MEC environment can only affect the edge servers where the botnets can access [46]. Therefore, it is difficult for the attacker to use their global botnets to launch a tremendous DDoS attack on a specific target edge server. However, the MEC has deployed a large number of latency-tolerant applications, disrupting legitimate users' access to services will greatly affect the QoS of MEC systems. It is difficult to detect such malicious network activities because of the direct connection between MEC systems and end devices [27]. Novel botnet-type DDoS attacks will also affect the availability of UALCMP and CFS Portal in the MEC system level.

Eavesdropping and Hijacking: A wireless channel between UE equipment and edge servers is prone to cyber risks, such as eavesdropping and hijacking due to the broadcast nature of the wireless medium. As a result, wireless communication channels are vulnerable to being hijacked to retrieve information by cyberattacks, such as Man-in-the-Middle (MitM), replay, advanced persistent threat (APT), Sybil, and spoofing attempts [52]. Traditional wireless channels mainly use some encryption protocols to secure the transmission process [52]. However, the process of information encryption and private key exchanges between channels largely hinders its applicability for massive latency-sensitive applications in the MEC environment.

b) MEC system level: MEC system level as part of the core network of MEC system connecting to end devices, MEH, and 5G core network. It has the following three aspects of functions. First, it determines to grant services for further process and handle services life cycle forwarded from end devices. Second, it guards the resource utilization status, and the configuration of VMs and underlying hardware in the MEH. Finally, it switches control signals with a 5G core network via wireless, wired, or optical. Thus, the MEC system level is critical to attackers for gaining unauthorized access, manipulating or misusing ME applications, etc.

DoS/DDoS Attacks: As the UALCMP is the entity handling the life cycle of requests, while the OSS grants the approval for subscribers to use a particular MEC service. The attacks over them can be targeted at congesting the access interfaces

between them and UE so that both of them have to be protected from DoS/DDoS attacks.

Manipulate and Misuse: Since all the ME application subscribers should be registered in OSS or UALCMP, the attackers could attempt to inject craftily constructed information to manipulate or misuse the functionality of these two. At the same time, malicious UE can inject fake information to impersonate legitimate entities to disrupt the everyday activities of MEO and MEPM. The malicious intrusions are improbable at MEO since MEO is deployed at the in-depth MEC system level that is difficult for attackers to reach. However, resource allocation and service manipulation attacks are highly possible, such as DNS amplification and VM escape.

TCP/IP Attacks: Separate physical hosts for the 5G core network and MEO are more prevalent [53], which introduce typical TCP/IP attacks, such as eavesdropping, spoofing, DoS, replay, and reset attacks.

c) SDN threats: In the MEC environment, in order to reduce network management costs and improve the network scalability and flexibility, infrastructure providers have introduced SDN. SDN separates the network architecture into a three-tier architecture of application, control and data plane [54], controlling the network with software to achieve centralized management of the network and programmability of network applications. This convenient network management approach is not only convenient for operators but also for attackers. A capable attacker can preempt SDN applications or tamper with the flow tables in SDN controllers to chaos packets forwarding.

Applications Plane Threats: Network functions are implemented as applications in this plane, and these applications are created by using the VI's computing resources in the MEH. These applications aim to provide network services, such as QoE, monitoring, load balancing, security, etc. These services host as one type of special MEC application in the MEH and exchange data with the SDN controllers through the north-bound interface. Thus, an authorized malicious application can invade other applications to control the network. This malicious threat can be caused by the open APIs of network equipment, the lack of mutual authentication mechanism between the application plane and the control plane, or the implementation of the wrong access control method for third-party applications [55].

Control Plane Threats: The SDN control plane communicates with the application plane and the data plane through the north-bound interface and the south-bound interface, respectively, and there can be multiple controllers in SDN. In the ETSI MEC architecture, these controllers are as one of the support functions that deployed in the MEC platform [43]. The control plane adopts a centralized management architecture, which relies on programming to achieve overall control of SDN data plane devices. Thus, an attacker can directly send specifically crafted DDoS flows to overwhelm the resources of control plane [56] or IP packets with random header fields to disturb legitimate flow setup [57]. Also, it is incompetent to secure application authorization, resource usage, and tracking while malicious applications in control plane [58].

Data Plane Threats: The main elements of the data plane are switches and routers. They are simple packets forwarding elements without embedded intelligence to take autonomous decisions. These data plane devices communicate with the controller through a standard OpenFlow interface, ensuring the compatibility and interoperability of configuration and communication among different devices. Since the packet forwarding of the data plane depends on the flow tables issued by the control plane, an attacker can control the forwarding of data packets by intercepting or tampering with the flow tables sent to the SDN switches [9]. In addition, three types of attacks may be used to compromise the data plane, including device attack, protocol attack, and side-channel attack [59]. Device attacks aim to exploit SDN software or hardware vulnerabilities to compromise the SDN data plane. Protocol attacks exploit network protocol loopholes in the forwarding equipment (e.g., as border gateway protocol (BGP) attacks) to attack the data plane. Side-channel attacks infer the network's forwarding strategy by analyzing the performance metrics of the forwarding equipment. For example, by analyzing the processing time of data packets, attackers can identify forwarding strategies [60].

3) MEC Host-Level Security: In the MEC paradigm, the MEH is the primary host-level functional entity that performs computational, storage, and networking operations. Furthermore, the user plane function (UPF) is a 5G AN entity included within the MEH for integrating the 5G core network into the LADN. Since the SDN security threats and the MEC IoT platform have been illustrated above, we will focus on the virtualization threats and NFV threats in this section.

a) VM threats: Virtualization integrates physical resources into a resource pool providing various on-demand services, which decreases the high management complexity and operational expenses, and enhances the efficiency for usage and fine-grained control. However, novel security threats and vulnerabilities introduced in the following contents become one of the major concerns in the MEC environment.

The overall VM entities are actually composed of SDN applications, virtual network functions (VNFs), IoT applications, and the MEC applications, which provide services for users or MEC architecture. The threats and vulnerabilities target to the VMs are as follows.

Infected VM Images: VM image, as a prepackaged software file, contains the configuration templates used to initiate the VM instances on demand. The resource renters can either create their own VM images from an image software or downloading VM images stored in the third-party's repository [54]. Thus, this gives the attacker the opportunity to upload VM images injected with malicious code such as a Trojan horse to the third-party repository, and the victim will be infected with the hidden malware after uploading such malicious VM images.

Compromising VM Migration: VM migration allows network operators to optionally initiate, terminate a VM, or move VMs from one physical machine to another. The migration of VM benefits the workload balancing and system management. Due to the dynamic nature of VM and the

plaintext presented in migration data, an attacker can easily launch MitM attacks to sniff or tamper the traffic.

VM Hopping: The attacker gains access to a host with multiple VMs by renting or hacking a guest VM. From the persecuted guest VM, the attacker then compromises other guests' VMs through privileged access to the host. The reason for this could be that the memory management module (MMU) of the hypervisor allows attackers to perform illegal manipulation on the memory pages of other guest VMs based on the access rights they have obtained [10].

VM Escape: Any interaction between VM and hypervisor may become a potential attack vector [61]. VM escape refers to the program running in the virtual machine using the vulnerability of the VM to break through the virtual machine monitor (VMM) or hypervisor. After that, the adversaries may obtain the host OSS management authority, control other virtual machines running on the host machine [62], and completely destroy the original security architecture.

b) NFV threats: NFV, as one of the key emerging technologies for 5G networks, consolidates multiple network functions onto the software, running on a range of industry-standard hardware [10]. The infrastructure that hosts MEC and NFV is quite similar [2]. Thus, it will be beneficial to reuse the infrastructure and infrastructure management of MEC by hosting both VNFs and MEC applications on the same platform. VNFs can be recognized as software that encapsulate network functions in a VM. Therefore, most NFV threats also are inherited from VM threats. NFV-specific security threats may from the NFV management and orchestration (NFV MANO). It is mainly responsible for the orchestration and management of virtual resources, the creation of VNF, and NFV lifecycle management. ETSI has published an official document about NFV MANO to describe how the NFV and its interfaces work without specific details about interface design and implementation [10]. Thus, an adversary may exploit the insecure interface to obtain sensitive data or launch a cross site scripting (XSS) or cross site request forgery (CSRF) attack by injecting a well-constructed script into the Web surfaces provided by the management interface.

B. Privacy Threats

User data in the MEC environment, such as user identity information, location information, and sensitive data, is typically stored and processed by an honest-but-curious authorized entity (e.g., edge data center and infrastructure provider), and the user has no way of knowing whether these semitrusted authorized entities will secretly obtain the user's private information to achieve the purpose of illegal profit. At the same time, in the open ecosystem of MEC, multiple trust domains are dominated by different infrastructure providers, and users cannot identify which service provider is trustworthy. Thus, MEC, with the complexity and real-time nature of the service mode, multisource heterogeneous data, and resource-constrained end devices, has more delicacy privacy threats.

In this section, we will focus on three aspects of privacy concerns: 1) data; 2) location; and 3) identity privacy.

1) Data Privacy: Users outsource data to edge nodes (i.e., IoT devices or edge servers) with computing resources, which give edge nodes the opportunity to control over the data, introducing the same security risks as cloud computing. It is challenging to ensure the confidentiality and integrity of the data since the complicated communication link may cause data to be lost or maliciously modified. Additionally, through privilege escalation, unauthorized entities may exploit the uploaded data for their own gain. Compared to cloud servers, edge servers have partially circumvented the data security and privacy issues caused by the long-distance transmission of multihop routing. However, the applications that are dominated by different application vendors and the ANs that belong to different telecom operators have compelled MEC to introduce more severe data privacy issues, such as coexistence of multiple security domains and data in multiple formats.

2) Location Privacy: LBS refers to using a certain positioning technology (e.g., global positioning system, mobile phone positioning, and positioning through WiFi access points) to provide mobile users with personalization related to their current location service. However, these services frequently collect location data in the background without the user's knowledge or consent. As a result, figuring out how to protect a user's location privacy has become a pressing issue [63]. The leakage of location information can be divided into three major kinds of threats as follows.

1) *Tracking Threat:* The adversary may obtain continuous location updates, allowing him to pinpoint the user in real time. For example, in vehicle ad hoc networks (VANETs), an attacker can eavesdrop on the communication between vehicles to lock and track the target vehicle; the attacker can tamper with the traffic information in the network and publish false information on the network, which creates the illusion of road congestion for other vehicles and affects the route selection of other vehicles.
2) *Identification Threat:* Even if the adversary only accesses the user's location on a sporadic basis, he may be able to isolate the user's frequently visited locations, such as home and work. The adversary can use these locations as pseudo-identifiers to deduce the user's identity from anonymous location traces [64].
3) *Profiling Threat:* The mobility track of the user may not include places that reveal his identity but can be used to profile him by the adversary. Some location service providers (LSPs) collect and analyze the location attributes in the user's request to infer the user's personal information, behavioral preferences, and physical condition [65]. For instance, if a user sends a location request that frequently includes a specific hospital location, the LSP can infer the user's physical condition, and the user's home address can be inferred based on the user's resident location at night.

3) Identity Privacy: Personal identifiable information (PII), also known as user identity, is the information about a person that has been collected, assessed, or used on demand by edge or cloud services [65]. Compared with the cloud computing data center located in the core network, edge nodes

TABLE I
AI METHODS FOR MEC SECURITY AND PRIVACY

Classification	Method	Advantages	Disadvantages	Example Application Scenarios
Supervised learning	kNN	Simple, cheap and efficient in detection tasks	Weak transferability	Intrusion detection [66]–[68] and privacy-preserving [69], [70]
	SVMs	High performance with small samples, strong generalization ability	Hard to acquire the optimal kernel function, sensitive to missing data	Authentication [71], intrusion [72] and malware [73] detection
	LR	Cheap, capacity of description the relationship between input and output	Easy to be underfitting, weak performance with missing data and large feature	Intrusion detection [74], PUF [75]
	DTs	Interpretable inference, fast execution	Sensitive to noise, ignoring relationship among different attributions in dataset	Intrusion detection [76], privacy-preserving [77], [78]
	RFs	Parallel efficient framework, strong generalization ability, and robust to missing data	Performing overfitting on datasets with large noise	Intrusion [79]–[81] and anomaly [82]–[84] detection
	CNN	Low network complexity, reducing weight parameters, improving performance by BP algorithm and better scalability	Large training cost, requiring excessive computation resource and memory	Intrusion detection [85], attack recognition [86] and privacy-preserving [87]
	RNN	High performance through correlation extraction for sequential data	Gradient explosion in training process with long-time sequence dataset	Malware detection [88]
Unsupervised learning	K-Means	Fast execution, interpretability and performing well in clustering with unlabeled data	Sensitive to abnormal data, dependence on the value of K	Secure links decision [89], privacy-preserving [90], [91]
	PCA	Reducing the data dimension, eliminating the effect between different components	Ignoring the influence of non-principal components, requiring to be combined with other methods	Attack recognition [92], privacy-preserving [93]
	RBM	Strong representation capacity, high inference performance with unlabeled data	Overfull computation cost, challenges in deployment on-board	Anomaly detection [94]
	DBNs	Flexible and efficient parallel framework, better scalability	Overfull computation cost, slow convergence rate	Intrusion detection [16], [95], anomaly detection [34]
	AE	Strong generalization capacity providing privacy-preserving through the encoding-decoding process	Requiring benign data for training in malicious detection	Anomaly detection [96], privacy-preserving [97]
Semi-supervised learning	S^3VM	Improving the performance with scarce labeled data	High computational complexity, requiring to resolve the programming when new data is added	Anomaly detection [98]
	GANs	Generative model based on BP algorithm instead of Markov chain, generating clearer and faster samples	Unstable training process, inefficient for discrete data	Anomaly detection [99], PUF [100]
RL	Non-deep RL	Framework for sequential decision-making problem	Excessive training iterations, risk of falling into suboptimal solution	Malware detection [101], anti-jamming [102]
	DRL	Efficient inference, high performance in MEC offloading game	Complex model setting and high computation cost	Anti-jamming [103] threat edge game [104]
Non-ML based methods	Bayesian Networks	Capacity of dealing with the unsure information	High computational complexity	Intrusion detection [105], privacy-preserving [106]
	Evolutionary Algorithms	Efficient, potential scalability	Complex manual programming, unsuitable to large datasets	Malware detection [107]

are located at the edge of the network, enabling the collection of more high-value sensitive information of users, such as location information, lifestyle habits, social relationships, even health status, etc. A considerable privacy crisis will occur while an attacker can map each private data to the corresponding user. For instance, in early 2018, the Facebook data scandal resulted in the disclosure of 50 million users' PII to a third-party company, Cambridge Analytica, via service providers for "analysis" purposes.

IV. REVIEW OF AI APPROACHES ON MEC SECURITY AND PRIVACY

In this section, we make a comprehensive review of the MEC security and privacy from the perspective of AI. First, we introduce the classification of ML approaches and summarize advantages, disadvantages as well as applications of the typical approaches. Then, considering the entire AI categories, we present some other applications of non-ML approaches. The summary of the AI-based methods and their corresponding advantages and disadvantages are introduced in Table I.

A. Classification of ML Methods in MEC Security and Privacy

Discussions and research around AI have focused on the field of ML in recent years [108]. We divide ML approaches into four categories: 1) supervised learning; 2) unsupervised learning; 3) semisupervised learning; and 4) reinforcement learning (RL). In this section, the characteristics of these four categories, classic algorithms and their applications in MEC are discussed in detail.

1) Supervised Learning: Supervised learning has attracted more and more attention in the field of the security and privacy protection of MEC due to its outstanding performance in prediction and classification. Generally, utilizing the labeled data as the input of supervised learning, the model can predict the label of test data through learning from the distribution of labels. We discuss the classic algorithms of supervised learning

and their applications in the protection of security and privacy of MEC in the following.

a) k-nearest neighbor: Based on the unsurpassed performance on classification tasks, *k*-nearest neighbor (kNN) has been widely used in network malicious intrusion detection [66], [67], [68]. As to the MEC privacy, computational-hungry algorithms are not suitable for the practical applications with massive data and computation load. Therefore, a lightweight edge-based kNN (EBkNN) model was proposed to protect users' security and privacy in [69]. However, the transferability of kNN is poor. A trained model often needs to be retrained to determine the optimal *k* value when the data set changes, which will increase the resource consumption of the light-weighted edge devices.

b) Support vector machines (SVMs): Due to the advantages of SVMs in solving linear and nonlinear data classification problems, they have been widely used in abnormal intrusion detection in the MEC environment. The research [71] chose the SVM algorithm to detect cloning attacks and Sybil attacks in industrial edge networks. For the actual data without labels, this work adopted the threshold detection method to generate labels for training, so that the model obtained improved detection accuracy. Nevertheless, for the scenarios with complex data distribution, the optimal kernel function is hard to acquire in SVM, which may decrease the performance of intrusion detection system (IDS)[4] [97], and this disadvantage will be further amplified in the data-heavy MEC environment.

c) Logistic regression: Logistic regression (LR) has been widely used in protecting the security and privacy of MEC. In response to the vulnerability in physical unclonable function (PUF), Laguduva et al. [75] proved that PUF can be cloned without prior knowledge of its structure. Based on ML, this work proposed a noninvasive attack method against the PUF of edge nodes and the corresponding defence strategy, which adopted LR, random forests (RFs), artificial neural network (ANN), and merged algorithms, respectively. The attack effectively cloned the PUF in MEC and the countermeasure greatly improved the accuracy of recognizing authentic and cloned PUF. However, LR is with the drawback of falling into the dilemma of overfitting that limits the accuracy, and has poor performance with missing data and large feature in MEC environments.

d) Decision trees: For the data with high-dimensional features in MEC, decision trees (DTs) can reduce the number of features in the internal nodes while ensuring a certain performance in inferring. Based on this, the DTs-based algorithm has been widely used in the network intrusion detection [109]. In the edge–cloud computing (ECC) scenario, a private random DT framework based on differential privacy (DP) was used to analyze the impact of different applications on privacy, so as to better coordinate the overall system to ensure the data privacy [77]. Whereas, DTs are sensitive to the noise data and ignore the relationship among different attributions in the data set, and these disadvantages affect the performance of preserving security and privacy seriously.

e) Random forests: RFs have been widely used in the IoT system for intrusion detection [79], [80], [81] and anomaly detection [82], [83], [84]. Especially in the detection of DDoS attacks, a mass of previous work has analyzed the performance of RFs [82], [110]. A research [111] proposed an automatic anomaly detection system based on RFs; the proposed system effectively detected anomalies in distributed edge devices. The disadvantage of RFs is that it is prone to overfitting on data sets with large noise.

f) Convolutional neural network: As one of the most classic DL models, the convolutional neural network (CNN) has been fully employed in protecting the MEC security and privacy [112]. Using the CNN model, Tian et al. [85] proposed a Web server attack detection system based on a distributed framework. The uniform resource locator (URL) in the IoT-cloud environment was analyzed and a high detection rate was obtained. However, the problem of CNN is that the amount of calculation under the application of high-dimensional data is huge, and it is difficult to guarantee sufficient computing resources in some resource-constrained MEC environments. Some current neural-network-scale compression technologies and partition training technologies [113], [114] provide support for the deployment of lightweight CNN on mobile terminals.

g) Recurrent neural network: Another classic DL model called the recurrent neural network (RNN), which is aimed at sequence-type data, has also been fully used in the malware detection in the IoT network [88], [115], [116], [117]. Conventional RNN models use the backpropagation training time (BPTT) to extract sequence data, but it will cause gradient explosion in the training process with long-time sequence data. Previous work [118] applied LSTM to the cloud-edge scenario to achieve distributed network attack detection. In the experiment, the performance of the proposed model was compared with LR, which proved that the utilization of LSTM can fully extract the correlation information from the long-term continuous network state data. However, the RNN model generally has the disadvantage of gradient explosion, which will limit its performance in long-term sequence training.

2) Unsupervised Learning: Unsupervised learning is an ML method for unlabeled data sets, which clusters different classes by learning their associations [119]. In addition, unsupervised learning can mine the structure information in depth, which may be hidden by labels in supervised learning [120]. Thus, the unsupervised learning algorithm has obvious advantages in the classification task in MEC security and privacy with a large number of unlabeled data. For example, the next-gen cloud security company, Bitglass, exploits unsupervised learning approaches to provide advanced threat protection (ATP) services for detection both known and zero-day threats on cloud applications [121]. The emblematic unsupervised learning methods in MEC security and privacy are listed as follows.

a) K-means: Based on the excellent performance of *K*-Means in feature clustering on unlabeled data, it has been applied to the field attack detection and privacy protection

[4]It is a system that analyzes network traffic for suspicious or unusual activity and generates notifications when it detects it.

in MEC. In order to solve the complex link attack problem caused by the open and multisource characteristic in the MEC environment, Li et al. [89] proposed a link decision scheme based on attribute attack graphs, which used K-Means to deredundate complex network alarm information to construct the attribute attack graphs, and utilized the greedy algorithm to make decisions on the link defense by the solution of the minimum dominating set of the attribute attack graph. However, K-Means is sensitive to the abnormal data, and the progress of K-Means clustering mainly depends on the value of K, which is usually set manually. And these disadvantages cause the weak robustness of K-Means-based methods.

b) Principal component analysis: Due to the effective feature extraction and dimension reduction capacity for large data sets, principal component analysis (PCA) fits the MEC environment well. The previous work [92] proposed a multiattack detection model with low complexity in the cloud-edge environment. The computationally efficient and low-cost PCA was selected as the feature extractor, and the deep neural network (DNN) was used as the classifier. In the experiment, the model was used to detect ten types of attacks with higher accuracy rate than other ML methods. Nevertheless, PCA ignores the influence of nonprincipal components, which may play a crucial role in inferring. Discarding them directly will decrease the classification accuracy. And PCA is effective for feature extraction, while it requires to be combined with the other AI-based methods to fully finish the inference task.

c) Restricted Boltzmann machine: Restricted Boltzmann machine (RBM) is a deep generative neural network based on unsupervised learning, which is a variant of the Boltzmann machine, and has been adopted in MEC security protections. In terms of malicious detection classifiers, the adopted features are the key factor affecting its accuracy. Benchea and Gavriluţ [94] used RBM to improve the feature generation, and generated new features in a nonlinear manner to train the classifier and improve its performance. However, the drawback of RBM is also obvious. The training of RBM takes a mass of computation cost, and it is still a challenge to deploy the model on-board in the resource-limited edge sever.

d) Deep belief networks: Deep belief networks (DBNs) are widely employed in the intrusion detection in IoT networks due to the efficient performance with a mass of unlabeled data [16], [73], [95], [122]. In order to realize the attack detection in the edge transmission network, an unsupervised-based model based on DBNs was proposed in [8]. In this work, the Android data package files in edge devices were utilized to extract the permission information, sensitive program APIs, and dynamic information as features, which were adopted by DBNs for attack detection. Unfortunately, the disadvantages of DBNs are manifested in their high computational cost and slow convergence rate, which limit the performance of DBNs in MEC security and privacy-preserving.

e) Deep autoencoder: Deep autoencoder (AE) is widely used in IDS for the MEC environment due to its feature extraction capabilities [97], and the process of feature extraction-reconstruction also provides privacy protection for user information in the original data. The previous work [96] utilized Deep AE to realize anomaly detection based on the characteristics of MEC. The model was trained in a distributed manner on multiple edge severs. The abnormal data from each edge sever was aggregated to update the model on the central server, and then the updated model was sent to each edge sever to reduce the load on the central server. While in terms of malicious detection, it requires a large amount of benign data for training an AE model, which is unrealistic in practice.

3) Semisupervised ML Learning: Semisupervised learning is a technology that falls between supervised learning and unsupervised learning [123]. It can solve the inability to construct a reliable supervised classifier caused by scarce labeled data. Thus, the semisupervised learning method is well-suited to the MEC scenario, in which a large amount of raw data is generated at any time and in any location but is incapable of being used efficiently. This inherent advantage makes it easier for Semisupervised learning algorithms to address security and privacy concerns in MEC.

a) Semisupervised SVM: The discovery of the distributed semisupervised SVM (S^3VM) architecture [98] enables it to solve the security problems in the MEC environment which owns tremendous unlabeled raw data and a small amount of labeled data. Wang et al. [98] proposed a hybrid approach including S^3VM-based appliance pattern matching classifier and the hidden Markov model (HMM)-based energy consumption habit classifier to defend against anomaly intrusion.

b) Other non-DL semisupervised approaches: MEC has a number of characteristics, including real-time processing, low computation costs, and a distributed architecture. Some of the above characteristics are already present in security and privacy-preserving methods based on semisupervised learning. In [124], an ELM-based semisupervised fuzzy C-Means (ESFCM) attack detection framework was proposed to secure the IoT system. At the same time, an optimized, low-cost semisupervised IDS model which combines an active learning SVM (ASVM) and fuzzy C-Means (FCM) clustering was proposed in [72]. To defend Sybil attacks that appeared in distributed environments, recently, Gong et al. [125] designed a Markov random fields-based, noisy-tolerated and scalable semisupervised learning approach. The above security defense mechanisms are more or less satisfying the conditions for deployment in the MEC environment. Despite the fact that researchers dislike the current semisupervised learning method, it cannot be ruled out that it has the potential to become a new mainstream solution to the security and privacy issues in the MEC environment.

c) Generative adversarial networks: Generative adversarial networks (GANs) have been recently implemented in MEC security and privacy. In [99], a supply chain risk management architecture that mixes ML, cryptographic hardware monitoring, and distributed system coordination techniques was proposed to detect normal and abnormal system behaviors in IoT systems. GANs can learn the distribution of attack samples from existing attacks in MEC, thereby generating zero-attack samples that do not exist in the set of known attack samples. Even for degenerate distributions, GANs can generate sharp samples without Markov chains requirement. As a DL-based semisupervised approach, GANs are suitable for training classifiers with limited ground-truth data sets. Hence,

GAN is a promising method for developing security and privacy applications in the MEC environment. The latest article about compressed CGANs models [126] made it feasible to deploy GANs in the MEC environment.

However, the disadvantages of GANs are also obvious, such as "the Helvetica scenario" and the instability and difficulty of the training process. We must avoid these drawbacks in order to maximize the benefits of GANs and use them to protect the security and privacy of MEC.

d) Other DL-based semisupervised approaches: DNN can fit various nonlinear functions appropriately, so it has outstanding contributions in the fields of pattern recognition, data analysis, and control. Simultaneously, semisupervised learning can be well adapted to a small number of labeled data together with a massive number of unlabeled data. Semisupervised learning combined with the DNN method began to be widely concerned by researchers in order to make good use of the advantage of these two approaches. Although semisupervised DL sounds like an effective solution to deal with a large amount of unlabeled data in the MEC environment, the detection accuracy is hard to exceed supervised DL. Therefore, only a small amount of work now uses semisupervised DL in terms of MEC security and privacy. For example, Abdel-Basset et al. [127] proposed a semisupervised DL approach for intrusions detection in IoT networks, which combines multiscale residual temporal convolutional (MS-Res) module to finetune the network capability and traffic attention mechanism to help the model to concentrate on important information in the learning process.

4) Reinforcement Learning: Researchers have extensively studied popular topics in communication fields, such as data caching and offloading in the MEC environment. In recent years, a large number of papers are using DRL algorithms to solve these problems [104], [128], [129], [130]. Previous studies (e.g., Integer programming and game-theory methods) only consider one-shot optimization [131]. However, the process of caching or offloading tasks in the MEC environment is a continuous process. Fortunately, RL provides a promising approach to maximize long-term rewards. At the same time, security and privacy issues have become the bottleneck of mobile caching/offloading, as edge server is always deployed close to users, making attackers also easier to reach the vulnerable position of MEC. Therefore, Xiao et al. [103] formally defined a repeated game between MEC systems and attackers and build the RL-based security solutions to defend against jamming and smart attacks in mobile offloading/caching.

However, the RL agent sampling data from the environment is a very inefficient process, and designing an appropriate reward function for unknown environments is challenging. Even if we successfully overcome the aforementioned difficulties, the local optimal is still hard to escape [132].

B. Other AI Approaches

In addition to the previously introduced ML methods, some other non-ML methods (in AI categories, but not in ML categories) have also been applied in MEC security and privacy preserving.

1) Bayesian Networks in MEC Security and Privacy: Bayesian networks, also named belief networks, are a kind of directed acyclic AI graph model, which are widely used in classification tasks [133]. To intuitively reflect the relationship between various vulnerabilities, the attack graph in the network can be modeled in IoT networks. Based on the attack graph, some previous works [134], [135] proposed to adopt Bayesian network model to find the correlation between vulnerabilities and network status, thereby providing security measurement for the network. However, the structure characteristic brings Bayesian networks with high computational complexity, which increases the burden of resource-limited edge severs.

2) Evolutionary Algorithms in MEC Security and Privacy: Evolutionary algorithms are also an important subset of AI, including the genetic algorithm (GA), evolution strategy, neuro-evolution, and so on. In the MEC network, the dependency graph that is used to characterize the dependencies between objects can be employed to represent the relationship between different malware. Based on this, Kim and Moon [107] proposed a malware detection model by GA, which turned the problem into searching the largest subgraph isomorphism problem in the dependency graph. Meanwhile, the complexity of the proposed model was greatly decreased by reducing the size of the dependency graph. Nevertheless, the disadvantages of evolutionary algorithms are also obvious. The manual programming is complex, which makes the evolutionary algorithms are not suitable in the large data set.

V. AI Approaches for Layer-Based MEC Security

In this section, we provide a hierarchical introduction to technologies involving AI approaches for protecting the MEC security. Furthermore, we provide a layer-based security solutions in Table II.

A. IoT System Security

1) Perception-Layer Security: Some attackers try to manipulate the device authentication or jam the radio frequency environment in the data collection progress to launch DoS attacks. Lightweight cryptography is widely used in IoT-enabled devices for authentication [144], [145], [146]. However, time-varying features in the network limit the performance of such traditional authentication algorithms in large-scale IoT systems. To alleviate these burdens, some research based on AI methods have been proposed to address these security threats, which are introduced in the following.

a) IoT device level: Physical-layer authentication (PLA) ensures that the connected device is not malicious. In PLA, physical-layer features, such as the received signal strength (RSS), channel impulse response (CIR), and channel state information (CSI) are adopted to detect spoof attacks [14], [71]. Aiming at the MEC environment, Chen et al. [136] proposed a PLA framework that combined clustering and traditional lightweight symmetric cryptography. Based on the advantages of unsupervised clustering approaches in the data without prior knowledge, this framework extracted and clustered CSI features, and then used the symmetric cryptography to achieve

TABLE II
SUMMARY OF AI-BASED WORKS AT EACH LAYER OF MEC FOR SECURITY

Layers	Works	Attack methods	AI-based methods					Results	Applications or scenarios
			Supervised learning	Unsupervised learning	Semi-supervised learning	RL	Non-ML methods		
IoT Systems	[136]	Spoofing attacks, replay attacks		Clustering				Authentication rate: 100%	Authentication
	[137]	Spoofing attacks	DNN					Authentication rate: 100%	Authentication, PUF
	[102]	DoS attacks				RL		Normalized accumulated reward: 2.25 msec	Anti-jamming
	[15]	DoS attacks, TCP/IP attacks	NB	PCA				Accuracy: 92.48% Detection rate: 95.35%	IDS
	[74]	DDoS attacks, MitM attacks	LR					Average accuracy: 74%	IDS
	[76]	DDoS attacks	DT				GA	None quantitative result	IDS
	[107]	Malware					GA	Detection rate: 88.89%	Malware detection
	[101]	Malware				Q-learning		Improved detection accuracy: 40%	Malware detection
	[103]	DoS/DDoS attacks, spoofing attacks				RL		None quantitative result	Edge caching
	[122]	Malware		DBN				Accuracy: 96.66%	Android malware detection
	[73]	Malware	SVM	DBN				Accuracy: 94.7%	Android malware detection
MEC System Level	[138]	Physical attacks	CNN	AE				PR-AUC: 99.20%	Anomaly detection
	[97]	DoS attacks, manipulate and misuse		AE				Accuracy: 95.4%	IDS
	[8]	DoS/DDoS attacks, eavesdropping and hijacking		DBN				Improved detection accuracy: 6%	IDS
	[89]	DoS/DDoS attacks, privilege escalation, eavesdropping and hijacking		K-Means				Overall redundant alarm compression rate: 97.2%	IDS
	[124]	DoS attacks, TCP/IP attacks, manipulate and misuse			ESFCM			Accuracy: 86.53%	IDS
	[139]	DoS attacks, manipulate and misuse	RF	K-Means				Accuracy: 96.03% FPR: 1.18%	IDS for SDN
	[140]	DoS attacks, manipulate and misuse	DNN					Accuracy: 75.75%	IDS for SDN
	[141]	DoS attacks, manipulate and misuse, TCP/IP attacks	RF					Precision: 96% Recall: 53.2% Accuracy: 97%	IDS for SDN
MEC Host Level	[18]	Eavesdropping and hijacking	MLP					Accuracy: 96.256%	Hardware Trojan Detection
	[100]	Manipulate and misuse			GAN			Precision: 95% Recall:20%	PUF
	[118]	DoS attacks,malware, physical attacks, eavesdropping and hijacking	LSTM					Accuracy: 99.91%	IDS
	[142]	Manipulate and misuse				RL		Robustness accuracy: 80%	Secure FL
	[143]	DoS/DDoS attacks, TCP/IP attacks	ELM					Precision: 98% Error rate: 1%	IDS
	[86]	DoS attacks, manipulate and misuse	CNN					Accuracy: 95.8%	Attacks recognition
	[92]	DoS/DDoS attacks, eavesdropping and hijacking	DNN	PCA				Accuracy: 99.9% Detection rate: 100%	Attacks recognition
	[115]	Malware	RNN					Accuracy: 93%	NFV malware detection

authentication of IoT devices. The proposed model significantly improved the authentication rate of the system with a low-complexity structure, and effectively combated spoofing attacks, replay attacks and small integer attacks.

However, the traditional ML-based methods require to be trained with a mass of samples, which occupies plentiful computationally resources and increases the time consumption. In response to this, Liao et al. [137] employed the data augmentation algorithm to accelerate the authentication process, and combined with DNN to enhance the accuracy of PLA and the model training efficiency. In the experiment, the author illustrated the proposed model attained a lightweight authentication with higher accuracy compared with the traditional threshold-based PLA method.

b) Raw data collection level: The raw data collected by sensors in MEC is usually transmitted through the cognitive radio network (CRN). Some attackers transmit false signals in the CRN to expend network bandwidth, which causes the received IoT devices to consume a mass of public resources. Taking the advantage of RL in game-based problems, Aref et al. [102] proposed an anti-jamming model based on multiagent RL. Multiple CRNs were allowed to transmit signals on the same frequency band, so each CRN was required to discriminate jamming signals and other normal CRN signals. Q-learning was used to generate the jamming-free subband frequency strategy, which was with lower complexity and achieved better performance in anti-jamming.

2) Network-Layer Security: Numerous communication protocols in the network layer can provide and manage the connectivity between sensors and edge servers. However, the open communication protocol makes it effortless for attackers to seek out the vulnerability during communication.

a) Data transmission level: An effective solution to anomalies in data transmission is to build an IDS to

automatically detect various attacks during communications. In a previous study [15], PCA was used to reduce the dimension of the original data and represent it as new data with the minimum attribute through the principle component load matrix, and the improved NB classifier was utilized to classify network attacks. Another work [147] applied the replacement missing value filter to the feature extraction and utilized the RF classifier to classify attacks with network traffic data. Whereas, the index of classification accuracy alone is not enough to objectively evaluate the performance of the IDS classifier. It needs to be combined with indexes, such as precision, recall rate, and false detection rate to make a comprehensive measurement [15], [82], [147].

DoS attacks are the most common type of intrusion in the network layer. In response to this, a previous research [74] adopted LR-based classifiers to detect DDoS attacks and MitM attacks. This work utilized smartphones as edge servers in IoT networks for training. Using the power consumption ratio information of attacking devices and nonattacking devices to predict intrusion, the proposed model achieved a relatively ideal accuracy rate. Another study [76] proposed an intelligent anomaly detection architecture based on GA and the DT classifier. The architecture detected illegal links based on the IP packet header information.

b) Offloading and caching level: Ensuring the safety of the offloading and caching process is an important task for maintaining the network-layer security. A previous research [103] extended RL to the MEC environment to protect the security of edge caching. In response to attacks in the ME cache, mobile devices adopted RL to make the optimal decision from a limited set of actions to protect the offloading process from interference or perform device authentication. This work proposed a security framework for MEC offloading with RL-based methods, which has enlightening significance for protecting the security of edge offloading and caching.

3) Application-Layer Security: The application layer directly provides users with intelligent services. However, the inherent vulnerabilities in the operating system (OS) of smart device pose security threats to the application. In a previous research [122], an Android malware intelligent detection system named DroidDelver based on the DBN model was introduced. In this work, API calls were extracted from the smali code, and different API csalls were categorized into blocks according to their functions. Based on the generated API call blocks, DBN learned the relationship between benign software and malware to detect malware in the system with improved accuracy. Xu et al. [73] proposed a hybrid analysis malware detection system of Android, which extracted features through DBN and combined them with original features to construct a vector set for classification. SVM was utilized as a classifier, and the kernel matrix constructed by the similarity between features was applied for classifying the benign software and malicious applications.

B. MEC System-Level Security

MEC system level provides UE and CFS Portal with available edge sever services. In this section, we summarize the AI-based approaches in MEC system level in the following two aspects.

1) End-Edge Level: The collaboration working mode of mobile ends and edge servers provides an efficient solution for intrusion detection in UEs and ANs. In a previous work [138], the Deep AE model was used for anomaly detection in the industrial IoT network in an end-edge collaborative mode. This work applied a CNN-variational AE model to extract features from the time series state information of sensors and perform anomaly detection. What is more, the memory and computing consumption deployed on the edge sever were reduced by compressing the neural network size. In another research [97], an IDS architecture based on the Deep AE was applied to detect abnormal traffic from sensors to edge servers. In order to improve the accuracy of classification, the model utilized isolation forest to further classify the output of the Deep AE model to find the points of classification errors. This optimized detection architecture has an enlightening effect on reducing the false alarm rate.

The AN between the edge sever and the mobile end is also with potential vulnerabilities to several attacks. Based on previous works, Chen et al. [8] proposed a network attack detection system based on DBN in MEC. The dynamic features were extracted by exploiting the Android file package in the edge server and used for the model training. The loss function was established according to the difference between inputs and the actual outputs, and the loss was reduced by using the BP algorithm to fine-tune the parameters of the neural network.

Moreover, a mass of high-dimensional and multisource alarm information is gathered in the edge server [89], which is a huge challenge for the processing and storage capabilities of the serve Li et al. [89] proposed an attack detection method based on the end-edge framework. The alarm information of the edge equipment was de-redundant through the *K*-Means method. The authors established attribute attack graphs through the correlation between the alarm information to comprehensively analyzed to the potential vulnerabilities of all edge nodes. A greedy decision-making algorithm was utilized to solve the problem of the minimum dominance set in the attribute attack graph to replace the generation of the complex attack linkage strategy. Given these points, this work provides ideas for the feature extraction of high-dimensional and multisource alarm information in the end-edge level and the consideration of complex attack relevance.

2) SDN Level: The separation of data plane and control plane in SDN also brings various security threats. In a previous study [139], an IDS in SDN based on the combination of *K*-Means and an improved RF classifier was proposed. The bat algorithm was used for feature selection, where the position of bat was set as the selected feature subset, the iteration of position movement, and target search were summarized as the feature selection process, and the fitness function was adopted to evaluate the feature selection process. After the selected features were clustered by *K*-Means, the flows were classified by using the RF algorithm based on the weighted voting mechanism. The proposed model achieved satisfactory classification accuracy due to the advanced feature selection method while reducing the training cost. Tang et al. [140] applied DNN to

build an abnormal flow detection system in SDN, which was deployed in the control plane. It monitored the data of all OpenFlow switches and utilized the global network status to detect intrusions. The flow table was modified to propagate the security policy to the switch when detecting intrusions. The DNN model of this system was trained through information, such as protocol, duration, and flow bytes as features and achieved improved classification accuracy.

Kirutika et al. [141] proposed an external monitor based on the RF algorithm for possible attacks in the SDN control plane. The number of incoming/outgoing data packets and lost packets, duration, and other network state information were utilized as features to train the RF classifier. The external monitor detected the malicious data and interrupted the controller timely. Consequently, this work ensured the security of SDN from the control plane to prevent the entire SDN system from being corroded by attackers.

C. MEC Host-Level Security

MEC host level is mainly responsible for managing various functions of the host, as well as the collaboration with other hosts and the cloud. We introduce the AI-related works in the MEC host level from the above aspects in the following.

1) Edge–Host Level: The edge host security directly affects the robustness of the overall MEC architecture. Khalid et al. [18] proposed a hardware Trojan detection framework in edge hosts by multilayer perceptrons (MLPs). The single power-port current acquisition block was designed in the time-division multiplexed current sensor to reduce the cost of data acquisition. Through the analysis of hardware Trojan benchmarks at the register transfer level in the System on Chip (SoC), four LEON3 processors from the other infrastructure providers were integrated to provide the solution. The proposed model greatly improved the detection accuracy of hardware Trojan compared with the existing methods, and reduced the power consumption on per unit area of the chip. In [100], a GAN-based self-adversarial agent model was proposed to improve the hardware security of edge hosts. The agent utilized vanilla GAN and conditional GAN, respectively, to attack edge hosts, and the public PUF was used to evaluate the quality of the attack by generating realistic secret keys. The agent reconstructed its underlying security primitives into the public PUF through feedback to improve the security entropy of the system when the attack quality exceeds the setting threshold.

2) Edge–Edge Level: The secure collaboration mode between edge hosts is crucial for MEC host level. In a research [118], LSTM was utilized to detect cyber attacks in the sensors network through the Fog-to-Things architecture. The entire architecture assigned attack detection tasks to edge hosts, then cyber attacks in the cover area were detected through the coordination of parameter update, storage and control between each edge host. In each edge host, the sequential stochastic gradient descent (SGD) algorithm was adopted to calculate local parameters of the LSTM. And the coordination host was responsible for aggregating local parameters from all edge hosts, then updating and returning the global parameters back for updating the local model.

Pan et al. [142] proposed a gradient aggregation agent (GAA) model suitable for the MEC environment against Byzantine attacks, using RL to protect the robustness of the distributed learning framework. The proposed GAA model learned the experience from the interaction among workers and the auxiliary information in the master, and decided the contribution weight of its generated parameters to the overall parameters according to the credit of the worker. For different aggressive environments, the robustness of the GAA model proved its performance in MEC multiedge collaborative level.

3) Edge–Cloud Level: The edge–cloud collaboration mode solves the problem of insufficient edge device resources. In [143], ELM was employed to detect attacks in IoT networks. Features about duration, IP, source and destination port numbers were converted into random multidimensional space vectors. However, plenty of features occupied the memory and the model training process generated huge energy consumption in the edge sever. To achieve an efficient detection architecture, edge devices were used to randomly project raw data to generate private data and upload the data to the cloud with sufficient resources for model training.

Ran et al. [86] proposed a CNN-based edge–cloud collaborative attack recognition architecture. The initial data sets of multiple edge hosts were set to the same. Then, edge hosts were responsible for the raw data preprocessing and feature selection, and utilized the trained CNN locally for attack detection. The edge host with the highest detection rate uploaded its own data to the cloud in a fixed period of time, and the cloud was in charge of updating the data set and retraining the CNN and sending the improved model to each edge host. This edge–cloud collaborative learning mode improves the detection accuracy and the robustness of the system. In addition, in order to release the workload of the cloud, AI algorithms (such as PCA [92] and Deep AE [97]) can be deployed in edge hosts to prelearn data, and perform secondary learning through the cloud to fully improve performance.

4) NFV Level: NFV provides resource virtualization services for edge hosts. Guizani and Ghafoor [115] proposed an anti-malware system based on NFV, which used RNN to predict malicious software so that NFV can deploy relevant resources in a timely manner to resist attacks. The raw data collected by the sensor was preprocessed and evaluated in terms of available functions, and feature reduction and data cleaning were employed to filter valuable features. Then, the system applied the RNN-LSTM model to learn the preprocessed data, and utilized NFV to virtualize the patch distribution mechanism.

VI. AI Approaches for Layer-Based MEC Privacy

Traditional IoT network privacy protection methods are implemented through cryptography or steganography [155], [156], [157], [158], while advanced AI technology gives more possibilities for privacy-preserving. In this section, we give a hierarchical introduction to AI-based works related to the privacy in the MEC environment. Related works are summarized

TABLE III
SUMMARY OF AI-BASED WORK AT EACH LAYER OF MEC FOR PRIVACY

Layers	Works	Privacy issues	Related AI methods					Results	Applications or scenarios
			Supervised learning	Unsupervised learning	Semi-supervised learning	RL	Non-ML methods		
IoT System	[90]	Location privacy		K-Means				None quantitative result	WSN privacy
	[148]	Data privacy	CNN					Accuracy: 98.27% Encryption runtime: 0.344 s Decryption runtime: 0.056 s	Mobile sensing privacy
	[69]	Data privacy	kNN					Encryption runtime: 0.13 s	Encrypted cloud database
	[87]	Data privacy	CNN					Error detection rate: 99% Encryption runtime: 0.012 s Decryption runtime: 0.025 s	Offloading privacy
	[149]	Data privacy, location privacy, identity privacy		AE				Caching efficiency: 15%	Caching privacy
	[91]	Data privacy, location privacy		K-Means				Runtime: 1209 ms Communication overhead: 1085KB	Caching privacy
	[150]	Identity privacy	CNN					Accuracy: 97%	Blockchain
	[151]	Data privacy			GAN			Accuracy: 96%	Video streams privacy
MEC System Level	[152]	Position privacy		Skip-gram				Identification rate: 83%	Anti-poisoning attacks
	[153]	Data privacy	DNN					Forward pass time: 0.6s Backward pass time: 0.2s	Facial recognition privacy
	[78]	Position privacy	DTs, kNN					Position confidentiality assurance: 90%	LBS privacy
MEC Host Level	[154]	Data privacy	DNN					Accuracy: 94.23% Compression ratio: 8.77%	FL privacy
	[77]	Data privacy	DTs					None quantitative result	Edge-cloud privacy
	[93]	Data privacy		PCA				None quantitative result	Edge-cloud privacy
	[70]	Data privacy	kNN					None quantitative result	Querying encrypted data

in the same hierarchical manner as Section V and listed in Table III.

A. IoT System Privacy

1) Perception-Layer Privacy: In the raw data collected by sensors, the location privacy is of great importance. Han et al. [90] proposed a location privacy-preserving framework based on *K*-Means in the IoT network. Fake source nodes were established to transmit packets from real source node, so as to make the attacker confused about the real location of the source node. And fake sink nodes were utilized to conceal the real route through generating fake packets and transmitting them in different directions. In this work, the public and private keys were applied in each node for authentication, and only the real packet was able to be transmitted to the real sink node through the specific route.

As the subject of the perception layer, the privacy protection of the sensor is also of the essence. In a previous research [148], a lightweight privacy-preserving framework of the mobile sensor was proposed based on CNN. In order to reduce the computing overhead, storage and communication costs on mobile terminals, this work transferred data processing tasks from the cloud to the edge server. The CNN model was publicly available to each edge server for feature extraction. The sensor randomly divided the collected image data into two parts for encrypting, which can be restored by superimposing them together. The two encrypted divided parts were, respectively, used for feature extraction and learning of two edge servers, and the interaction of the two edge servers was controlled by a trusted third party.

2) Network-Layer Privacy: In this section, we introduce the AI-enabled methods of privacy preserving from the following aspects.

a) Offloading level: ML-based inference tasks on mobile devices are often offloaded to the nearby edge server instead of the cloud sever, which greatly reduces the communication pressure and computational burden. In [69], kNN was used to construct a lightweight encrypted cloud database architecture with edge offloading. In this work, data owners uploaded data to the cloud server, data users sent query requests to the cloud, and the cloud utilized kNN for classifying. Pailier encryption system was employed to encrypt data and labels to provide database security, query privacy, and data access mode hiding. In another work, Tian et al. [87] proposed a lightweight scheme to protect the privacy of the CNN-based inference task in the offloading process. Multiple pairs of encryption and decryption keys were generated and stored in the IoT device in an offline form, and each pair corresponded to a CNN inference task request. In the inference phase, the IoT device offloaded the CNN training task to the nearby edge server online, and the CNN model provided data integrity checks to ensure the correctness of the data.

b) Caching level: Edge caching provides additional resources for mobile devices to relieve storage pressure. Yu et al. [149] proposed a proactive content caching method based on federated learning (FL), which performed training with no require of the centralized collection user data. The method used a centralized server to aggregate parameters of distributed edge servers for updating, and each edge server utilized Deep AE for local training. In the absence of user data, the performance of this model was still better than other methods, and it played a vital role in protecting user privacy. In [91], a privacy-preserving method based on *K*-Means was used in the caching process of the next generation cellular network. Through the FL system, a distributed architecture was adopted to upload user data instead of encrypted user data

to the cloud server for training based on the SGD method, so that avoiding data leakage caused by directly uploading data.

3) Application-Layer Privacy: Blockchain is a well-known shared database in the application layer. Due to the data in blockchain cannot be tampered, it can be applied for the privacy-preserving of user information. Zhao et al. [150] proposed a CNN-based FL system combined with blockchain to provide a user data privacy protection model for home appliance manufacturers. First, the user's smart phone collected data from home appliances, and conducted local-level training of the model in a collaborative manner with the edge server. Then, users signed the trained model and uploaded the model to the blockchain to protect it from tampering. In addition, the manufacturer used the DP technology to calculate an average model for the models collected from users and preserve their privacy.

In addition, video streaming and analytics (VSM) is also a popular application in IoT system. In order to protect the user privacy contained in the required data in the VSA, Wu et al. [151] proposed an enhanced privacy protection system. Using the steganography of the GAN model at the front end, the system was able to reverse the privacy-enhanced video without changing the back end without auxiliary data. In addition, by combining with system optimization technology, the system fully reduced the network bandwidth and realized efficient real-time processing on the hardware.

B. MEC System-Level Privacy

Aiming at protecting edge servers from privacy damage caused by malicious users poisoning attacks, Zhao et al. [152] proposed a privacy-preserving model based on a feature learning model named Skip-gram. In the edge server, the social relationship between different users was extracted and an inferred social graph was established. The model predicted the location of the poisoning through the best map between the social graph and the social graph. In other work, Mao et al. [153] proposed a privacy-preserving model of DNN facial recognition based on DP mechanism. In order to reduce training costs, the DNN model was partitioned from the convolutional layer, where the first part was deployed on the user side and the second part was arranged on the edge server side. The output of the user-side model was input to the edge server, and the calculated loss information was fed back to the user side to update the gradient.

Moreover, in a previous research [78], a combined approach based on DTs and kNN was used to identity the user's position, and the HMM was applied to estimate the user's destination and location tracking sequence. This approach was implemented in the MEC environment to ensure the timeliness and confidentiality of the delivery of LBS and provide privacy-preserving LBS for roaming users. *G*-Means clustering method was adopted to extract effective location features from the redundant information of the device, and DTs and kNN were merged to solve the position tracking sequence ordering problem. Then, the edge sever utilized excessive locations to predict the location through HMMs.

C. MEC Host-Level Privacy

In the MEC host level, related privacy-preserving research focuses on the realization of two collaborative levels: 1) edge–edge and 2) edge–cloud.

1) Edge–Edge Level: Lu et al. [154] proposed a joint learning mechanism based on FL to provide privacy protection support for joint work scenarios of multiple edge servers. Edge servers were used to train the ML-based model, and the parameter server was responsible for collecting parameters of all edge servers and updating global parameters. In this work, the parameter privacy of the entire model was protected because the edge server only communicated with the parameter server and had no authority to obtain the information from other edge servers. Meanwhile, in order to improve the training efficiency of the model, a gradient sparsification method was adopted to compress the interaction between edge servers and the parameter server so as to reduce the communication consumption and obtain effective compression space at a tiny cost of accuracy.

2) Edge–Cloud Level: The edge–cloud hybrid architecture is an efficient privacy-preserving framework. In a research [77], a privacy-preserving mechanism in the edge–cloud environment based on DTs was proposed. On account of the DP algorithm, the edge server as the collector of the adversary data set was required to not distinguish the difference from the normal data set. In this mechanism, the cloud server was used to build a private random DTs model, and the edge server was responsible for collecting data and adding it to the random DTs.

In another work, Osia et al. [93] proposed an edge–cloud hybrid privacy-preserving framework that used edge server resources to reduce the cloud processing latency and improve the processing performance. The framework consisted of two modules. The privacy edge module was responsible for extracting features from the original image data, and the cloud server module was in charge of inferring and feeding back to users. In the privacy edge module, a Siamese privacy framework was adopted, and the PCA model was employed to reduce the dimension of the data to extract features and ensure the accuracy of feature extraction through fine-tuning. In the cloud server, the CNN model was utilized to perform inference based on the extracted low-dimensional features, and the use of low-dimensional features instead of the raw user data to transfer to the cloud also ensured the secure user data privacy.

VII. Lessons Learned, Future Works, and Challenges

In this section, according to the security and privacy concerns and related AI-based solution, we summarize specific research problems and propose promising future directions to give our insights to future researchers.

A. AI Methods for MEC Security

The integration of various complex heterogeneous technologies in MEC makes it face multitudinous threats. Additionally, although powerful AI methods can provide efficient classifiers and feature extraction tools for MEC security protection, the deployment of these methods in the MEC environment is also

extremely challenging. In this section, we summarize the existing research issues and future directions in using AI methods to solve various security threats in MEC.

1) IoT System Security: The heavily deployed IoT networks provide MEC with a wide range of interconnection and low-latency applications, while also brings the following challenges: 1) attackers can steal sensitive information from IoT devices in collecting raw data through side channel attacks, or directly perform physical attacks to disrupt the normal operation of IoT devices, and the extensively applied gateway devices also increase the risk of TCP/IP attacks and 2) for the reason that the application layer of IoT directly provides users with related services, it is also indispensable to perform effective malware detection in IoT networks.

a) General IDS architecture for enhancing trust-based approaches in ad hoc networks: Research Problems: Existing IDS for the security of wired networks could be used in wireless contexts. However, the intrinsic characteristics of MANET may limit their application, such as the absence of centralized infrastructure, limited bandwidth, mobility of the nodes, etc.

Future Directions: Considering the intrinsic characteristics of MANET, a general IDS architecture should have the following properties in the future.

1) A self-defence mechanism is of immense importance for defending repeated false alarms by sending a flood of irrelevant packets to the IDS host.
2) It should require little system resources to execute and not hinder system performance by adding overhead.
3) It should run continuously and keep up transparent to the system and the users.
4) The IDS should abide by standard to be cooperative and open, such as the standard alert format intrusion detection message exchange format (IDMEF) and a protocol for transporting such alerts intrusion detection exchange protocol (IDXP) [159].

b) Efficient malware detection tools in IoT networks: Research Problems: The malware detection tools can effectively prevent threatening objects in the IoT network from destroying software-level security. However, various widely popular applications in the practical IoT environment accelerate the evolution of malware, which greatly limits the performance of traditional detection methods, and cannot guarantee the false alarm rate for benign system files.

Future Directions: Future research on malware detection should mainly focus on combining traditional malware detection tools with powerful AI models. The key to detection lies in heuristic features, such as file properties, code fragments, and file hashes that distinguish benign from malware. Therefore, some DL-based natural language processing (NLP) models can be used to extract and process various feature information in OpCode, and construct malware detection as a binary classification problem. In addition, in order to adapt to the continuous emergence of system files and malware in the IoT network, the dependency graph can be used to express the relationship between software, and solve the state-of-the-art graph neural network (GNN) model to extract features and graph reduction operations to reduce the amount of computation.

c) MEC-IoT security protocols with ML: Research Problems: The intelligent transportation system (ITS) is a typical application of MEC-IoT systems, which consists of advanced sensors and control systems. ITS relies on the interconnectivity of various devices to process real-time data flow and transmit it to assure secure and efficient digital services. Such data flow is plaintext that is prone to eavesdropping and hijacking. Unmanned aerial vehicles (UAVs) are another example in which the aim is to conserve battery life while offloading computational or storage information to MEC servers for processing. As a result, using strong cryptographic primitives or lengthy security processes would be impossible [160]. The recent works either consider security protocols or ML technologies to secure the MEC-IoT systems. However, an elaborately designed method by aligning security protocols with ML can be more effective in protecting the security of MEC-IoT systems.

Future Directions: Although the ML-based IDS can cope well with abnormal data traffic, the required massive data collection in the interconnective IoT systems raises privacy concerns during both the training and prediction stages. One promising solution is secretly sharing the data with lightweighted cryptography protocols and evaluating the data with three-party computation. However, there are several research challenges to be solved in the future. First, the three-party computation is only suitable for computation over a $\mathbb{Z}_{2^k}$ ring. But both the training data and intermediate parameters of ML are decimal values that are unable to handle modular arithmetic. Second, secret sharing is costly and quickly becomes a major performance bottleneck in resource-constrained MEC-IoT systems.

2) Mobile Core Network Security: The mobile core network is the provider of the main functions of MEC. Although the in-depth modules of the MEC system (e.g., MEO) are difficult to be invaded by attackers, manipulated and misused attacks can inject fake data into the authentication module to disrupt the normal operation of the edge server. Moreover, threats in SDN are also hot research topics in the mobile core network.

a) Ensuring reliable end-edge connections in AN: Research Problems: As mentioned above, the end-edge reliable connection constitutes the first barrier to the AN security. In addition to the rational use of service request authentication provided by modules, such as OSS and MEO at the MEC system level, AI-based IDS should also be combined to protect the connection from being disrupted.

Future Directions: In the future, researchers can focus on choosing optimal AI models and features for detecting attacks between ends and edges. In order to obtain the sensitive information uploaded by mobile ends to the edge server, attackers usually inject malicious data packets into the end-edge connection. Therefore, IDS at the end-edge level should focus on the analysis of network traffic, analyze network dynamic information from the captured data packets, and take countermeasures when abnormalities are detected. In order to perform efficient feature extraction on large-scale data streams, DL models (such as DNN and DBN) can be used as favorable tools for building IDS in the future.

b) Detecting and mitigating DDoS attack on SDN control plane: Research Problems: SDN offers unparalleled programming which enables network administrators to dynamically customize and control their networks within the MEC environment. One of the security concerns is DDoS attacks which drain the network capacity of SDN control plane by sending heavy traffic.

Future Directions: Although the advantage of the SDN control plane is that it can get the global view of the entire network, the control plane is insufficiently scalable to support high-frequency flow requests. A promising solution is to leverage the scalability and easy customization of virtualized software functions and adopt appropriate ML technologies and rule-based schemes to safeguard the centralized SDN control plane.

3) Mobile Host-Level Security: When considering MEC security, the multifaceted threats cannot be ignored: 1) VMs composed of various heterogeneous technologies may suffer from infected VM images, compromising VM migration, VM hopping, VM escape, and VM DoS attacks; 2) there are specific threats in NFV MANO when adopting the NFV technology to encapsulate the MEC function into VMs; and 3) the hypervisor responsible for managing VMs may have traditional TCP/IP attacks, eavesdropping and hijacking attacks, and incur VM hopping, VM escape, and VM DoS.

a) Coping with inherent threats in VMs: Research Problems: The various inherent threats in VMs are open challenges, which bring risks to taking advantages of VM to virtualize edge network resources. Consequently, coping with inherent threats in VMs is getting more attention from researchers.

Future Directions: In order to deal with infected VM images, future researchers can enhance the security of VM images by combining traditional encryption algorithms and the firewall technology. For compromising VM migration, future research should also focus on the encryption of the plaintext transmitted during the migration process, and it is also meaningful for intrusion detection in the migration channel. For VM hopping, future countermeasures may be to deploy effective authentication algorithms in MMU to prevent malicious intrusion. And for VM escape, a valuable research direction is to enhance the security of the VMM and maintain the management authority of the VM by integrating additional monitors. In addition, for the above threats as well as VM DoS, an effective potential solution is to monitor through the AI-based IDS. For example, some commonly used lightweight AI classifiers, such as kNN, RFs, and NB can be adapted to the IDS for VMs.

b) MEC secure resource management by NFV: Research Problems: The key function of NFV is to provide flexible resource management services for the network. In order to safely use NFV to manage MEC resources, the threats in the VM must first be resolved, which has been explained in detail above. And it is also a promising research problem to solve the threats of NFV.

Future Directions: In the future, we can regard for implementing NFV in MEC by adopting DRL that has been proven to perform well in the wireless resource allocation game. The security of the NFV interface should be enhanced to counter the threats from the NFV MANO. In addition, an adaptive and powerful hypervisor can be established, and IDS based on ML approaches (e.g., DNN, K-Means, DBN, etc.) with powerful feature extraction ability can be embedded in it to detect VM threats and possible TCP/IP attacks as well as eavesdropping and hijacking attacks in the future.

c) Developing intelligent IDS by NFV: Research Problems: Generally, AI-based IDS consumes a lot of computing resources and occupies a certain amount of storage space, which makes the deployment in resource-constrained edge devices challenging. Therefore, adopting NFV to provide flexible resource management for intelligent IDS is a meaningful research topic.

Future Directions: Future researchers can consider using the flexible resource management framework brought by NFV to assist in the realization of delay-sensitive and computationally intensive attack detection and defense tasks. In order to improve the emergency response speed of IDS, future research directions can adopt the NFV technology to deploy the required resources in time to counter attacks. For different types of attacks, the corresponding defense methods (e.g., configuring system patches and performing reasonable security domain partition) can be virtualized, and the bound virtualization function can be directly deployed to the abnormal area when a certain type of attack is detected by the intelligent IDS.

4) Mobile User Security: The security of mobile users' devices is also a challenging subject. The possible threats on the UE side include malware and hijacking, DoS/DDoS attacks, and TCP/IP attacks, which bring a new research perspective to the security protection of MEC.

a) Ensuring UE hardware-level security: Research Problems: Ensuring the mobile users security from the hardware level of devices will also become an enlightenment for the future work. PUF technology embedded in mobile devices is attracting attentions. It can provide identity authentication for devices through electronic circuits. Unfortunately, advanced nonintrusive attacks [75] have been able to clone the PUF in MEC, which greatly threaten network security.

Future Directions: It is a prospective research direction to utilize ML method with great performance in classification to recognize the cloned PUF. In addition, deploying PUF on the chip also faces the dilemma of a tradeoff between performance and energy consumption. Future researchers can devote themselves to applying ML methods to generate the optimal PUF configuration for different application scenarios.

B. AI Methods for MEC Privacy

As mentioned in Section III, the open MEC ecological environment has introduced multiparty infrastructure providers, which makes MEC services face the following privacy issues: 1) tasks offloading makes user's sensitive information may be stolen or tampered in the communication link during the offloading process; 2) the LBS brings a great test to MEC's privacy-preserving; and 3) users' highly sensitive PII arouses subscribers' concerns about MEC.

1) Preserving Data Collection Privacy (Research Problems): In IoT networks, raw data collected by a

large number of sensors urgently require to be uploaded to edge nodes for processing to offer corresponding services. However, the existing communication technologies in sensor networks (e.g., Bluetooth, WiFi, NFC, etc.) cannot provide users with completely trusted communication links. Therefore, protecting the privacy of sensitive information contained in the raw data is a meaningful research problem.

Future Directions: In the future, an appropriate research direction is to use the DP technology between sensors and edge nodes to add random disturbances to the raw data for encryption without affecting the task effect. The end-to-end encryption technology HE can also provide researchers with a solution to protect the privacy of the raw data, which has a lower computational complexity. In addition, it is also an effective privacy-preserving solution to divide the raw data into multiple parts at the sensor, and upload encrypted parts to multiple edge nodes for further processing. It is worth noting that this solution requires a trusted third party to coordinate the collaboration between different edge nodes, and it requires careful design by future researchers.

2) Ensuring LBS Privacy (Research Problems): LBS is a basic service of IoT, and the location information is also required for multiple services. As mentioned in Section III, the user's location involves key personal information, and the current smart devices make it easy to leak location privacy. More importantly, the leakage of location privacy in some application scenarios that rely on location information (such as VANET) will cause unpredictable consequences. Therefore, the LBS that provides privacy guarantees in the MEC is imminent.

Future Directions: Future research directions can focus on using encryption technologies, such as DP and HE to encrypt user location information in UE. In addition, because the routing information of the network also contains location privacy, attackers can trace the source according to the routing information. Therefore, future researchers can devote themselves to developing methods to encrypt routing information, such as adding fake nodes and injecting fake source data to ad hoc networks to confuse attackers about the real route.

3) Privacy-Preserving Approaches for Edge Offloading and Caching (Research Problems): Edge offloading and caching are key functions of the MEC architecture, which enables mobile users to obtain nearby resources to handle delay-sensitive and computationally intensive tasks. A large amount of data containing sensitive information is offloaded or cached on the edge server, which brings numerous privacy issues.

Future Directions: Future researchers can devote themselves to the research of edge encryption databases, and store user information that is offloaded or cached to edge servers with guaranteed privacy. And some ML models (such as kNN, CNN, SVM, etc.) can be selected as the classifier of the encrypted database to classify according to the popularity of the content. The encrypted data and labels will provide reliable query and access services for edge offloading and caching. In addition, privacy-preserving research for some classic edge offloading and caching applications (for example, video streaming [161]) is also a promising research direction.

C. Enhanced Approaches in MEC Security and Privacy

In addition to utilizing AI approaches to solve the security and privacy issues in MEC, how to enhance the performance of AI-based methods are also with great promise.

1) Integrating AI and Third-Party Technology (Research Problems): Lightweight edge servers have less abundant computing power and storage than cloud servers with centralized resources. Therefore, it is also an important challenge to deploy computationally intensive AI models at the edge reasonably, which determines the practicality of the research model. Especially for the DL model, the training of the neural network requires a large amount of computing resources, and this is overloaded for a single edge server. Therefore, the DL model deployment scheme based on partition training [162], [163] and neural network hardware acceleration technology [113], [114] has become a trend.

Future Directions: The pivotal challenge of adopting partition training is how to adaptively divide the neural network into multiple partitions for training on multiple edge servers. And for the neural network hardware acceleration technology, the main methods include matrix decomposition, pruning, layer reduction, etc., which are mainly based on the idea of scale compression. The main challenge is to select an appropriate scale compression method for the model, and combine it with partition training to reduce model complexity and storage space while assuring the accuracy. The deployment of neural networks at the hardware level is also an attractive topic. Research [164], [165] of using memristors to realize DNN will provide basic support for improving the performance of the neural networks on MEC severs.

VIII. Conclusion

MEC is becoming the main computation paradigm with its lightweight and efficient architecture. In order to duly extract valuable information from a mass of raw data and make relevant decisions, AI is deployed in MEC to provide intelligent data-related services. As the security and privacy issues attract more attention, AI-based technologies play a crucial role in protecting the security and privacy in the MEC environment.

This survey introduces the MEC architecture from a holistic perspective and separates the layers from the MEC-enabled IoT System and edge sever system levels. For the functions of each layer, related security, and privacy threats are explained in detail. After that, the methods related to MEC security and privacy are summarized from the perspective of AI, the ML-based methods are classified and mainly explained, and the inherent challenges in ML are also discussed. In addition, the non-ML AI approaches for MEC security and privacy are also summarized. Then, the related MEC security and privacy-preserving AI-based works are systematically discussed in each layer. Finally, this survey summarizes the ideas and challenges of existing works from the three aspects of AI-based methods, MEC framework, and security and privacy challenges in AI, and envisages future research directions. This survey aims to provide researchers with an overview of MEC security and privacy preserving from the perspective of AI, and to encourage the development of related research.

REFERENCES

[1] 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.

[2] Y. C. Hu, M. Patel, D. Sabella, N. Sprecher, and V. Young, "Mobile edge computing—A key technology towards 5G," ETSI, Sophia Antipolis, France, White Paper, 2015.

[3] T. Ouyang, Z. Zhou, and X. Chen, "Follow me at the edge: Mobility-aware dynamic service placement for mobile edge computing," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 10, pp. 2333–2345, Oct. 2018.

[4] S. M. A. Huda and S. Moh, "Survey on computation offloading in UAV-enabled mobile edge computing," *J. Netw. Comput. Appl.*, vol. 201, May 2022, Art. no. 103341.

[5] H. Zhou, Z. Zhang, Y. Wu, M. Dong, and V. C. Leung, "Energy efficient joint computation offloading and service caching for mobile edge computing: A deep reinforcement learning approach," *IEEE Trans. Green Commun. Netw.*, vol. 7, no. 2, pp. 950–961, Jun. 2023.

[6] 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.

[7] B. Bracken. "Cyberattacks on healthcare spike 45% since November." Accessed: Aug. 3, 2021. [Online]. Available: https://threatpost.com/cyberattacks-healthcare-spike-ransomware/162770/

[8] Y. Chen, Y. Zhang, S. Maharjan, M. Alam, and T. Wu, "Deep learning for secure mobile edge computing in cyber-physical transportation systems," *IEEE Netw.*, vol. 33, no. 4, pp. 36–41, Jul./Aug. 2019.

[9] K. Benton, L. J. Camp, and C. Small, "Openflow vulnerability assessment," in *Proc. 2nd ACM SIGCOMM Workshop Hot Topics Softw. Defined Netw.*, 2013, pp. 151–152.

[10] M. Pattaranantakul, R. He, Q. Song, Z. Zhang, and A. Meddahi, "NFV security survey: From use case driven threat analysis to state-of-the-art countermeasures," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 4, pp. 3330–3368, 4th Quart., 2018.

[11] A. D. Joseph, P. Laskov, F. Roli, J. D. Tygar, and B. Nelson, "Machine learning methods for computer security (Dagstuhl perspectives workshop 12371)," in *Dagstuhl Manifestos*, vol. 3. Wadern, Germany: Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2013, pp. 1–30.

[12] M. I. Jordan and T. M. Mitchell, "Machine learning: Trends, perspectives, and prospects," *Science*, vol. 349, no. 6245, pp. 255–260, 2015.

[13] "AV-TEST: Malware." 2022. [Online]. Available: http://www.av-test.org/en/statistics/malware/

[14] N. Wang, T. Jiang, S. Lv, and L. Xiao, "Physical-layer authentication based on extreme learning machine," *IEEE Commun. Lett.*, vol. 21, no. 7, pp. 1557–1560, Jul. 2017.

[15] S. Manimurugan, "IoT-fog-cloud model for anomaly detection using improved Na¨ıve Bayes and principal component analysis," *J. Ambient Intell. Humanized Comput.*, vol. 2021, pp. 1–10, 2021.

[16] G. Zhao, C. Zhang, and L. Zheng, "Intrusion detection using deep belief network and probabilistic neural network," in *Proc. IEEE Int. Conf. Comput. Sci. Eng. (CSE) IEEE Int. Conf. Embedded Ubiquitous Comput. (EUC)*, vol. 1, 2017, pp. 639–642.

[17] M. Roopak, G. Y. Tian, and J. Chambers, "Deep learning models for cyber security in IoT networks," in *Proc. IEEE 9th Annu. Comput. Commun. Workshop Conf. (CCWC)*, 2019, pp. 452–457.

[18] F. Khalid, S. R. Hasan, S. Zia, O. Hasan, F. Awwad, and M. Shafique, "MacLeR: Machine learning-based runtime hardware trojan detection in resource-constrained IoT edge devices," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 39, no. 11, pp. 3748–3761, Nov. 2020.

[19] A. Libri, A. Bartolini, and L. Benini, "pAElla: Edge AI-based real-time malware detection in data centers," *IEEE Internet Things J.*, vol. 7, no. 10, pp. 9589–9599, Oct. 2020.

[20] W. Shi, J. Cao, Q. Zhang, Y. Li, and L. Xu, "Edge computing: Vision and challenges," *IEEE Internet Things J.*, vol. 3, no. 5, pp. 637–646, Oct. 2016.

[21] A. Singh, S. C. Satapathy, A. Roy, and A. Gutub, "AI-based mobile edge computing for IoT: Applications, challenges, and future scope," *Arab. J. Sci. Eng.*, vol. 47, pp. 9801–9831, Jan. 2022.

[22] M. Satyanarayanan, "The emergence of edge computing," *Computer*, vol. 50, no. 1, pp. 30–39, Jan. 2017.

[23] B. Varghese, N. Wang, S. Barbhuiya, P. Kilpatrick, and D. S. Nikolopoulos, "Challenges and opportunities in edge computing," in *Proc. IEEE Int. Conf. Smart Cloud (SmartCloud)*, 2016, pp. 20–26.

[24] 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.

[25] F. Giust et al., "MEC deployments in 4G and evolution towards 5G," ETSI, Sophia Antipolis, France, White Paper, 2018.

[26] K. Peng, V. 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, Oct. 2018, Art. no. 8267838.

[27] 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.

[28] 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.

[29] E. Ahmed and M. H. Rehmani, "Mobile edge computing: Opportunities, solutions, and challenges," *Future Gener. Comput. Syst.*, vol. 70, pp. 59–63, May 2017.

[30] Y. Ai, M. Peng, and K. Zhang, "Edge computing technologies for Internet of Things: a primer," *Digit. Commun. Netw.*, vol. 4, no. 2, pp. 77–86, 2018.

[31] F. Bonomi, "Connected vehicles, the Internet of Things, and fog computing," in *Proc. 11th ACM Int. Workshop Veh. Inter-Netw. (VANET)*, 2011, pp. 13–15.

[32] N. Fernando, S. W. Loke, and W. Rahayu, "Mobile cloud computing: A survey," *Future Gener. Comput. Syst.*, vol. 29, no. 1, pp. 84–106, 2013.

[33] P. Subramaniam and M. J. Kaur, "Review of security in mobile edge computing with deep learning," in *Proc. Adv. Sci. Eng. Technol. Int. Conf. (ASET)*, 2019, pp. 1–5.

[34] Y. Chen, Y. Zhang, and S. Maharjan, "Deep learning for secure mobile edge computing," 2017, *arXiv:1709.08025*.

[35] M. A. Al-Garadi, A. Mohamed, A. K. Al-Ali, X. Du, I. Ali, and M. Guizani, "A survey of machine and deep learning methods for Internet of Things (IoT) security," *IEEE Commun. Surveys Tuts.*, vol. 22, no. 3, pp. 1646–1685, 3rd Quart., 2020.

[36] M. K. Shambour and A. Gutub, "Progress of IoT research technologies and applications serving Hajj and Umrah," *Arab. J. Sci. Eng.*, vol. 47, no. 2, pp. 1253–1273, 2022.

[37] 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.

[38] F. Giust, X. Costa-Perez, and A. Reznik, "Multi-access edge computing: An overview of ETSI MEC ISG," *IEEE 5G Tech Focus*, vol. 1, no. 4, p. 4, Dec. 2017.

[39] S. Kekki et al., "MEC in 5G networks," ETSI, Sophia Antipolis, France, White Paper, 2018.

[40] K. Lounis and M. Zulkernine, "Attacks and defenses in short-range wireless technologies for IoT," *IEEE Access*, vol. 8, pp. 88892–88932, 2020.

[41] L. Zanzi et al., "Evolving multi-access edge computing to support enhanced IoT deployments," *IEEE Commun. Stand. Mag.*, vol. 3, no. 2, pp. 26–34, Jun. 2019.

[42] C. Campolo, R. Dos Reis Fontes, A. Molinaro, C. E. Rothenberg, and A. Iera, "Slicing on the road: Enabling the automotive vertical through 5G network softwarization," *Sensors*, vol. 18, no. 12, p. 4435, 2018.

[43] E. Schiller, N. Nikaein, E. Kalogeiton, M. Gasparyan, and T. Braun, "CDS-MEC: NFV/SDN-based application management for MEC in 5G systems," *Comput. Netw.*, vol. 135, pp. 96–107, Apr. 2018.

[44] C. Cimpanu. "You can now rent a Mirai botnet of 400,000 bots." 2016. [Online]. Available: BleepingComputer.com

[45] D. Zeng, S. Guo, and Z. Cheng, "The Web of things: A survey," *J. Commun.*, vol. 6, no. 6, pp. 424–438, 2011.

[46] Q. He et al., "A game-theoretical approach for mitigatingedge DDoS attack," *IEEE Trans. Dependable Secure Comput.*, vol. 19, no. 4, pp. 2333–2348, Jul./Aug. 2022.

[47] K.-W. Huang and H.-M. Wang, "Identifying the fake base station: A location based approach," *IEEE Commun. Lett.*, vol. 22, no. 8, pp. 1604–1607, Aug. 2018.

[48] N. Farooqi, A. Gutub, and M. O. Khozium, "Smart community challenges: Enabling IoT/M2M technology case study," *Life Sci J*, vol. 16, no. 7, pp. 11–17, 2019.

[49] X. Liu, Z. Zhou, W. Diao, Z. Li, and K. Zhang, "When good becomes evil: Keystroke inference with smartwatch," in *Proc. 22nd ACM SIGSAC Conf. Comput. Commun. Security*, 2015, pp. 1273–1285.

[50] E. Fernandes, J. Jung, and A. Prakash, "Security analysis of emerging smart home applications," in *Proc. IEEE Symp. Security Privacy (SP)*, 2016, pp. 636–654.

[51] S. Chen, R. Ma, H.-H. Chen, H. Zhang, W. Meng, and J. Liu, "Machine-to-machine communications in ultra-dense networks—A survey," *IEEE IEEE Commun. Surveys Tuts.*, vol. 19, no. 3, pp. 1478–1503, 3rd Quart., 2017.

[52] X. Lu, D. Niyato, N. Privault, H. Jiang, and P. Wang, "Managing physical layer security in wireless cellular networks: A cyber insurance approach," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 7, pp. 1648–1661, Jul. 2018.

[53] X. Costa-Perez et al., "5G-Crosshaul: An SDN/NFV integrated fronthaul/backhaul transport network architecture," *IEEE Wireless Commun.*, vol. 24, no. 1, pp. 38–45, Feb. 2017.

[54] I. Ahmad, S. Shahabuddin, T. Kumar, J. Okwuibe, A. Gurtov, and M. Ylianttila, "Security for 5G and beyond," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 4, pp. 3682–3722, 4th Quart., 2019.

[55] D. Kreutz, F. M. Ramos, and P. Verissimo, "Towards secure and dependable software-defined networks," in *Proc. 2nd ACM SIGCOMM Workshop Hot Topics Softw. Defined Netw.*, 2013, pp. 55–60.

[56] S. Shin and G. Gu, "Attacking software-defined networks: A first feasibility study," in *Proc. 2nd ACM SIGCOMM Workshop Hot Topics Softw. Defined Netw.*, 2013, pp. 165–166.

[57] P. Fonseca, R. Bennesby, E. Mota, and A. Passito, "A replication component for resilient OpenFlow-based networking," in *Proc. IEEE Netw. Oper. Manage. Symp.*, 2012, pp. 933–939.

[58] T. Nadeau and P. Pan, "Software driven networks problem statement," Network Working Group Internet-Draft, IETF, Sep. 2011.

[59] A. Shaghaghi, M. A. Kaafar, R. Buyya, and S. Jha, "Software-defined network (SDN) data plane security: Issues, solutions, and future directions," in *Handbook of Computer Networks and Cyber Security*. Cham, Switzerland: Springer, 2020, pp. 341–387.

[60] S. Scott-Hayward, S. Natarajan, and S. Sezer, "A survey of security in software defined networks," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 1, pp. 623–654, 1st Quart., 2016.

[61] J. Szefer, E. Keller, R. B. Lee, and J. Rexford, "Eliminating the hypervisor attack surface for a more secure cloud," in *Proc. 18th ACM Conf. Comput. Commun. Security*, 2011, pp. 401–412.

[62] P. Dubrulle, R. Sirdey, P. Dore, M. Aichouch, and E. Ohayon, "Blind hypervision to protect virtual machine privacy against hypervisor escape vulnerabilities," in *Proc. IEEE 13th Int. Conf. Ind. Informat. (INDIN)*, 2015, pp. 1394–1399.

[63] K. C. Lee, B. Zheng, C. Chen, and C.-Y. Chow, "Efficient index-based approaches for skyline queries in location-based applications," *IEEE Trans. Knowl. Data Eng.*, vol. 25, no. 11, pp. 2507–2520, Nov. 2013.

[64] K. Fawaz and K. G. Shin, "Location privacy protection for smartphone users," in *Proc. ACM SIGSAC Conf. Comput. Commun. Security*, 2014, pp. 239–250.

[65] 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.

[66] G. Liu, H. Zhao, F. Fan, G. Liu, Q. Xu, and S. Nazir, "An enhanced intrusion detection model based on improved KNN in WSNs," *Sensors*, vol. 22, no. 4, p. 1407, 2022.

[67] A. R. Syarif and W. Gata, "Intrusion detection system using hybrid binary PSO and K-nearest neighborhood algorithm," in *Proc. 11th Int. Conf. Inf. Commun. Technol. Syst. (ICTS)*, 2017, pp. 181–186.

[68] H. Shapoorifard and P. Shamsinejad, "Intrusion detection using a novel hybrid method incorporating an improved KNN," *Int. J. Comput. Appl*, vol. 173, no. 1, pp. 5–9, 2017.

[69] Y. Tan, W. Wu, J. Liu, H. Wang, and M. Xian, "Lightweight edge-based KNN privacy-preserving classification scheme in cloud computing circumstance," *Concurr. Comput. Pract. Exp.*, vol. 32, no. 19, 2020, Art. no. e5804.

[70] Q. Chen, K. Fan, K. Zhang, H. Wang, H. Li, and Y. Yang, "Privacy-preserving searchable encryption in the intelligent edge computing," *Comput. Commun.*, vol. 164, pp. 31–41, Dec. 2020.

[71] S. Chen, Z. Pang, H. Wen, K. Yu, T. Zhang, and Y. Lu, "Automated labeling and learning for physical layer authentication against clone node and Sybil attacks in industrial wireless edge networks," *IEEE Trans. Ind. Informat.*, vol. 17, no. 3, pp. 2041–2051, Mar. 2021.

[72] V. V. Kumari and P. R. K. Varma, "A semi-supervised intrusion detection system using active learning SVM and fuzzy c-means clustering," in *Proc. Int. Conf. I-SMAC (IoT Social, Mobile, Analytics Cloud)(I-SMAC)*, 2017, pp. 481–485.

[73] L. Xu, D. Zhang, N. Jayasena, and J. Cavazos, "HADM: Hybrid analysis for detection of malware," in *Proc. SAI Intell. Syst. Conf.*, 2016, pp. 702–724.

[74] A. J. Majumder, J. D. Miller, C. B. Veilleux, and A. A. Asif, "Smart-power: A smart cyber-physical system to detect IoT security threat through behavioral power profiling," in *Proc. IEEE 44th Annu. Comput., Softw., Appl. Conf. (COMPSAC)*, 2020, pp. 1041–1049.

[75] V. Laguduva, S. A. Islam, S. Aakur, S. Katkoori, and R. Karam, "Machine learning based IoT edge node security attack and countermeasures," in *Proc. IEEE Comput. Soc. Annu. Symp. VLSI (ISVLSI)*, 2019, pp. 670–675.

[76] C. Sinclair, L. Pierce, and S. Matzner, "An application of machine learning to network intrusion detection," in *Proc. 15th Annu. Comput. Security Appl. Conf. (ACSAC)*, 1999, pp. 371–377.

[77] X. Wu, X. Xu, F. Dai, J. Gao, G. Ji, and L. Qi, "An ensemble of random decision trees with personalized privacy preservation in edge-cloud computing," in *Proc. Int. Conf. Internet Things (iThings) IEEE Green Comput. Commun. (GreenCom) IEEE Cyber, Phys. Social Comput. (CPSCom) IEEE Smart Data (SmartData) IEEE Congr. Cybermat. (Cybermatics)*, 2020, pp. 779–786.

[78] A. K. Sangaiah, D. V. Medhane, T. Han, M. S. Hossain, and G. Muhammad, "Enforcing position-based confidentiality with machine learning paradigm through mobile edge computing in real-time industrial informatics," *IEEE Trans. Ind. Informat.*, vol. 15, no. 7, pp. 4189–4196, Jul. 2019.

[79] Z. Liu, N. Su, Y. Qin, J. Lu, and X. Li, "A deep random forest model on spark for network intrusion detection," *Mobile Inf. Syst.*, vol. 2020, Dec. 2020, Art. no. 6633252.

[80] R. K. Singh, S. Dalal, V. K. Chauhan, and D. Kumar, "Optimization of far in intrusion detection system by using random forest algorithm," in *Proc. 2nd Int. Conf. Adv. Comput. Softw. Eng. (ICACSE)*, 2019, pp. 1–4.

[81] Z. Guowei, Y. Hui, Y. Zhuang, G. Yue, X. Zhang, and Q. Shuang, "Research on network intrusion detection method of power system based on random forest algorithm," in *Proc. 13th Int. Conf. Meas. Technol. Mechatronics Autom. (ICMTMA)*, 2021, pp. 374–379.

[82] M. Hasan, M. M. Islam, M. I. I. Zarif, and M. Hashem, "Attack and anomaly detection in IoT sensors in IoT sites using machine learning approaches," *Internet Things*, vol. 7, Sep. 2019, Art. no. 100059.

[83] M. Wu, Z. Song, and Y. B. Moon, "Detecting cyber-physical attacks in CyberManufacturing systems with machine learning methods," *J. Intell. Manuf.*, vol. 30, no. 3, pp. 1111–1123, 2019.

[84] W. Zhang, J. Wang, G. Han, S. Huang, Y. Feng, and L. Shu, "A data set accuracy weighted random forest algorithm for IoT fault detection based on edge computing and blockchain," *IEEE Internet Things J.*, vol. 8, no. 4, pp. 2354–2363, Feb. 2021.

[85] Z. Tian, C. Luo, J. Qiu, X. Du, and M. Guizani, "A distributed deep learning system for Web attack detection on edge devices," *IEEE Trans. Ind. Informat.*, vol. 16, no. 3, pp. 1963–1971, Mar. 2020.

[86] Q. Ran, X. Ju, X. Zhang, and Y. Zhang, "Cloud edge cooperative attack recognition based on CNN," in *Proc. J. Phys. Conf. Series*, 2020, Art. no. 12143.

[87] Y. Tian, J. Yuan, S. Yu, and Y. Hou, "LEP-CNN: A lightweight edge device assisted privacy-preserving CNN inference solution for IoT," 2019, *arXiv:1901.04100*.

[88] G. Radhakrishnan, K. Srinivasan, S. Maheswaran, K. Mohanasundaram, D. Palanikkumar, and A. Vidyarthi, "WITHDRAWN: A deep-RNN and meta-heuristic feature selection approach for IoT malware detection," *Mater. Today Proc.*, to be published.

[89] Q. Li, S. Meng, S. Zhang, J. Hou, and L. Qi, "Complex attack linkage decision-making in edge computing networks," *IEEE Access*, vol. 7, pp. 12058–12072, 2019.

[90] G. Han, H. Wang, M. Guizani, S. Chan, and W. Zhang, "KCLP: A k-means cluster-based location privacy protection scheme in WSNs for IoT," *IEEE Wireless Commun.*, vol. 25, no. 6, pp. 84–90, Dec. 2018.

[91] Y. Liu, Z. Ma, Z. Yan, Z. Wang, X. Liu, and J. Ma, "Privacy-preserving federated k-means for proactive caching in next generation cellular networks," *Inf. Sci.*, vol. 521, pp. 14–31, Jun. 2020.

[92] T. T. Huong, T. P. Bac, D. M. Long, B. D. Thang, T. D. Luong, and N. T. Binh, "An efficient low complexity edge-cloud framework for security in IoT networks," in *Proc. IEEE 8th Int. Conf. Commun. Electron. (ICCE)*, 2021, pp. 533–539.

[93] S. A. Osia, A. S. Shamsabadi, A. Taheri, H. R. Rabiee, and H. Haddadi, "Private and scalable personal data analytics using hybrid edge-to-cloud deep learning," *Computer*, vol. 51, no. 5, pp. 42–49, May 2018.

[94] R. Benchea and D. T. Gavriluţ, "Combining restricted Boltzmann machine and one side perceptron for malware detection," in *Proc. Int. Conf. Conceptual Struct.*, 2014, pp. 93–103.

[95] Q. Tian, D. Han, K.-C. Li, X. Liu, L. Duan, and A. Castiglione, "An intrusion detection approach based on improved deep belief network," *Appl. Intell.*, vol. 50, pp. 3162–3178, May 2020.

[96] J. Schneible and A. Lu, "Anomaly detection on the edge," in *Proc. IEEE Mil. Commun. Conf. (MILCOM)*, 2017, pp. 678–682.

[97] K. Sadaf and J. Sultana, "Intrusion detection based on autoencoder and isolation forest in fog computing," *IEEE Access*, vol. 8, pp. 167059–167068, 2020.

[98] X. Wang, I. Yang, and S.-H. Ahn, "Sample efficient home power anomaly detection in real time using semi-supervised learning," *IEEE Access*, vol. 7, pp. 139712–139725, 2019.

[99] R. E. Hiromoto, M. Haney, and A. Vakanski, "A secure architecture for IoT with supply chain risk management," in *Proc. 9th IEEE Int. Conf. Intell. Data Acquisit. Adv. Comput. Syst. Technol. Appl. (IDAACS)*, vol. 1, 2017, pp. 431–435.

[100] J. Yoon and H. Lee, "PUFGAN: Embracing a self-adversarial agent for building a defensible edge security architecture," in *Proc. IEEE Conf. Comput. Commun.*, 2020, pp. 904–913.

[101] L. Xiao, Y. Li, X. Huang, and X. Du, "Cloud-based malware detection game for mobile devices with offloading," *IEEE Trans. Mobile Comput.*, vol. 16, no. 10, pp. 2742–2750, Oct. 2017.

[102] M. A. Aref, S. K. Jayaweera, and S. Machuzak, "Multi-agent reinforcement learning based cognitive anti-jamming," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, 2017, pp. 1–6.

[103] 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.

[104] H. Li, J. Wu, H. Xu, G. Li, and M. Guizani, "Explainable intelligence-driven defense mechanism against advanced persistent threats: A joint edge game and AI approach," *IEEE Trans. Dependable Secure Comput.*, vol. 19, no. 2, pp. 757–775, Mar./Apr. 2021.

[105] J. L. Thames, R. Abler, and A. Saad, "Hybrid intelligent systems for network security," in *Proc. 44th Annu. Southeast Regional Conf.*, 2006, pp. 286–289.

[106] R. Wright and Z. Yang, "Privacy-preserving Bayesian network structure computation on distributed heterogeneous data," in *Proc. 10th ACM SIGKDD Int. Conf. Knowl. Disc. Data Min.*, 2004, pp. 713–718.

[107] K. Kim and B.-R. Moon, "Malware detection based on dependency graph using hybrid genetic algorithm," in *Proc. 12th Annu. Conf. Genet. Evol. Comput.*, 2010, pp. 1211–1218.

[108] P. K. Roy, S. Saumya, J. P. Singh, S. Banerjee, and A. Gutub, "Analysis of community question-answering issues via machine learning and deep learning: State-of-the-art review," *CAAI Trans. Intell. Technol.*, vol. 8, no. 1, pp. 95–117, 2022.

[109] D. Santhadevi and B. Janet, "EIDIMA: Edge-based intrusion detection of IoT malware attacks using decision tree-based boosting algorithms," in *High Performance Computing and Networking*. Singapore: Springer, 2022, pp. 449–459.

[110] M. B. Farukee, M. Z. Shabit, M. R. Haque, and A. S. Sattar, "DDoS attack detection in IoT networks using deep learning models combined with random forest as feature selector," in *Proc. Int. Conf. Adv. Cyber Security*, 2020, pp. 118–134.

[111] I. Alrashdi, A. Alqazzaz, E. Aloufi, R. Alharthi, M. Zohdy, and H. Ming, "Ad-IoT: Anomaly detection of IoT cyberattacks in smart city using machine learning," in *Proc. IEEE 9th Annu. Comput. Commun. Workshop Conf. (CCWC)*, 2019, pp. 305–310.

[112] Z. Wang, J. Tian, H. Fang, L. Chen, and J. Qin, "LightLog: A lightweight temporal convolutional network for log anomaly detection on the edge," *Comput. Netw.*, vol. 203, Feb. 2022, Art. no. 108616.

[113] E. Li, L. Zeng, Z. Zhou, and X. Chen, "Edge AI: On-demand accelerating deep neural network inference via edge computing," *IEEE Trans. Wireless Commun.*, vol. 19, no. 1, pp. 447–457, Jan. 2020.

[114] Z. Zhao, K. M. Barijough, and A. Gerstlauer, "DeepThings: Distributed adaptive deep learning inference on resource-constrained IoT edge clusters," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 37, no. 11, pp. 2348–2359, Nov. 2018.

[115] N. Guizani and A. Ghafoor, "A network function virtualization system for detecting malware in large IoT based networks," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 6, pp. 1218–1228, Jun. 2020.

[116] M. Arivukarasi and A. Antonidoss, "Performance analysis of malicious URL detection by using RNN and LSTM," in *Proc. 4th Int. Conf. Comput. Methodol. Commun. (ICCMC)*, 2020, pp. 454–458.

[117] C.-Y. Chen, L.-A. Chen, Y.-Z. Cai, and M.-H. Tsai, "RNN-based DDoS detection in IoT scenario," in *Proc. Int. Comput. Symp. (ICS)*, 2020, pp. 448–453.

[118] A. Diro and N. Chilamkurti, "Leveraging LSTM networks for attack detection in fog-to-things communications," *IEEE Commun. Mag.*, vol. 56, no. 9, pp. 124–130, Sep. 2018.

[119] Z. Ghahramani, "Unsupervised learning," in *Proc. Summer School Mach. Learn.* 2003, pp. 72–112.

[120] A. I. Károly, R. Fullér, and P. Galambos, "Unsupervised clustering for deep learning: A tutorial survey," *Acta Polytechnica Hungarica*, vol. 15, no. 8, pp. 29–53, 2018.

[121] R. Wason, S. Aghili, and P. Zavarsky, "An integrated CASB implementation model to enhance enterprise cloud security," Faculty Graduate Stud., Concordia Univ. Edmonton, Edmonton, AB, Canada, Project Rep., 2020.

[122] S. Hou, A. Saas, Y. Ye, and L. Chen, "DroidDelver: An android malware detection system using deep belief network based on API call blocks," in *Proc. Int. Conf. Web-Age Inf. Manage.*, 2016, pp. 54–66.

[123] X. J. Zhu, "Semi-supervised learning literature survey," Dept. Comput. Sci., Univ. Wisconsin, Madison, WI, USA, Rep. TR 1530, 2005.

[124] S. Rathore and J. H. Park, "Semi-supervised learning based distributed attack detection framework for IoT," *Appl. Soft Comput.*, vol. 72, pp. 79–89, Nov. 2018.

[125] N. Z. Gong, M. Frank, and P. Mittal, "SybilBelief: A semi-supervised learning approach for structure-based Sybil detection," *IEEE Trans. Inf. Forensics Security*, vol. 9, pp. 976–987, 2014.

[126] M. Li, J. Lin, Y. Ding, Z. Liu, J.-Y. Zhu, and S. Han, "GAN compression: Efficient architectures for interactive conditional GANs," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2020, pp. 5284–5294.

[127] M. Abdel-Basset, H. Hawash, R. K. Chakrabortty, and M. J. Ryan, "Semi-supervised spatio-temporal deep learning for intrusions detection in IoT networks," *IEEE Internet Things J.*, vol. 8, no. 15, pp. 12251–12265, Aug. 2021.

[128] Y. Qian, R. Wang, J. Wu, B. Tan, and H. Ren, "Reinforcement learning-based optimal computing and caching in mobile edge network," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 10, pp. 2343–2355, Oct. 2020.

[129] X. Wang, Z. Ning, and S. Guo, "Multi-agent imitation learning for pervasive edge computing: A decentralized computation offloading algorithm," *IEEE Trans. Parallel Distrib. Syst.*, vol. 32, no. 2, pp. 411–425, Feb. 2021.

[130] L. Chen, S. Tang, V. Balasubramanian, J. Xia, F. Zhou, and L. Fan, "Physical-layer security based mobile edge computing for emerging cyber physical systems," *Comput. Commun.*, vol. 194, pp. 180–188, Oct. 2022.

[131] X. Qiu, W. Zhang, W. Chen, and Z. Zheng, "Distributed and collective deep reinforcement learning for computation offloading: A practical perspective," *IEEE Trans. Parallel Distrib. Syst.*, vol. 32, no. 5, pp. 1085–1101, May 2021.

[132] A. Irpan. "Deep reinforcement learning doesn't work yet." 2018. [Online]. Available: https://www.alexirpan.com/2018/02/14/rl-hard.html

[133] O. Pourret, P. Naʹım, and B. Marcot, *Bayesian Networks: A Practical Guide to Applications*. Chichester, U.K.: Wiley, 2008.

[134] M. Frigault, L. Wang, A. Singhal, and S. Jajodia, "Measuring network security using dynamic Bayesian network," in *Proc. 4th ACM Workshop Qual. Protect.*, 2008, pp. 23–30.

[135] M. Frigault and L. Wang, "Measuring network security using Bayesian network-based attack graphs," in *Proc. 32nd Annu. IEEE Int. Comput. Softw. Appl. Conf.*, 2008, pp. 698–703.

[136] Y. Chen et al., "Clustering based physical-layer authentication in edge computing systems with asymmetric resources," *Sensors*, vol. 19, no. 8, p. 1926, 2019.

[137] R.-F. Liao et al., "Multiuser physical layer authentication in Internet of Things with data augmentation," *IEEE Internet Things J.*, vol. 7, no. 3, pp. 2077–2088, Mar. 2020.

[138] D. Kim et al., "Squeezed convolutional variational autoencoder for unsupervised anomaly detection in edge device Industrial Internet of Things," in *Proc. Int. Conf. Inf. Comput. Technol. (ICICT)*, 2018, pp. 67–71.

[139] J. Li, Z. Zhao, R. Li, and H. Zhang, "AI-based two-stage intrusion detection for software defined IoT networks," *IEEE Internet Things J.*, vol. 6, no. 2, pp. 2093–2102, Apr. 2019.

[140] T. A. Tang, L. Mhamdi, D. McLernon, S. A. R. Zaidi, and M. Ghogho, "Deep learning approach for network intrusion detection in software defined networking," in *Proc. Int. Conf. Wireless Netw. Mobile Commun. (WINCOM)*, 2016, pp. 258–263.

[141] K. Kirutika, V. Vetriselvi, R. Parthasarathi, and G. S. V. Rao, "Controller monitoring system in software defined networks using random forest algorithm," in *Proc. Int. Carnahan Conf. Security Technol. (ICCST)*, 2019, pp. 1–6.

[142] X. Pan, M. Zhang, D. Wu, Q. Xiao, S. Ji, and Z. Yang, "Justinian's GAAvernor: Robust distributed learning with gradient aggregation agent," in *Proc. 29th USENIX Security Symp. (USENIX Security)*, 2020, pp. 1641–1658.

[143] R. Kozik, M. Choraś, M. Ficco, and F. Palmieri, "A scalable distributed machine learning approach for attack detection in edge computing environments," *J. Parallel Distrib. Comput.*, vol. 119, pp. 18–26, Sep. 2018.

[144] N. Alassaf, A. Gutub, S. A. Parah, and M. Al Ghamdi, "Enhancing speed of SIMON: A light-weight-cryptographic algorithm for IoT applications," *Multimedia Tools Appl.*, vol. 78, no. 23, pp. 32633–32657, 2019.

[145] N. Alassaf and A. Gutub, "Simulating light-weight-cryptography implementation for IoT healthcare data security applications," *Int. J. E-Health Med. Commun.*, vol. 10, no. 4, pp. 1–15, 2019.

[146] N. Kheshaifaty and A. Gutub, "Engineering graphical Captcha and AES crypto hash functions for secure online authentication," *J. Eng. Res.*, to be published.

[147] N. Farnaaz and M. A. Jabbar, "Random forest modeling for network intrusion detection system," *Procedia Comput. Sci.*, vol. 89, pp. 213–217, Aug. 2016.

[148] K. Huang, X. Liu, S. Fu, D. Guo, and M. Xu, "A lightweight privacy-preserving CNN feature extraction framework for mobile sensing," *IEEE Trans. Dependable Secure Comput.*, vol. 18, no. 3, pp. 1441–1455, May/Jun. 2021.

[149] Z. Yu et al., "Federated learning based proactive content caching in edge computing," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, 2018, pp. 1–6.

[150] Y. Zhao et al., "Privacy-preserving blockchain-based federated learning for IoT devices," *IEEE Internet Things J.*, vol. 8, no. 3, pp. 1817–1829, Feb. 2021.

[151] H. Wu et al., "PECAM: Privacy-enhanced video streaming and analytics via securely-reversible transformation," in *Proc. 27th Annu. Int. Conf. Mobile Comput. Netw.*, 2021, pp. 229–241.

[152] P. Zhao, H. Huang, X. Zhao, and D. Huang, "P3: Privacy-preserving scheme against poisoning attacks in mobile-edge computing," *IEEE Trans. Comput. Social Syst.*, vol. 7, no. 3, pp. 818–826, Jun. 2020.

[153] Y. Mao, S. Yi, Q. Li, J. Feng, F. Xu, and S. Zhong, "A privacy-preserving deep learning approach for face recognition with edge computing," in *Proc. USENIX Workshop Hot Topics Edge Comput.(HotEdge)*, 2018, pp. 1–6.

[154] X. Lu, Y. Liao, P. Lio, and P. Hui, "Privacy-preserving asynchronous federated learning mechanism for edge network computing," *IEEE Access*, vol. 8, pp. 48970–48981, 2020.

[155] M. Alotaibi, D. Al-Hendi, B. Alroithy, M. AlGhamdi, and A. Gutub, "Secure mobile computing authentication utilizing hash, cryptography and steganography combination," *J. Inf. Security Cybercrimes Res.*, vol. 2, no. 1, pp. 73–82, 2019.

[156] M. Shambour and A. Gutub, "Personal privacy evaluation of smart devices applications serving Hajj and Umrah rituals," *J. Eng. Res.*, to be published.

[157] H. Samkari and A. Gutub, "Protecting medical records against cybercrimes within Hajj period by 3-layer security," *Recent Trends Inf. Technol. Appl.*, vol. 2, no. 3, pp. 1–21, 2019.

[158] S. Almutairi, A. Gutub, and M. Al-Ghamdi, "Image steganography to facilitate online students account system," *Rev. Bus. Technol. Res*, vol. 16, no. 2, pp. 43–49, 2019.

[159] P. Albers, O. Camp, J.-M. Percher, B. Jouga, L. Me, and R. S. Puttini, "Security in ad hoc networks: A general intrusion detection architecture enhancing trust based approaches," in *Proc. Wireless Inf. Syst.*, 2002, pp. 1–12.

[160] 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.

[161] W. Zhang et al., "ELF: Accelerate high-resolution mobile deep vision with content-aware parallel offloading," in *Proc. 27th Annu. Int. Conf. Mobile Comput. Netw.*, 2021, pp. 201–214.

[162] T. Mohammed, C. Joe-Wong, R. Babbar, and M. Di Francesco, "Distributed inference acceleration with adaptive DNN partitioning and offloading," in *Proc. IEEE Conf. Comput. Commun.*, 2020, pp. 854–863.

[163] L. Zeng, X. Chen, Z. Zhou, L. Yang, and J. Zhang, "CoEdge: Cooperative DNN inference with adaptive workload partitioning over heterogeneous edge devices," *IEEE/ACM Trans. Netw.*, vol. 29, no. 2, pp. 595–608, Apr. 2021.

[164] O. Krestinskaya, A. P. James, and L. O. Chua, "Neuromemristive circuits for edge computing: A review," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 31, no. 1, pp. 4–23, Jan. 2020.

[165] C. Lammie, W. Xiang, B. Linares-Barranco, and M. R. Azghadi, "MemTorch: An open-source simulation framework for memristive deep learning systems," 2020, *arXiv:2004.10971*.

Cheng Wang received the B.S. and M.S. degrees from Lanzhou University, Lanzhou, China, in 2017 and 2020, respectively. He is currently pursuing the Ph.D. deegree with Hubei Engineering Research Center on Big Data Security, School of Cyber Science and Engineering, Huazhong University of Science and Technology, Wuhan, China.

His research interests include Web security, edge computing, and recommender system.

Zenghui Yuan received the M.S. degree from Huazhong University of Science and Technology, Wuhan, China, in 2021, where he is currently pursuing the Ph.D. degree with Hubei Engineering Research Center on Big Data Security, School of Cyber Science and Engineering.

His main research interests include deep learning, reinforcement learning, and edge computing.

Pan Zhou (Senior Member, IEEE) received the B.S. degree from the Advanced Class, Huazhong University of Science and Technology (HUST), Wuhan, China, in 2006, the M.S. degree from the Department of Electronics and Information Engineering, HUST in 2008, and the Ph.D. degree from the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA, USA, in 2011.

He is currently a Full Professor and a Ph.D. Advisor with Hubei Engineering Research Center on Big Data Security, School of Cyber Science and Engineering, HUST. He held honorary degree in his bachelor and merit research award of HUST in his master study. He was a Senior Technical Member with Oracle Inc., Austin, TX, USA, from 2011 to 2013, and worked on Hadoop and distributed storage system for big data analytics at Oracle Cloud Platform. He has published more than 170 refereed papers in international leading journals and key conferences in the area of security and privacy, big data analytics, machine learning, mobile computing, and networks, including: IEEE/ACM TRANSACTIONS ON NETWORKING, IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, IEEE TRANSACTIONS ON DEPENDABLE AND SECURE COMPUTING, IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY, IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS, IEEE TRANSACTIONS ON MOBILE COMPUTING, IEEE TRANSACTIONS ON PARALLEL AND DISTRIBUTED SYSTEMS, IEEE TRANSACTIONS ON INFORMATION THEORY, IEEE TRANSACTIONS ON IMAGE PROCESSING, IEEE TRANSACTIONS ON SOFTWARE ENGINEERING, IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS, IEEE TRANSACTIONS ON COMPONENTS, PACKAGING AND MANUFACTURING TECHNOLOGY, IEEE TRANSACTIONS ON MULTIMEDIA, IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS, IEEE TRANSACTIONS ON ARTIFICIAL INTELLIGENCE, IEEE TRANSACTIONS ON COMMUNICATIONS, IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, IEEE TRANSACTIONS ON NETWORK SCIENCE AND ENGINEERING, IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, IEEE TRANSACTIONS ON EMERGING TOPICS IN COMPUTATIONAL INTELLIGENCE, ACM TRANSACTIONS ON STORAGE, and ACM TRANSACTIONS ON KNOWLEDGE DISCOVERY FROM DATA, and ICDE, INFOCOM, CVPR, ICCV, ICDCS, ICPP, ACM MM, NEURIPS, AAAI, IJCAI, NAACL, COLING, PoPETs/PETS, SECON, CIKM, and ECAI.

Prof. Zhou received the "Rising Star in Science and Technology of HUST" in 2017. He is currently an Associate Editor of the IEEE TRANSACTIONS ON NETWORK SCIENCE AND ENGINEERING.

Zichuan Xu (Member, IEEE) received the B.Sc. and M.E. degrees in computer science from Dalian University of Technology, Dalian, China, in 2008 and 2011, respectively, and the Ph.D. degree in computer science from the Australian National University, Canberra, ACT, Australia, in 2016.

He was a Research Associate with the University College London, London, U.K. He is currently an Associate Professor with the School of Software, Dalian University of Technology. His research interests include cloud computing, software-defined networking, wireless sensor networks, algorithmic game theory, and optimization problems.

Ruixuan Li (Member, IEEE) received the B.S., M.S., and Ph.D. degrees in computer science from Huazhong University of Science and Technology, Wuhan, China, in 1997, 2000, and 2004, respectively.

From 2009 to 2010, he was a Visiting Researcher with the Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON, Canada. He is currently a Professor with the School of Computer Science and Technology, Huazhong University of Science and Technology. His research interests include cloud and edge computing, big data management, and distributed system security.

Prof. Li is a member of ACM.

Dapeng Oliver Wu (Fellow, IEEE) received the B.E. degree in electrical engineering from Huazhong University of Science and Technology, Wuhan, China, in 1990, the M.E. degree in electrical engineering from Beijing University of Posts and Telecommunications, Beijing, China, in 1997, and the Ph.D. degree in electrical and computer engineering from Carnegie Mellon University, Pittsburgh, PA, USA, in 2003.

He is currently a Professor with the Department of Computer Science, City University of Hong Kong, Hong Kong. His research interests are in the areas of networking, communications, signa processing, computer vision, machine learning, smart grid, and information and network security.

Dr. Wu received the University of Florida Term Professorship Award in 2017, the University of Florida Research Foundation Professorship Award Limits in 2009, the AFOSR Young Investigator Program (YIP) Award in 2009, the ONR YIP Award in 2008, the NSF CAREER Award in 2007, the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY Best Paper Award in 2001, and the Best Paper Awards in IEEE GLOBECOM 2011 and the International Conference on Quality of Service in Heterogeneous Wired/Wireless Networks 2006. He was an Elected Member of the Multimedia Signal Processing Technical Committee and the IEEE Signal Processing Society from January 2009 to December 2012. He was elected as a Distinguished Lecturer by the IEEE Vehicular Technology Society in 2016. He has served as the Technical Program Committee Chair for IEEE INFOCOM 2012, the IEEE International Conference on Communications (ICC 2008), and the Signal Processing for Communications Symposium; and a member of executive committee and/or technical program committee of over 100 conferences. He has served as the Chair for the Award Committee, the Mobile and Wireless Multimedia Interest Group, the Technical Committee on Multimedia Communications, and the IEEE Communications Society. He has served as the Editor-in-Chief for the IEEE TRANSACTIONS ON NETWORK SCIENCE AND ENGINEERING, an Editor-at-Large for the IEEE OPEN JOURNAL OF THE COMMUNICATIONS SOCIETY, the Founding Editor-in-Chief for the *Advances in Multimedia*, and an Associate Editor for the IEEE TRANSACTIONS ON CLOUD COMPUTING, IEEE TRANSACTIONS ON COMMUNICATIONS, IEEE TRANSACTIONS ON SIGNAL AND INFORMATION PROCESSING OVER NETWORKS, *IEEE Signal Processing Magazine*, IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, and IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY. He is also a Guest Editor for the Special Issue on Cross-Layer Optimized Wireless Multimedia Communications and the Special Issue on Airborne Communication Networks of IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS.