Cloud-Edge Orchestration for the Internet of Things: Architecture and AI-Powered Data Processing

Yulei Wu, *Senior Member, IEEE*

***Abstract*—The Internet of Things (IoT) has been deeply penetrated into a wide range of important and critical sectors, including smart city, water, transportation, manufacturing, and smart factory. Massive data are being acquired from a fast growing number of IoT devices. Efficient data processing is a necessity to meet diversified and stringent requirements of many emerging IoT applications. Due to the constrained computation and storage resources, IoT devices have resorted to the powerful cloud computing to process their data. However, centralized and remote cloud computing may introduce unacceptable communication delay since its physical location is far away from IoT devices. Edge cloud has been introduced to overcome this issue by moving the cloud in closer proximity to IoT devices. The orchestration and cooperation between the cloud and the edge provides a crucial computing architecture for IoT applications. Artificial intelligence (AI) is a powerful tool to enable the intelligent orchestration in this architecture. This article first introduces such a kind of computing architecture from the perspective of IoT applications. It then investigates the state-of-the-art proposals on AI-powered cloud-edge orchestration for the IoT. Finally, a list of potential research challenges and open issues is provided and discussed, which can provide useful resources for carrying out future research in this area.**

***Index Terms*—Artificial intelligence (AI), cloud computing, edge computing, Internet of Things (IoT), offloading.**

I. Introduction

DUE to the advanced semiconductor and related technologies, including microelectromechanical systems and sensors, the Internet of Things (IoT) has gained significant attentions by a wide range of sectors, e.g., smart city, water, transportation, manufacturing, and smart factory [1]–[6]. A report from Ericsson presents that the estimated number of connected devices in a typical smart factory is 0.5 per square meter,[1] and in dense areas, this number could be increased up to 1 per square meter. According to the estimation by Statista, the number of IoT devices will reach 75.44 billion worldwide by 2025.[2] The popularity of the IoT in different sectors is owing to its essential abilities of ubiquitous data acquisition and the intelligent decision making drawn from the data [7], [8]. The data generated and collected by massive IoT devices have a set of key features, including large volume, high velocity, multimodes, various veracities, and heterogeneity [9]–[13]. Essentially, these data need to be transmitted to a computing facility for data processing and knowledge extraction, and eventually decisions can be made for autonomous operations of IoT applications [14]–[16]. The unique features of IoT data can therefore cause a significant increase in the burden on the communication networks (mainly wireless networks) and the computing facilities.

Cloud computing has gained popularity at a rapid pace for providing computing and storage resources to many applications including IoT applications [17]–[19]. Many large companies, such as Microsoft, Google, and Amazon, have produced their cloud computing platforms (e.g., Microsoft Azure IoT Suite, AWS IoT Platform, and IBM Watson IoT Platform) and have hosted numerous essential IoT services to help with business digital transformation. However, centralized cloud computing has been struggling to meet the demands of large-scale IoT networks, creating the scalability issues [20], [21]. Traditional centralized cloud computing that is located far away from IoT devices, introduces remarkable delay overheads due to the long transmission distance [22]. It also causes a significant increase in the bandwidth consumption of communication networks between the cloud and IoT devices [23]. The increase in the delay cannot be tolerated by many delay-sensitive IoT applications, e.g., augmented reality (AR), virtual reality (VR), mixed reality (MR), and autonomous vehicles. Besides, the privacy issue is another key concern when IoT data need to be offloaded to the centralized cloud computing for processing [24]–[26].

Edge computing brings the cloud in proximity to IoT devices [27], [28]. It intrinsically solves the above issues due to its decentralization nature [29], [30]. In practice, edge servers can be deployed at anywhere located closer to IoT devices, e.g., within a house, on top of a building, at the side of a road, and along with a base station. In principle, edge servers have less computing and storage capacities than the centralized cloud servers (also known as data centers) [31]. Therefore,

Manuscript received April 15, 2020; revised June 18, 2020 and July 10, 2020; accepted August 4, 2020. Date of publication August 7, 2020; date of current version August 6, 2021. This work was supported by the Engineering and Physical Sciences Research Council of United Kingdom under Grant EP/R030863/1.

The author is with the College of Engineering, Mathematics and Physical Sciences, University of Exeter, Exeter EX4 4QF, U.K. (e-mail: y.l.wu@exeter.ac.uk).

Digital Object Identifier 10.1109/JIOT.2020.3014845

[1]https://www.ericsson.com/en/mobility-report/articles/realizing-smart-manufact-iot

[2]https://www.statista.com/statistics/471264/iot-number-of-connected-devices-worldwide/

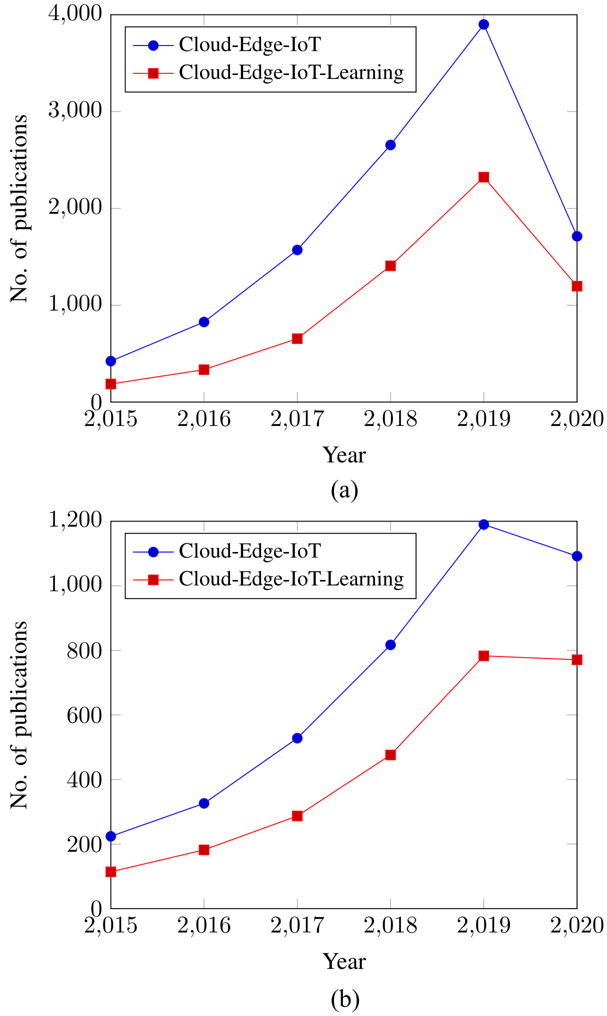


Fig. 1. Number of related publications from 2015 to June 2020 in the (a) IEEE Xplore database and (b) ScienceDirect database, using the keywords "cloud," "edge," and "IoT," and the keywords "cloud", "edge", "IoT", and "learning" searched in the full text.

edge servers may not be able to completely satisfy the computing and storage requirements of some IoT applications. The orchestration between edge servers and cloud servers is a necessity in certain IoT scenarios to meet diversified and stringent application requirements [23]. Different from centralized cloud servers, edge servers may belong to different network providers [32]. This creates additional challenges for cloud-edge orchestration, since different network providers may have different business models, operation rules, etc. In addition, this factor also brings privacy issues, since IoT data may be transferred to edge servers that are owned by different network providers [33], [34].

In cloud-edge orchestration, data are transmitted between the entities of IoT devices, edge servers, and cloud servers, for efficient processing and intelligent decision making. Related topics have received numerous attentions in recent years. Fig. 1 shows the number of publications in the IEEE Xplore database[3] and the ScienceDirect database.[4] The decrease from Year 2019 to Year 2020 is due to the count of only half a year in 2020 (i.e., until June 2020—the time of writing this article). There are two key aspects affecting the performance of cloud-edge orchestration: 1) architecture and 2) the associated data processing algorithms [35]. The architecture of cloud-edge orchestration varies with different IoT applications. The communication network is an important media for data transmission among the cloud, the edge and the IoT devices [36]. The fifth-generation (5G) communication system has been designed to provide data transmission with the requirements of high throughput, high reliability and low latency, and it also enables massive connections of IoT devices [2], [37], [38]. Artificial intelligence (AI), especially deep learning, has become powerful tools for data processing [39], [40]. The nature of pervasive AI in 5G and beyond 5G (B5G), coupled with the cloud-edge orchestration architecture, can enable on-demand local/remote data processing and real-time decision making for intelligent IoT ecosystems [41]–[44]. In this article, the state-of-the-art proposals of these two important aspects of cloud-edge orchestration will be investigated, followed by a list of potential research challenges and open issues. The aim of this survey article is to provide researchers a deeper understanding of these two aspects and enables them to continuously make contributions to bridge gaps in the research of cloud-edge orchestration.

The remainder of this article is organized as follows. Section II briefly introduces the concept and use cases of the IoT. Section III presents the architecture of cloud-edge orchestration for the IoT and its recent studies. The state-of-the-art proposals of AI-powered data processing for cloud-edge orchestration are investigated in Section IV. Potential research challenges and open issues are discussed in Section V. Finally, Section VI concludes this article.

II. Internet of Things

The IoT is to connect any devices (essentially everything in the physical world), with an "ON" and "OFF" switch, to the Internet [45]. It has been widely introduced into a wide range of sectors to enhance the automation in business operations. In industry 4.0, the IoT has been deeply penetrated into the areas of agriculture, power grid, water systems, autonomous vehicles, healthcare, and factories, to make these verticals smarter, creating the so-called Industrial IoT (IIoT) [46]–[50].

The use of the IoT in healthcare, called Internet of Healthcare Things or Internet of Medical Things, has been widely adopted to carry out digital transformation in the healthcare industries [51]–[53]. This thanks to the fast development of wearable and biosensor technologies, along with essential medical devices. Remote health monitoring is one of the most common and important application areas of the IoT in healthcare. There are many cases where remote health monitoring is necessary. For example, patients with chronic diseases have to stay at home due to the shortage of healthcare staffs in hospitals (e.g., during special periods like Covid-19), and the health monitoring data is vital in this case. Remote health monitoring becomes essential for people living in rural areas. The IoT is becoming more important in healthcare when it is

[3]https://ieeexplore.ieee.org/Xplore/home.jsp
[4]https://www.sciencedirect.com

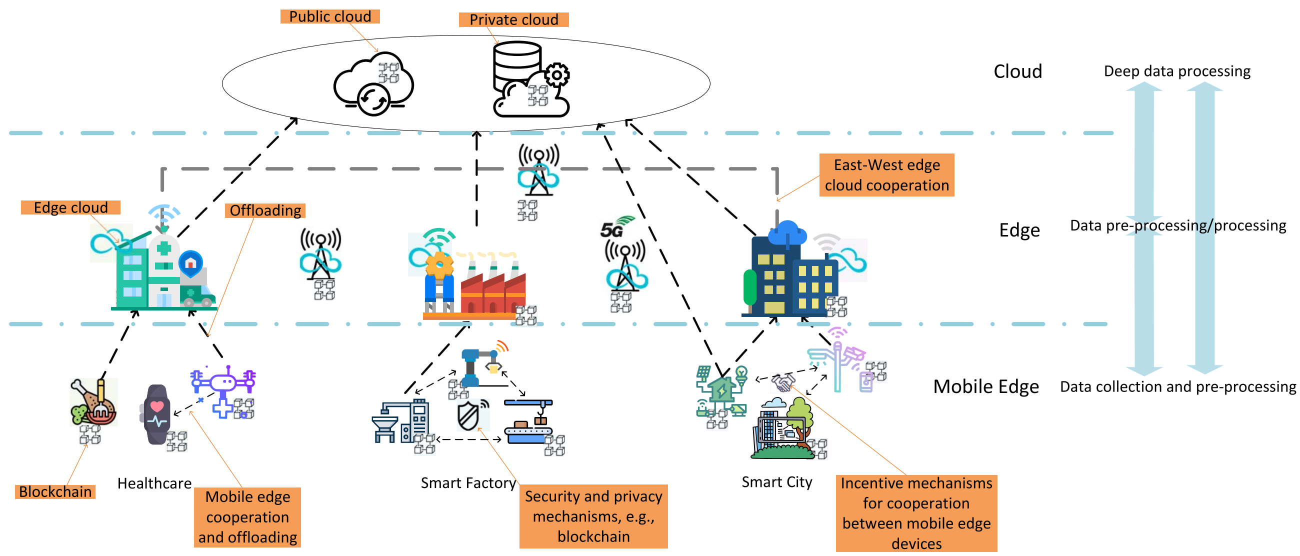


Fig. 2. Architecture for the cooperation among the cloud, the edge, and the mobile-edge computing.

coupled with medical robotics. This is particularly useful for remote surgery, especially in the era of 5G and B5G providing the necessary communication environment.

Another important use case of the IoT is in manufacturing, toward the so-called smart factory [54]. Coupling the IoT technologies with the manufacturing processes, enables automatic data collection and analysis. Results from the data analysis can allow the factory management to make better-informed decisions and achieve an optimized production process. The remote control of IoT devices, together with certain techniques, e.g., smart contract [55], provides the ability of automatic remote monitoring for the production environment and automatic decision making for any changes of the production plans. It can also monitor the whole lifecycle of product supply chain, enabling a more streamline and automatic management. IoT technologies can also be integrated into the products, creating the so-called smart products. On the one hand, the feedback information from the connected smart products can allow the factory management to improve the efficiency and intelligence of manufacturing processes. On the other hand, smart products can collect user data, e.g., how the products are used by a user, for their further improvement to meet the needs of the competitive market.

Efficient data handling, including data collection, data processing, and decision making, is necessary and crucial for many IoT applications [56]. First, to meet the diversified and stringent requirements from various IoT applications, e.g., ultralow latency and ultrahigh reliability, efficient data handling is a key factor. Second, the automation in business operations due to the involvement of the IoT, also relies on efficient data handling. With the fast growth of the number of IoT devices in different sectors, data handling process has placed significant pressures on computing and communication facilities. This calls for the innovation on computing and communication architectures for the IoT. Section III will present a promising architecture for the IoT through cloud-edge orchestration and investigate related works from the perspectives of IoT applications and practical issues.

III. Cloud-Edge Architecture for the IoT

In this section, the cloud-edge architecture will be elaborated and the related architecture proposals will be surveyed. The discussion will be mainly focused on the architecture, and the related data processing will be lightly touched in this section but will be discussed with more details in Section IV.

A. Cloud, Edge, and Mobile-Edge Computing

Cloud computing is able to provide on-demand availability of computing and storage resources. It is essentially located far away from computation tasks and belongs to one service provider, e.g., Microsoft and Google. Edge computing brings the cloud in proximity closer to the computation tasks. It essentially has less computing and storage resources than the remote cloud, and belongs to multiple service providers. Mobile-edge computing refers to the edge computing capabilities provided by mobile devices, e.g., IoT devices. Due to various computing and storage capabilities at different computing facilities (e.g., cloud, edge, and mobile edge), and the security and privacy concerns where data need to be transferred outside their origin for processing [57], the cooperation among the cloud, the edge, and the mobile-edge computing is necessary in practice, in order to satisfy the diversified and stringent requirements of various IoT applications.

Fig. 2 shows a typical architecture for the cooperation among the cloud, the edge, and the mobile-edge computing. The IoT devices with the dual roles of both data collection and data preprocessing, establish *mobile-edge computing layer*. IoT devices come from different verticals, e.g., wearable devices for healthcare, robotic arms for smart factory, mobile phones for Telecom, and AR/VR/MR for mixed verticals. Due to limited computing capabilities of IoT devices, computation tasks may have to be offloaded to mobile-edge computing devices (e.g., IoT devices), edge servers, and cloud data centers for processing. Incentive mechanisms are essential for the cooperation between IoT devices as they have different owners, and between edge servers as they may belong to different service providers [58]. Security and privacy preservation mechanisms need to be in place to guarantee that the data transmission can satisfy the corresponding requirements of IoT applications and comply data protection laws [59]–[63]. This can be considered from different aspects. For example, appropriate security mechanisms, e.g., blockchain, can be adopted to guarantee the transparency and security of IoT data and data transmission [64]–[69]. In addition, appropriate machine learning frameworks, e.g., federated learning [29], [70], can be enabled to ensure the data privacy by carrying out the data analysis locally. The basic idea of federated learning in the context of IoT, is to learn and improve the model in each IoT device using its own data. Any changes of the model in an IoT device are then summarized and sent to a cloud (e.g., edge cloud and remote data centers), using encrypted communication. The cloud then averages the updated results (e.g., model parameters) from all IoT devices as soon as they receive (no storage of the updates at the cloud), and improves the shared model. All the training data remain on the IoT devices, ensuring data privacy.

A number of recent studies have paid attention on the architecture for the cooperation of the cloud, the edge, and the mobile-edge computing. These works focus on different application areas (e.g., IIoT, mobile networks, and vehicular networks), various practical issues (e.g., multidomain and microservices), and network security and privacy considerations. In what follows, the state-of-the-art and typical studies of these architecture proposals will be investigated and discussed. A summary of these proposals is provided in Table I.

B. Industrial Networks

Due to the significant increase in the implementation of the IoT in industrial devices, industrial networks are moving toward automation in many aspects, e.g., industrial resource management, service upgrade, and operation policy implementation. A unique representation model for heterogeneous IIoT data is essential [88], [89]. Unlike modern IIoT devices, there are many legacy industrial devices that cannot be programmed. These legacy devices are therefore not easily being controlled, hindering the automatic implementation of service upgrades and policy changes. In addition, industrial networks have been deployed and operated for many years. They have their own architecture for legacy industrial automation. It is crucial to smoothly integrate modern control architecture and systems into the legacy control architecture of industrial networks.

Xia *et al.* [71] proposed an architecture for IIoT with a hierarchical control structure in the cloud-edge architecture. Software-defined networking (SDN) [90], [91] is used to separate the control plane and data plane in this architecture. The hierarchical SDN controllers are able to improve the intelligence and flexibility of the control plane of the architecture. Remote radio heads form a mobile-edge computing-based radio access network, improving the scalability and cooperation at data plane of the architecture. Deep learning techniques are implemented in the mobile-edge computing to enhance the edge intelligence.

Dai *et al.* [72] studied the industrial edge computing and how cloud-edge collaboration on edge servers can help meet the requirements of IIoT applications. The authors first discussed the mapping of the 5-level reference architecture of legacy industrial automation systems with the cloud-edge architecture. Levels 0–2 in the legacy industrial systems, focusing on industrial automation control and monitoring, have more real-time requirements, and thus they require edge computing to satisfy the needs. Levels 3 and 4, being used for manufacturing operation management and enterprise operations, have low real-time requirements, and thus cloud computing is more appropriate to meet the computing needs. The authors then proposed a reference architecture for the industrial systems with cloud and edge computing. The proposed architecture contains the following three layers.

1) *Top Layer:* The industrial cloud computing platforms are adopted to support a range of IIoT applications.
2) *Middle Layer:* The industrial edge gateway is in charge of managing data collection processes from edge servers, and balancing networking, computing, and storage resources.
3) *Base Layer:* This layer consists of edge servers.

C. Mobile and Vehicular Networks

The IoT devices in the context of mobile and vehicular networks may experience varying network conditions and Quality of Service (QoS) when moving across edge servers. It is therefore necessary to take into account such context when making decisions for the usage of the cloud and the edge resources. In addition, service continuity is another important factor that needs to be considered to ensure the required QoS for IoT applications in the context of mobile networks. How to ensure the end-to-end mobility support in the cloud-edge architecture is a hot research topic.

Guo *et al.* [73] proposed a context-aware object detection algorithm for vehicular networks, based on edge-cloud cooperation. The authors leveraged deep learning techniques to build an object detection model in the cloud server. The context information and the captured images extracted at the edge servers were used to train model parameters locally, and the results were used to adjust the object detection model in the cloud. The cooperation between the edge and the cloud improves the performance and adaptation of the detection model under various real-world settings.

Ghosh *et al.* [74] proposed a mobility-driven framework for real-time cloud–fog–edge collaboration.[5] It contains four layers, i.e., IoT layer, edge layer, fog layer, and cloud layer. The proposed framework exploited the mobility nature of a moving node, by analyzing the global positioning system (GPS) log

[5]It is worth mentioning that the term of fog computing was proposed to bring service provisioning in closer proximity to users [92]. It focuses on service provisioning instead of data processing which is the focus of edge computing.

TABLE I
SUMMARY OF THE STATE-OF-THE-ART PROPOSALS OF CLOUD-EDGE ARCHITECTURES FOR THE IOT

Applications and Practical Issues	**Existing Architecture Proposals**	**Features**
Industrial networks	Xia et al. [71]	A hierarchical control architecture based on SDN
	Dai et al. [72]	Mapping the reference architecture of legacy industrial automation systems with cloud-edge architecture
Mobile and vehicular networks	Guo et al. [73]	Context-awareness at the edge and deep analytics at the cloud
	Ghosh et al. [74]	Mobility-driven framework with real-time prediction of vehicle locations
	Liu et al. [75]	A three-layer architecture based on SDN
	Shah et al. [76]	Enhanced orchestration and management for mobile edge computing to provide better end-to-end mobility support
Healthcare	Muhammad et al. [77]	A classic cloud-edge architecture
	Baktir et al. [78]	Dynamic management based on SDN
Multi-domain cloud-edge	Taleb et al. [79]	Content delivery network (CDN) slice over multiple domains
Microservice	Villari et al. [80]	Osmotic computing
	Alam et al. [81]	A scalable and modular architecture based on lightweight virtualization
	Castellano et al. [82]	Each application can have its own orchestration strategy
	Dai et al. [83]	A microservice-based and knowledge-driven architecture to enable plug-and-play function components
	Yousefpour et al. [84]	QoS-aware service provisioning for dynamic deployment and release of services
Security and privacy	Diro and Chilamkurti [85]	A distributed deep learning model for the detection of cyber attacks in the environment of fog-to-things computing
	Nie et al. [25]	A differentially private tensor computing model for SDN-based IoT big data
	Guo et al. [86]	A trusted and automatic service function chain orchestration approach based on consortium blockchain and deep reinforcement learning
	Yang et al. [87]	A distributed machine learning architecture to enable fog intelligence for intelligent wireless network management with privacy preservation

data, the spatiotemporal mobility data, and other contextual information. Then, it relied on advanced machine learning algorithms to predict the location of the moving node (i.e., IoT and edge devices) in real time. The proposed framework can provide better QoS performance for IoT applications with real-time requirements.

Liu *et al.* [75] proposed a cloud-edge network based on SDN for mobile vehicles, with the aim of load balancing and low response delay. They introduced three layers: 1) *data center layer* that is used for performing function releasing, devices orchestration, and data aggregation; 2) *middle routing layer* that is designed for planning routers; and 3) *vehicle network layer* for transmitting data packets and services among devices.

Shah *et al.* [76] enhanced the orchestration and management of mobile-edge computing to provide better end-to-end mobility support that is needed to maintain service continuity when mobile users move across edge servers. The proposed solution integrated SDN and virtualization techniques with mobile-edge computing architecture.

D. Healthcare

Healthcare IoT devices are usually highly heterogeneous, since there are many different types of monitoring tasks in healthcare ecosystems. It is important to have customized orchestrations for different types of healthcare IoT devices in the cloud-edge architecture. Similar to the industrial networks, an efficient control to the significant number of healthcare IoT devices is crucial to ensure QoS of orchestration in the cloud-edge architecture.

Muhammad *et al.* [77] proposed a pathology detection system based on deep learning, edge computing, and cloud computing for smart healthcare. Specifically, electroencephalogram signals of a person were collected by sensors and sent to an edge server. The server preprocessed the received signals before transmitting them to a remote cloud server. The cloud server extracted deep features from the electroencephalogram signals using a tree-based deep learning model.

Baktir *et al.* [78] proposed a multitier computing and communication architecture amongst healthcare IoT devices, edge servers, and cloud data centers. SDN is adopted to carry out dynamic management of the proposed architecture, policy implementation, and service orchestration of healthcare IoT applications.

E. Multidomain Cloud-Edge Architecture

The cloud-edge architecture may be deployed at multiple network domains in a wide area network scenario, e.g., the Internet. For massive IoT data, some of them need to be processed at the edge, but lots of them need to be transferred to the cloud for deep analysis. The results from the cloud can be pushed back to the edge and even IoT for many operations, including configuration and optimization. This process involves the North–South connectivity. Distributed edge deployment may span multiple domains, resulting in the East–West connectivity.

Taleb *et al.* [79] developed an architecture for providing the video content delivery network functionality as a service (called CDN slice) over the cross-domain cloud-edge environment in the context of 5G mobile networks. The authors used network functions virtualization (NFV) and edge computing to drive resource allocation and management for ensuring QoS of CDN slices over multiple domains.

F. Service/Microservice-Oriented Architecture

Service-oriented architecture is promising to enhance flexibility and interoperability between cloud and edge services. Osmotic computing [80] was a term proposed to support data transfer protocols that enable seamless communication between the cloud and the edge. It aims to achieve dynamic management of services and microservices across the edge and the cloud, solving the issues related to deployment, networking, and security. These efforts were made to ensure the guaranteed QoS for IoT applications. Osmotic computing provides a federated environment for cloud providers, edge providers, IoT providers, and application providers, allowing them to work together to make the success of all the parties. The authors, who first proposed this computing paradigm, have identified a list of research directions, e.g., runtime microservice deployment, microservice configuration, microservice networking, microservice security, edge computing, microservice workload contention and interference evaluation, and microservice orchestration and elasticity control [80].

Alam *et al.* [81] proposed a modular and scalable architecture based on lightweight virtualisation. The proposed architecture runs on cloud, fog, and edge devices, and offers containerized services and microservices for the IoT using the Docker technique. The architecture ensures data collection and processing at the most appropriate places between the IoT devices and the cloud, making the computation and intelligence distributed at appropriate places over the entire network.

Castellano *et al.* [82] proposed a service-based approach for cloud-edge service orchestration. With the proposed approach, each application can then have its own orchestration strategy, e.g., an application can have different optimization criteria and use different reactions when handling the same event.

From the perspective of industrial edge computing, Dai *et al.* [83] focused on the feature of plug-and-play function components for industrial edges and proposed a microservice-based architecture to enable this feature. The function components can be dynamically configured for a service, based on the orchestration of microservices with the knowledge base and the reasoning process.

Yousefpour *et al.* [84] proposed a framework solution for QoS-aware dynamic service provisioning for fog computing. It was designed for achieving elastic deployment and release of application services on fog nodes, to meet application requirements, e.g., low latency.

G. Security and Privacy

Security and privacy are another main concerns for the cloud-edge architecture [93]–[95]. The cloud is usually operated by a third party and edge clouds may belong to different service providers. In order to achieve a smart decision to help with the automation of business operations, various data including user data, network operation data, and business operation data, may need to be transferred to the cloud and the edge for processing. Malicious data will result in incorrect decisions, affecting the performance of IoT applications eventually. Excessive exposure of data (or private data) can improve the effectiveness of the learning models, but it may violate data protection regulations/laws. How to ensure the security of data transmission and the privacy of user and business data, is the key factor to make the success and sustainability of the edge-cloud architecture.

Abeshu and Chilamkurti [85] considered the low accuracy and less scalability issues, caused by traditional cloud computing, for cyberattack detection mechanisms in the environment with massive IoT devices. They proposed a distributed deep learning model for the detection of cyberattacks in the environment of fog-to-things computing. Specifically, the authors developed a deep learning scheme based on stacked autoencoder as unsupervised deep learning. The deep learning scheme at the fog level needs to handle the model, parameters, and data distribution and update. A host-centric training scheme was also proposed to facilitate the efficient handling of distributed IoT data.

Nie *et al.* [25] considered the adoption of SDN for the management of IoT networks. Due to numerous communication demands to be taken place for efficient management, data privacy is a big concern. To tackle this problem, a differentially private tensor computing model was proposed to model and analyze such kinds of big data. They further devised an algorithm, called differentially private tensor decomposition, to achieve secure computing in SDN-based IoT. The proposed algorithm was able to maximize the resource usage and cooperation of cloud-edge computing.

Guo *et al.* [86] proposed a trusted and automatic service function chain orchestration approach for the cloud-edge environment. The approach used deep reinforcement learning and consortium blockchain techniques to minimize orchestration cost and improve QoS. In addition, a time-slotted model was devised to support dynamic service migration for IoT networks with high mobility in the heterogeneous cloud-edge environment.

Yang *et al.* [87] developed a distributed machine learning architecture to enable fog intelligence for intelligent wireless network management. It integrated both distributed edge processing and centralized cloud computing. The architecture also considered the privacy preservation while maintaining its scalability. The proposed architecture was particularly designed for network anomaly detection.

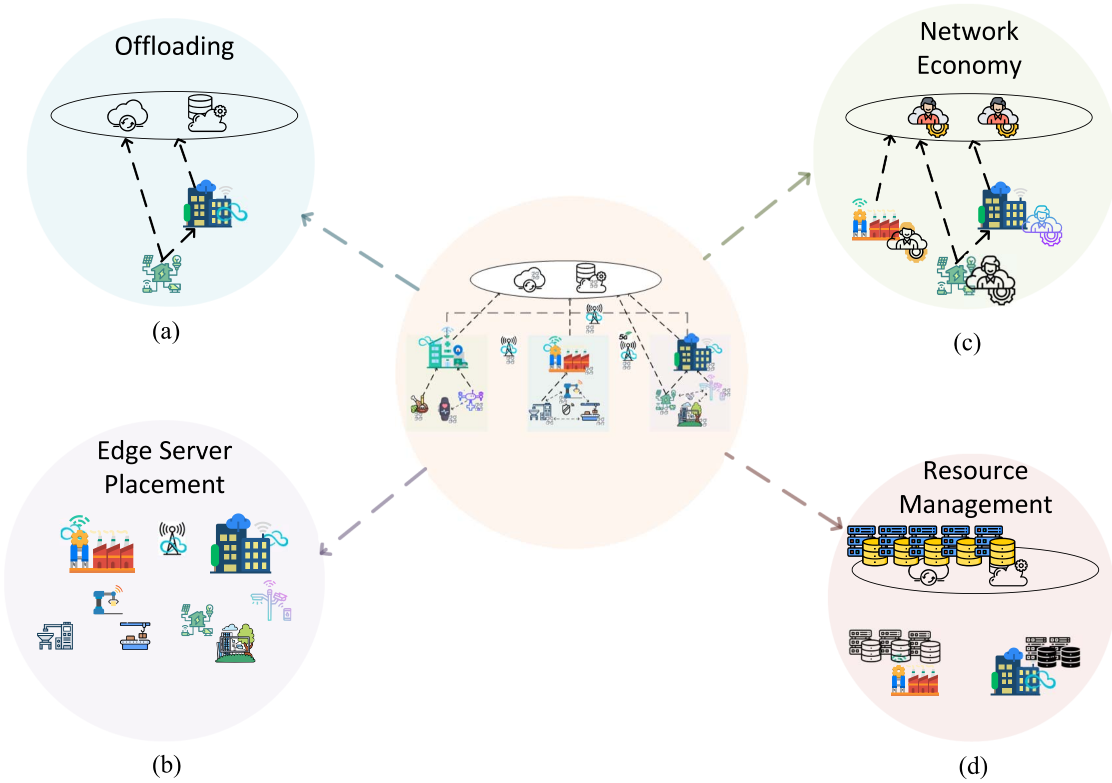


Fig. 3. Typical data processing use cases for cloud-edge architecture. (a) Offloading. (b) Edge server placement. (c) Network economy. (d) Resource management.

IV. AI-Powered Data Processing for the Cloud-Edge Architecture

In what follows, the state-of-the-art proposals and representative works of AI-powered data processing for cloud-edge orchestration will be surveyed and discussed. A set of typical use cases will be investigated, as shown in Fig. 3. A summary of these proposals is provided in Table II.

A. *Offloading*

Task offloading is one of the core components that enable the success of cloud-edge orchestration. It determines how much computation tasks need to be transferred to other mobile-edge devices, edge servers, and/or cloud servers [see Fig. 3(a)], in order to satisfy the stringent requirements of diversified IoT applications. The offloading decision is essentially subject to a number of constraints and factors, including the computing and storage capability of mobile-edge devices, edge servers and cloud servers, the communication delay (usually wireless communications), power consumption, and the requirements of IoT applications [111]–[113]. Machine learning has been a promising tool in recent years that can be adopted to make intelligent decisions for computation offloading.

Sun *et al.* [96] investigated the cloud-edge orchestration problem for IIoT. Since service accuracy is important to IIoT applications, in addition to delay and power consumption that have been widely taken into account in the literature, the authors considered service accuracy as a new performance metric. They proposed an AI-powered offloading framework with the aim of maximizing the service accuracy. The proposed framework was able to intelligently distribute the coming traffic from IIoT devices to either edge servers or remote cloud. The basic idea of the proposed framework was to enable a three-layer architecture: 1) IIoT layer; 2) edge layer; and 3) cloud layer. In the cloud layer, remote cloud pretrains network models. Edge servers in the edge layer then load the pretrained models from the cloud layer and further get them trained with domain data. After the assessment of service accuracy of the trained model in the edge layer, the tasks at the IIoT devices can be offloaded to appropriate edge servers.

Chen *et al.* [97] studied the use of edge computing to control traffic flow and investigated the way of transmitting only useful data to the remote cloud. The authors proposed a learning-based traffic control algorithm at the edge side to allow limited data to be offloaded to the cloud, while still maintaining the appropriate level of intelligence at the cloud side. An important contribution of this study is the label-less learning, as unlabeled data collection has become more practical in the networking environment. The basic idea of label-less learning includes the following steps.

1) A small amount of data with labels are used to train a model in order to enable the initial intelligence of the model.
2) The pretrained model is used to label the unlabeled data. A selection of the newly labeled data is added back into the training data set.
3) The additional selection of newly added data is obtained based on the mutual verification of multimodal data.
4) The model is trained again using the newly added labeled data.

TABLE II
SUMMARY OF THE STATE-OF-THE-ART PROPOSALS OF DATA PROCESSING FOR CLOUD-EDGE ARCHITECTURES

Practical Issues	**Existing Data Processing Proposals**	**Features**	**Application Areas**
Offloading	Sun et al. [96]	Taking into account delay, power consumption and service accuracy when making offloading decisions	Industrial networks
	Chen et al. [97]	A learning-based traffic control model with little labelled data, to offload only useful data	N/A
	Wang et al. [98]	Minimizing the power consumption of vehicles and computational facilities; A deep learning model was devised to obtain the optimal workload allocation	Vehicular networks
	Huang et al. [99]	Considering computational consumption, communication consumption and latency when making offloading decisions	N/A
Edge server placement	Rodrigues et al. [100]	To decrease the operational costs for service providers and the service price for clients, while minimizing the service delay	N/A
	Cao et al. [101]	Considering the heterogeneity of edge and cloud servers and the fairness of response time of base stations	Mobile networks
Network economy	Zhang et al. [23]	Considering a business model based on wholesale and buy-back scheme	N/A
	Luong et al. [102]	Leveraging deep learning techniques to develop an optimal auction scheme as an incentive mechanism for resource allocation	N/A
Microservice	Morshed et al. [103]	A holistic distributed deep learning approach that provides a detailed consideration of how deep learning technologies can be orchestrated across the edge and the cloud	N/A
	Renart et al. [104]	Automatically splitting and orchestration of IoT applications across the edge and the cloud resources	N/A
Resource management	Muñoz et al. [105]	Integrating the control of the packet and optical networks with the distributed edge and cloud resources to dynamically deploy IoT services, with the aim of reducing network bandwidth utilization	Video applications
	Zhou et al. [106]	An online orchestration framework for cross-edge service function chaining, with the aim of maximizing the cost efficiency through jointly optimizing the resource provisioning and traffic routing	N/A
	Na et al. [107]	A resource orchestration approach between edge gateways and edge servers	N/A
	Gilly et al. [108]	An edge computing orchestrator that can arrange location-based vehicular edge services through hierarchical dynamic resource management	Vehicular networks
	Roig et al. [109]	Applying deep reinforcement learning for the management and orchestration of the physical resources of virtual network functions	N/A
	Chien et al. [110]	Using suitable network models through AI to empower network intelligence, and leveraging the integration of edge computing and cloud computing to improve computing performance	Telecom networks

Wang *et al.* [98] developed an offloading algorithm in the cloud-edge environment for Internet of Vehicles. The design philosophy of the algorithm was to minimize the energy consumption of vehicles and the associated computing facilities. They formulated the offloading as an optimization problem and proposed a heuristic algorithm to solve. A deep learning model was devised to obtain the optimal workload allocation which was part of the heuristics considered in the proposed algorithm.

Huang *et al.* [99] discussed the challenges of SDN-based computing platform for IoT applications, in terms of meeting a set of requirements, including low latency and high reliability. They proposed a task offloading approach for the cloud-edge architecture with service orchestration. The proposed scheme considered computational complexity, communication overhead, and offloading latency. It can reach offloading decisions for tasks with diversified resource demands and delay requirements.

B. *Edge Server Placement*

Due to the decentralized nature of edge computing, edge servers are essentially geographically distributed to support a significant growing number of IoT devices. The placement of edge servers directly affects the performance of IoT applications, where Fig. 3(b) depicts a typical scenario. There are many factors that need to be considered in the design of an edge server placement algorithm, including scalability of deployment, heterogeneity of IoT devices, fairness of service provision, the density of servers, and service placement strategies [30], [113]–[115].

Rodrigues *et al.* [100] proposed an edge server deployment policy for cloud-edge computing, based on *k*-means clustering and particle swarm optimization. The objective of the policy is to decrease the operational costs for service providers and the service price for clients, while minimizing the service delay. In addition, the deployment policy should be scalable with the dramatic increase in the number of IoTs and the frequency of user requests in B5G/6G IoT environment. The proposed solution can simultaneously decide where edge servers should be deployed and how the computing resources are allocated to each IoT user.

Cao *et al.* [101] further investigated the edge server placement problem, by considering two important factors. One is the heterogeneity of edge and cloud servers, and the other is the fairness of response time of base stations. These two factors can significantly affect the performance of server placement algorithms, and thus the performance of IoT applications. The authors proposed an approach consisting of both an online stage and an offline stage. The offline stage adopted an integer linear programming (ILP) technique to achieve an optimal placement strategy of heterogeneous edge servers. The online stage developed a game theory-based algorithm to capture dynamic characteristics of user mobility.

C. Network Economy

Network economy considers economic factors in the information society. It is a key element affecting the success of the ecosystems of many IoT applications, especially for the cloud-edge architecture where many providers may be involved as shown in Fig. 3(c). Edge servers may belong to different providers and they may be selfish in serving certain IoT devices. Many solutions have been developed to address these issues. Considering the economic factors in these solutions is an efficient way of making them practical.

Zhang *et al.* [23] paid more attentions on the cooperation of cloud-edge architecture and the associated business models. The authors proposed a framework to allow the edge and the cloud to share their computing resources with each other based on the wholesale and buyback scheme. They considered two cases, where the edge and the cloud belong to the same provider and different providers, respectively. Then, they formulated the resource management between the edge and the cloud as a profit maximization problem. The authors solved this optimization problem from two points of view. One is from the perspective of social welfare maximization, through the optimal cloud resource management. The other is from the point of view of profit maximization for the edge and the cloud, through the optimal pricing and cloud resource management.

Luong *et al.* [102] investigated the study of incentive mechanisms for service providers to achieve the full potential of fog computing. The authors leveraged deep learning techniques to develop an optimal auction scheme as an incentive mechanism for the resource allocation in fog computing. The deep learning model was designed using feedforward neural networks. The economic and pricing model was considered in auction-based resource allocation. An application based on blockchain was used to validate the effectiveness of the proposed auction scheme.

D. Microservices and Data Transmission

Microservice is a variant of the service-oriented architecture and has been a widely used means for service deployment. In this architecture, an application is orchestrated as a suite of loosely coupled services. This modular way of service deployment facilitates the service management, in terms of configuration, fault detection, and prediction. Many challenges still exist in practical implementation, e.g., business awareness and decentralized control.

Morshed *et al.* [103] discussed the concept of osmotic computing and identified the current issues of this emerging paradigm. The authors then proposed a distributed deep learning model that provides a detailed consideration of how deep learning technologies can be orchestrated, and how this orchestration can take advantage of the cloud, the edge, and the mobile-edge environment. They also presented a list of challenges that need to be tackled in the development of deep learning models, in terms of the following aspects.

1) The heterogeneity of data sets coming from geographically distributed data sources, introduces significant complexity for deep learning.
2) The expression of a context through different terminologies, including clinical notes and test results, makes it hard for a deep learning model to identify patterns.
3) The deployment and configuration of an application, based on distributed deep learning, across the edge and the cloud nodes is a challenge.
4) It is not an easy task to evaluate which deep learning framework, such as TensorFlow and Keras, is suitable for the application based on a given deep learning model.
5) Deep learning models developed and used in different institutions may not be available to each other.

Renart *et al.* [104] investigated the splitting and orchestration of IoT applications across the edge and the cloud resources. They tackled the problem by exploring heterogeneous cloud-edge infrastructure for deploying dataflow applications, and considering the limitations of IIoT devices in terms of CPU, memory, and network bandwidth. The authors proposed a programming model based on R-Pulsar[6] to enable developers to define dataflow splitting across the edge and the cloud.

E. Resource Management

Proper distribution of IoT computations between the cloud and the edge can improve the efficient usage of network resources. In order to cope with various computation tasks at different levels of cloud-edge architecture, efficient and heterogeneous resource allocation and management schemes are

[6]https://rpulsar.rdi2.rutgers.edu

necessary, as shown in Fig. 3(d). Muñoz *et al.* [105] integrated the control and management of the packet and optical networks with the edge and cloud resources to enable the dynamic IoT service deployment. In the control and management, the authors developed flow monitors and congestion avoidance techniques, and also container-based edge nodes. Experimental results show up to 90% reduction of network bandwidth utilization through a case study of video analytics.

Zhou *et al.* [106] proposed an online orchestration framework to carry out cross-edge service function chaining. It aims to maximize the cost efficiency, through joint optimization of resource provisioning and traffic routing. The proposed solution was managing to fully unleash the benefits of service function chaining in geographically dispersed edge clouds. This cost optimization problem was solved by combining an online optimization technique with an approximate optimization method in a joint optimization framework.

Na *et al.* [107] considered an IoT scenario with edge servers and edge gateways, where an IoT device may connect to an edge server without the support of edge gateways. The authors proposed a resource orchestration approach between edge gateways and edge servers. The proposed approach can allocate optimal resources by taking into account of computing capacities of edge gateways and edge servers. It can also manage interference among the gateways to maximize the efficiency of IoT systems.

Gilly *et al.* [108] considered the demands of vehicular low-latency offloading in 5G and proposed an edge computing orchestrator that can arrange location-based vehicular edge services through hierarchical dynamic resource management. The proposed orchestrator is able to ensure low-latency responses due to energy-efficient service allocation and migration.

Roig *et al.* [109] applied deep reinforcement learning for the management and orchestration (MANO) of the physical resources of virtual network functions (VNFs). A central unit was proposed to learn to autonomously reconfigure computing and storage resources, deploy VNF instances, and offload the computation tasks to the edge and the cloud. The learning process considered the network conditions, available resources at the edge, and the VNF requirements. The optimization aim was to minimize a cost function that considers economical cost, latency and the Quality of Experience (QoE) of users.

Chien *et al.* [110] proposed an architecture for B5G heterogeneous networks. The architecture intelligently optimized network resource usage and network performance. It used suitable network models through AI to empower the network intelligence, and leveraged the integration of edge computing and cloud computing to improve computing performance. The authors also provided recommendations of which deep learning models are useful to handle various network issues.

V. Research Challenges and Open Issues

Although many studies have been presented to cover various aspects of cloud-edge orchestration for the IoT, there are still some challenges and open issues that need to be investigated and researched to unveil the full potential of this computing architecture and enable the success of IoT in wider sectors. In what follows, a list of potential challenges and open issues will be discussed and explained.

1) *Space, Air, Ground, and Sea Mobile Networks:* Most of the existing works are focused on air and ground networks, e.g., fixed edge servers at base stations or roof of the buildings, flying drones equipped with edge servers, and mobile IoT devices. B5G/6G expects to have collaboration among space, air, ground, and sea mobile networks [116]. How this mobile edge, edge and cloud cooperation architecture can be extended for the B5G/6G environment is still an open issue. The issues in the current cloud-edge orchestration in terms of offloading and security and privacy, would become more challenging in this complex architecture due to different communication and computing environment at space (e.g., satellite networks), air, ground, and sea mobile networks (e.g., marine-based edge computing and under sea networks).
2) *Mobility Awareness and Context Awareness:* Since IoT devices at the mobile-edge computing layer are mobile in nature, mobility-aware and context-aware solutions are needed when an IoT device is making decisions on task offloading to other nearby IoT devices [117], [118]. The decisions may include the size of a task to be partially offloaded, which nearby IoT devices need to be offloaded, and so on. In addition, mobility awareness and context awareness also need to be considered when a mobile IoT device offloads computation tasks to edge servers, especially, in the presence of mobile-edge servers (e.g., drones). Mobility patterns of IoT devices are usually not predictable in certain circumstances, making accurate and efficient offloading decisions a challenging task.
3) *Ultradense Environment:* One of the main use cases of 5G is to support massive IoT connections. Thus, the computation offloading needs to consider ultradense networks, where multiple IoT devices compete for the constrained computation resources and communication channels, e.g., offloading the computation tasks of mass IoT devices to one edge server or one mobile IoT device. How to selectively offload some computations to these resource-constrained devices, while maintaining the required QoS requirements of IoT applications, is still a difficult task.
4) *QoE Guarantee:* Application QoS may not directly affect the perceived QoE of IoT devices. To guarantee QoE, a number of questions need to be answered, e.g.,
   a) how to define QoE metrics of different applications in the cloud-edge orchestration architecture;
   b) how to link QoS with QoE by considering all possible operations over caching, prefetching, and task offloading at mobile-edge IoT devices, edge servers, and cloud servers;
   c) how to handle these issues when mobility presents at the mobile-edge layer.
5) *Mapping of Physical IoT Applications and Virtual Computing Resources:* IoT devices are cooperated in the

physical world to accomplish an application task. The task offloading to the virtual world of the cloud-edge orchestration architecture needs to consider the behavior of the task in the physical world, so that the application requirements can be satisfied in a holistic manner. How the mapping between physical world and virtual world can be integrated into the decision making for various operations in cloud-edge orchestration architecture is an open issue.

6) *Balancing Accountability and Privacy:* Accountability and privacy are two important but contradictory factors in cyber security [119]. Service providers expect the user traffic to be accountable. To achieve this purpose, receivers need to know who should get penalty when needed. In contrast, end users usually manage to protect their privacy by hiding their identities. This issue exists in the cooperation architecture of cloud, edge, and mobile-edge computing. The computing and storage services are provided in the form of microservices, which are then used to serve IoT devices for e.g., computation tasks. How to balance accountability and privacy in the cloud-edge orchestration architecture is still a challenge.
7) *Network Control and Management:* NFV and SDN are two main technologies of 5G and B5G/6G. How the computing services can be deployed at the cloud layer, the edge layer and the mobile-edge layer, and how the services can be migrated between different layers, under the SDN paradigm, are important to the practical operations of the cloud-edge orchestration architecture and are still hot research topics [120]. In addition, how the management of this architecture can be integrated into the MANO framework [121] of the NFV architecture is essential to the real-world deployment of this architecture. Furthermore, how the management framework of the cloud-edge architecture can cope with the edge servers belonging to different network providers is still an open issue.
8) *Autonomous Network Management:* Due to the complexity of this cloud-edge orchestration architecture, coupled with dynamic-in-nature services/applications, network management is becoming much harder than ever. Manual network management is obviously not practical. Autonomous management is the future direction in this area [122]. However, how to streamline the data collection, data processing, and decision making, and feed the performance of decision making back to data collection and processing to improve the whole process, is still an open research question.
9) *Lightweight Deep Learning:* Data processing is the key enabler of this cooperation computing architecture. Machine learning, especially deep learning, is a significant way for efficient data processing. How to ensure the efficiency of data processing at different layers of this cloud-edge orchestration architecture, subject to the resource and power constraints at different layers is important. Lightweight deep learning has been a promising solution for data processing at the resource-constrained IoT devices, due to the wide adoption of federated learning in the IoT. The design of lightweight deep learning models is limited by a number of factors, including computation capability, data features and properties, and application requirements. The research of designing an efficient lightweight deep learning model for cloud-edge orchestration architecture is still in the infant stage.
10) *Label-Less Training:* The training of learning models has been a big challenge in cyber networks, due to the dilemma of online labeling. Semi-supervised learning and unsupervised learning have been a trend in cyber networks to solve this issue [123]. The accuracy of these learning paradigms is always a challenge that researchers endeavor to address. Inaccurate models would yield less accurate decisions for the operation, e.g., offloading, in the cloud-edge orchestration architecture. How to ensure efficient label-less training and accurate decision making is still a challenging issue.
11) *Ethics in Deep Learning:* Deep learning models have been widely adopted in the cloud-edge orchestration architecture to enable the automation of many operations, including resource management, offloading, and edge server deployment. A deep learning model highly depends on its training data. If there are biases in the training data (e.g., anomaly data is significantly larger than normal data), the trained model may make unfair decisions for automatic system operations. There are many ethical issues in AI-powered data processing [124], especially in the process of autonomous network management. Addressing ethics is an open and very important issue to ensure the success and sustainability of the AI-powered cloud-edge orchestration architecture.
12) *Standardization:* Standardization of this cloud-edge orchestration architecture is important to its practical implementation and usage. How the development of this architecture can be integrated into the standardization of mainstream 5G and B5G/6G architecture in ETSI, 3GPP and IEEE is still under way.

VI. Conclusion

This article investigated the representative and state-of-the-art proposals for the architecture and its associated AI-powered data processing approaches in cloud-edge orchestration for the IoT. The investigated studies laid important foundations for the success of cloud-edge architecture in support of emerging IoT applications. However, there are still many research challenges and open issues, which were discussed at the end of this article, that need to keep a watchful eye on. These challenges and issues provided useful guidance for future research in this area.

References

[1] M. Roy and A. Roy, “Nexus of Internet of Things (IoT) and big data: Roadmap for smart management systems (SMGs),” *IEEE Eng. Manag. Rev.*, vol. 47, no. 2, pp. 53–65, May 2019.
[2] Y. Wu, H. Huang, C. Wang, and Y. Pan, *5G-Enabled Internet of Things*. Boca Raton, FL, USA: CRC Press, 2019.

[3] H.-N. Dai, H. Wang, G. Xu, J. Wan, and M. Imran, "Big data analytics for manufacturing Internet of Things: Opportunities, challenges and enabling technologies," *Enterprise Inf. Syst.*, to be published. [Online]. Available: https://doi.org/10.1080/17517575.2019.1633689

[4] M. A. Jan, W. Zhang, M. Usman, Z. Tan, F. Khan, and E. Luo, "SmartEdge: An end-to-end encryption framework for an edge-enabled smart city application," *J. Netw. Comput. Appl.*, vol. 137, pp. 1–10, Jul. 2019. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1084804519300827

[5] M. Babar, F. Arif, M. A. Jan, Z. Tan, and F. Khan, "Urban data management system: Towards big data analytics for Internet of Things based smart urban environment using customized Hadoop," *Future Gener. Comput. Syst.*, vol. 96, pp. 398–409, Jul. 2019. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0167739X18321095

[6] Q. Wang, H. Dai, Q. Wang, M. K. Shukla, W. Zhang, and C. G. Soares, "On connectivity of UAV-assisted data acquisition for underwater Internet of Things," *IEEE Internet Things J.*, vol. 7, no. 6, pp. 5371–5385, Mar. 2020.

[7] F. Liang, W. Yu, X. Liu, D. Griffith, and N. Golmie, "Towards edge-based deep learning in industrial Internet of Things," *IEEE Internet Things J.*, vol. 7, no. 5, pp. 4329–4341, May 2020.

[8] B. Yin, Y. Wu, T. Hu, J. Dong, and Z. Jiang, "An efficient collaboration and incentive mechanism for Internet of Vehicles (IoV) with secured information exchange based on blockchains," *IEEE Internet Things J.*, vol. 7, no. 3, pp. 1582–1593, Mar. 2020.

[9] P. Braun *et al.*, "Pattern mining from big IoT data with fog computing: Models, issues, and research perspectives," in *Proc. 19th IEEE/ACM Int. Symp. Cluster Cloud Grid Comput. (CCGRID)*, 2019, pp. 584–591.

[10] X. Zhou, W. Liang, K. I. Wang, and L. T. Yang, "Deep correlation mining based on hierarchical hybrid networks for heterogeneous big data recommendations," *IEEE Trans. Comput. Soc. Syst.*, early access, May 8, 2020, doi: 10.1109/TCSS.2020.2987846.

[11] C. Lai, T. Wang, C. Liu, and L. Wang, "Probabilistic top-*k* dominating query monitoring over multiple uncertain IoT data streams in edge computing environments," *IEEE Internet Things J.*, vol. 6, no. 5, pp. 8563–8576, Jun. 2019.

[12] C. Hu, Y. Deng, and L. T. Yang, "On-demand capacity provisioning in storage clusters through workload pattern modeling," *IEEE Access*, vol. 5, pp. 24830–24841, 2017.

[13] L. Zhao, G. Han, Z. Li, and L. Shu, "Intelligent digital twin-based software-defined vehicular networks," *IEEE Netw.*, early access, Jun. 5, 2020, doi: 10.1109/MNET.011.1900587.

[14] M. Noura, A. Gyrard, S. Heil, and M. Gaedke, "Automatic knowledge extraction to build semantic Web of Things applications," *IEEE Internet Things J.*, vol. 6, no. 5, pp. 8447–8454, Oct. 2019.

[15] Y. Yin, L. Long, and X. Deng, "Dynamic data mining of sensor data," *IEEE Access*, vol. 8, pp. 41637–41648, 2020.

[16] S. Fu, Z. Li, K. Liu, S. Din, M. Imran, and X. Yang, "Model compression for IoT applications in industry 4.0 via multi-scale knowledge transfer," *IEEE Trans. Ind. Informat.*, vol. 16, no. 9, pp. 6013–6022, Sep. 2020.

[17] J. Osborne, "Internet of Things and Cloud Computing," in *Internet of Things and Data Analytics Handbook*, H. Geng, Eds. Hoboken, NJ, USA: Wiley, 2017, ch. 42.

[18] K. Li, "Quantitative modeling and analytical calculation of elasticity in cloud computing," *IEEE Trans. Cloud Comput.*, early access, Feb. 7, 2017, doi: 10.1109/TCC.2017.2665549.

[19] C. Huang, G. Min, Y. Wu, Y. Ying, K. Pei, and Z. Xiang, "Time series anomaly detection for trustworthy services in cloud computing systems," *IEEE Trans. Big Data*, early access, Jun. 1, 2017, doi: 10.1109/TBDATA.2017.2711039.

[20] A. A. Diro, N. Chilamkurti, and Y. Nam, "Analysis of lightweight encryption scheme for fog-to-things communication," *IEEE Access*, vol. 6, pp. 26820–26830, 2018.

[21] M. Ma, G. Shi, and F. Li, "Privacy-oriented blockchain-based distributed key management architecture for hierarchical access control in the IoT scenario," *IEEE Access*, vol. 7, pp. 34045–34059, 2019.

[22] O. A. Khashan, "Hybrid lightweight proxy re-encryption scheme for secure fog-to-things environment," *IEEE Access*, vol. 8, pp. 66878–66887, 2020.

[23] Y. Zhang, X. Lan, J. Ren, and L. Cai, "Efficient computing resource sharing for mobile edge-cloud computing networks," *IEEE/ACM Trans. Netw.*, vol. 28, no. 3, pp. 1227–1240, Jun. 2020.

[24] O. Alkadi, N. Moustafa, B. Turnbull, and K. R. Choo, "A deep blockchain framework-enabled collaborative intrusion detection for protecting IoT and cloud networks," *IEEE Internet Things J.*, early access, May 22, 2020, doi: 10.1109/JIOT.2020.2996590.

[25] X. Nie, L. T. Yang, J. Feng, and S. Zhang, "Differentially private tensor train decomposition in edge-cloud computing for SDN-based Internet of Things," *IEEE Internet Things J.*, vol. 7, no. 7, pp. 5695–5705, Jul. 2020.

[26] S. Ramesh and M. Govindarasu, "An efficient framework for privacy-preserving computations on encrypted IoT data," *IEEE Internet Things J.*, early access, May 28, 2020, doi: 10.1109/JIOT.2020.2998109.

[27] C. Gong, F. Lin, X. Gong, and Y. Lu, "Intelligent cooperative edge computing in the Internet of Things," *IEEE Internet Things J.*, early access, Apr. 8, 2020, doi: 10.1109/JIOT.2020.2986015.

[28] J. Zhang, Y. Wu, G. Min, F. Hao, and L. Cui, "Balancing energy consumption and reputation gain of UAV scheduling in edge computing," *IEEE Trans. Cogn. Commun. Netw.*, early access, Jun. 24, 2020, doi: 10.1109/TCCN.2020.3004592.

[29] X. Wang, C. Wang, X. Li, V. C. M. Leung, and T. Taleb, "Federated deep reinforcement learning for Internet of Things with decentralized cooperative edge caching," *IEEE Internet Things J.*, early access, Apr. 9, 2020, doi: 10.1109/JIOT.2020.2986803.

[30] L. Chen, C. Shen, P. Zhou, and J. Xu, "Collaborative service placement for edge computing in dense small cell networks," *IEEE Trans. Mobile Comput.*, early access, Oct. 7, 2019, doi: 10.1109/TMC.2019.2945956.

[31] J. Liang, Y. Long, Y. Mei, T. Wang, and Q. Jin, "A distributed intelligent Hungarian algorithm for workload balance in sensor-cloud systems based on urban fog computing," *IEEE Access*, vol. 7, pp. 77649–77658, 2019.

[32] A. Asheralieva and D. Niyato, "Hierarchical game-theoretic and reinforcement learning framework for computational offloading in UAV-enabled mobile edge computing networks with multiple service providers," *IEEE Internet Things J.*, vol. 6, no. 5, pp. 8753–8769, Oct. 2019.

[33] K. Xue, J. Hong, Y. Ma, D. S. L. Wei, P. Hong, and N. Yu, "Fog-aided verifiable privacy preserving access control for latency-sensitive data sharing in vehicular cloud computing," *IEEE Netw.*, vol. 32, no. 3, pp. 7–13, May/Jun. 2018.

[34] A. B. Amor, M. Abid, and A. Meddeb, "Secure fog-based e-learning scheme," *IEEE Access*, vol. 8, pp. 31920–31933, 2020.

[35] J. Ren, D. Zhang, S. He, Y. Zhang, and T. Li, "A survey on end-edge-cloud orchestrated network computing paradigms: Transparent computing, mobile edge computing, fog computing, and cloudlet," *ACM Comput. Surveys*, vol. 52, no. 6, p. 125, Oct. 2019. [Online]. Available: https://doi.org/10.1145/3362031

[36] L. Jiao, Y. Wu, J. Dong, and Z. Jiang, "Toward optimal resource scheduling for Internet of Things under imperfect CSI," *IEEE Internet Things J.*, vol. 7, no. 3, pp. 1572–1581, Mar. 2020.

[37] X. Ge, R. Zhou, and Q. Li, "5G NFV-based tactile Internet for mission-critical IoT services," *IEEE Internet Things J.*, vol. 7, no. 7, pp. 6150–6163, Jul. 2020.

[38] C. Luo, J. Ji, Q. Wang, X. Chen, and P. Li, "Channel state information prediction for 5G wireless communications: A deep learning approach," *IEEE Trans. Netw. Sci. Eng.*, vol. 7, no. 1, pp. 227–236, Jan.–Mar. 2020.

[39] S. Wan, J. Lu, P. Fan, and K. B. Letaief, "Towards big data processing in IoT: Path planning and resource management of UAV base stations in mobile-edge computing system," *IEEE Internet Things J.*, vol. 7, no. 7, pp. 5995–6009, Jul. 2020.

[40] 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, 2018.

[41] Y. Lu, L. Liu, J. Panneerselvam, B. Yuan, J. Gu, and N. Antonopoulos, "A GRU-based prediction framework for intelligent resource management at cloud data centres in the age of 5G," *IEEE Trans. Cogn. Commun. Netw.*, vol. 6, no. 2, pp. 486–498, Jun. 2020.

[42] N. Javaid, A. Sher, H. Nasir, and N. Guizani, "Intelligence in IoT-based 5G networks: Opportunities and challenges," *IEEE Commun. Mag.*, vol. 56, no. 10, pp. 94–100, Oct. 2018.

[43] X. Zhou, W. Liang, K. I. Wang, H. Wang, L. T. Yang, and Q. Jin, "Deep learning enhanced human activity recognition for Internet of healthcare things," *IEEE Internet Things J.*, vol. 7, no. 7, pp. 6429–6438, Jul. 2020, doi: 10.1109/JIOT.2020.2985082.

[44] Y. Wu, F. Hu, G. Min, and A. Zomaya, *Big Data and Computational Intelligence in Networking*. Boca Raton, FL, USA: CRC Press, 2017.

[45] C. Zhang, M. Dong, and K. Ota, "Enabling computational intelligence for green Internet of Things: Data-driven adaptation in LPWA networking," *IEEE Comput. Intell. Mag.*, vol. 15, no. 1, pp. 32–43, Feb. 2020.

[46] C. Perera, C. H. Liu, and S. Jayawardena, "The emerging Internet of Things marketplace from an industrial perspective: A survey," *IEEE Trans. Emerg. Topics Comput.*, vol. 3, no. 4, pp. 585–598, Dec. 2015.

[47] H. Xu, W. Yu, D. Griffith, and N. Golmie, "A survey on industrial Internet of Things: A cyber-physical systems perspective," *IEEE Access*, vol. 6, pp. 78238–78259, 2018.

[48] J. Wu, S. Luo, S. Wang, and H. Wang, "NLES: A novel lifetime extension scheme for safety-critical cyber-physical systems using SDN and NFV," *IEEE Internet Things J.*, vol. 6, no. 2, pp. 2463–2475, Apr. 2019.

[49] N. Zhang *et al.*, "Physical layer authentication for Internet of Things via WFRFT-based Gaussian tag embedding," *IEEE Internet Things J.*, early access, Jun. 11, 2020, doi: 10.1109/JIOT.2020.3001597.

[50] L. Cheng *et al.*, "SCTSC: A semicentralized traffic signal control mode with attribute-based blockchain in IoVs," *IEEE Trans. Comput. Soc. Syst.*, vol. 6, no. 6, pp. 1373–1385, Dec. 2019.

[51] Y. A. Qadri, A. Nauman, Y. B. Zikria, A. V. Vasilakos, and S. W. Kim, "The future of healthcare Internet of Things: A survey of emerging technologies," *IEEE Commun. Surveys Tuts.*, vol. 22, no. 2, pp. 1121–1167, 2nd Quart., 2020.

[52] M. Wehde, "Healthcare 4.0," *IEEE Eng. Manag. Rev.*, vol. 47, no. 3, pp. 24–28, Jul. 2019.

[53] X. Guo, H. Lin, Y. Wu, and M. Peng, "A new data clustering strategy for enhancing mutual privacy in healthcare IoT systems," *Future Gener. Comput. Syst.*, vol. 113, pp. 407–417, Dec. 2020.

[54] M. Compare, P. Baraldi, and E. Zio, "Challenges to IoT-enabled predictive maintenance for industry 4.0," *IEEE Internet Things J.*, vol. 7, no. 5, pp. 4585–4597, May 2020.

[55] W. Zou *et al.*, "Smart contract development: Challenges and opportunities," *IEEE Trans. Softw. Eng.*, early access, Sep. 24, 2019, doi: 10.1109/TSE.2019.2942301.

[56] S. Li, L. D. Xu, and X. Wang, "Compressed sensing signal and data acquisition in wireless sensor networks and Internet of Things," *IEEE Trans. Ind. Informat.*, vol. 9, no. 4, pp. 2177–2186, Jan. 2013.

[57] M. Babar *et al.*, "A secured data management scheme for smart societies in industrial Internet of Things environment," *IEEE Access*, vol. 6, pp. 43088–43099, 2018.

[58] L. Liu, N. Antonopoulos, M. Zheng, Y. Zhan, and Z. Ding, "A socioecological model for advanced service discovery in machine-to-machine communication networks," *ACM Trans. Embed. Comput. Syst.*, vol. 15, no. 2, p. 38, Mar. 2016. [Online]. Available: https://doi.org/10.1145/2811264

[59] L. Cui, S. Yang, Z. Chen, Y. Pan, Z. Ming, and M. Xu, "A decentralized and trusted edge computing platform for Internet of Things," *IEEE Internet Things J.*, vol. 7, no. 5, pp. 3910–3922, May 2020.

[60] L. Cui, S. Yang, Z. Chen, Y. Pan, M. Xu, and K. Xu, "An efficient and compacted DAG-based blockchain protocol for industrial Internet of Things," *IEEE Trans. Ind. Informat.*, vol. 16, no. 6, pp. 4134–4145, Jun. 2020.

[61] L. Malina, G. Srivastava, P. Dzurenda, J. Hajny, and R. Fujdiak, "A secure publish/subscribe protocol for Internet of Things," in *Proc. 14th Int. Conf. Availability Rel. Security*, 2019, pp. 1–10.

[62] G. Srivastava, R. M. Parizi, A. Dehghantanha, and K.-K. R. Choo, "Data sharing and privacy for patient IoT devices using blockchain," in *Proc. Int. Conf. Smart City Informatization*, 2019, pp. 334–348.

[63] G. Srivastava, J. Crichigno, and S. Dhar, "A light and secure healthcare blockchain for IoT medical devices," in *Proc. IEEE Can. Conf. Elect. Comput. Eng. (CCECE)*, 2019, pp. 1–5.

[64] Y. Xiao, N. Zhang, W. Lou, and Y. T. Hou, "A survey of distributed consensus protocols for blockchain networks," *IEEE Commun. Surveys Tuts.*, vol. 22, no. 2, pp. 1432–1465, 2nd Quart., 2020.

[65] S. Seng, X. Li, C. Luo, H. Ji, and H. Zhang, "A D2D-assisted MEC computation offloading in the blockchain-based framework for UDNs," in *Proc. IEEE Int. Conf. Commun. (ICC)*, 2019, pp. 1–6.

[66] H. Dai, Z. Zheng, and Y. Zhang, "Blockchain for Internet of Things: A survey," *IEEE Internet Things J.*, vol. 6, no. 5, pp. 8076–8094, Jun. 2019.

[67] S. Li, S. Zhao, P. Yang, P. Andriotis, L. Xu, and Q. Sun, "Distributed consensus algorithm for events detection in cyber-physical systems," *IEEE Internet Things J.*, vol. 6, no. 2, pp. 2299–2308, Mar. 2019.

[68] S. Li, K. R. Choo, Q. Sun, W. J. Buchanan, and J. Cao, "IoT forensics: Amazon echo as a use case," *IEEE Internet Things J.*, vol. 6, no. 4, pp. 6487–6497, Mar. 2019.

[69] J. Wu, M. Dong, K. Ota, J. Li, and W. Yang, "Application-aware consensus management for software-defined intelligent blockchain in IoT," *IEEE Netw.*, vol. 34, no. 1, pp. 69–75, Apr. 2020.

[70] B. Yin, H. Yin, Y. Wu, and Z. Jiang, "FDC: A secure federated deep learning mechanism for data collaborations in the Internet of Things," *IEEE Internet Things J.*, vol. 7, no. 7, pp. 6348–6359, Jul. 2020.

[71] W. Xia, J. Zhang, T. Q. S. Quek, S. Jin, and H. Zhu, "Mobile edge cloud-based industrial Internet of Things: Improving edge intelligence with hierarchical SDN controllers," *IEEE Veh. Technol. Mag.*, vol. 15, no. 1, pp. 36–45, Mar. 2020.

[72] W. Dai, H. Nishi, V. Vyatkin, V. Huang, Y. Shi, and X. Guan, "Industrial edge computing: Enabling embedded intelligence," *IEEE Ind. Electron. Mag.*, vol. 13, no. 4, pp. 48–56, Dec. 2019.

[73] J. Guo, B. Song, S. Chen, F. R. Yu, X. Du, and M. Guizani, "Context-aware object detection for vehicular networks based on edge-cloud cooperation," *IEEE Internet Things J.*, vol. 7, no. 7, pp. 5783–5791, Jul. 2020.

[74] S. Ghosh, A. Mukherjee, S. K. Ghosh, and R. Buyya, "Mobi-IoST: Mobility-aware cloud-fog-edge-IoT collaborative framework for time-critical applications," *IEEE Trans. Netw. Sci. Eng.*, early access, Sep. 16, 2019, doi: 10.1109/TNSE.2019.2941754.

[75] Y. Liu, Z. Zeng, X. Liu, X. Zhu, and M. Z. A. Bhuiyan, "A novel load balancing and low response delay framework for edge-cloud network based on SDN," *IEEE Internet Things J.*, vol. 7, no. 7, pp. 5922–5933, Jul. 2020.

[76] S. D. A. Shah, M. A. Gregory, S. Li, and R. D. R. Fontes, "SDN enhanced multi-access edge computing (MEC) for E2E mobility and QoS management," *IEEE Access*, vol. 8, pp. 77459–77469, 2020.

[77] G. Muhammad, M. F. Alhamid, and X. Long, "Computing and processing on the edge: Smart pathology detection for connected healthcare," *IEEE Netw.*, vol. 33, no. 6, pp. 44–49, Dec. 2019.

[78] A. C. Baktir, C. Tunca, A. Ozgovde, G. Salur, and C. Ersoy, "SDN-based multi-tier computing and communication architecture for pervasive healthcare," *IEEE Access*, vol. 6, pp. 56765–56781, 2018.

[79] T. Taleb, P. A. Frangoudis, I. Benkacem, and A. Ksentini, "CDN slicing over a multi-domain edge cloud," *IEEE Trans. Mobile Comput.*, vol. 19, no. 9, pp. 2010–2027, Sep. 2020.

[80] M. Villari, M. Fazio, S. Dustdar, O. Rana, and R. Ranjan, "Osmotic computing: A new paradigm for edge/cloud integration," *IEEE Cloud Comput.*, vol. 3, no. 6, pp. 76–83, Nov./Dec. 2016.

[81] M. Alam, J. Rufino, J. Ferreira, S. H. Ahmed, N. Shah, and Y. Chen, "Orchestration of microservices for IoT using docker and edge computing," *IEEE Commun. Mag.*, vol. 56, no. 9, pp. 118–123, Sep. 2018.

[82] G. Castellano, F. Esposito, and F. Risso, "A service-defined approach for orchestration of heterogeneous applications in cloud/edge platforms," *IEEE Trans. Netw. Service Manag.*, vol. 16, no. 4, pp. 1404–1418, Sep. 2019.

[83] W. Dai *et al.*, "Semantic integration of plug-and-play software components for industrial edges based on microservices," *IEEE Access*, vol. 7, pp. 125882–125892, 2019.

[84] A. Yousefpour *et al.*, "FogPlan: A lightweight QoS-aware dynamic fog service provisioning framework," *IEEE Internet Things J.*, vol. 6, no. 3, pp. 5080–5096, Jan. 2019.

[85] A. Abeshu and N. Chilamkurti, "Deep learning: The Frontier for distributed attack detection in fog-to-things computing," *IEEE Commun. Mag.*, vol. 56, no. 2, pp. 169–175, Feb. 2018.

[86] S. Guo, Y. Dai, S. Xu, X. Qiu, and F. Qi, "Trusted cloud-edge network resource management: DRL-driven service function chain orchestration for IoT," *IEEE Internet Things J.*, vol. 7, no. 7, pp. 6010–6022, Jul. 2020.

[87] K. Yang, H. Ma, and S. Dou, "Fog intelligence for network anomaly detection," *IEEE Netw.*, vol. 34, no. 2, pp. 78–82, Apr. 2020.

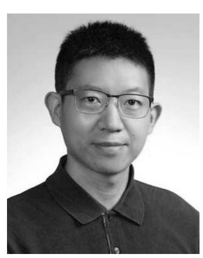
[88] X. Wang, L. T. Yang, Y. Wang, L. Ren, and M. J. Deen, "ADTT: A highly-efficient distributed tensor-train decomposition method for IIoT big data," *IEEE Trans. Ind. Informat.*, early access, Jan. 20, 2020, doi: 10.1109/TII.2020.2967768.

[89] X. Wang, L. T. Yang, L. Song, H. Wang, L. Ren, and J. Deen, "A tensor-based multi-attributes visual feature recognition method for industrial intelligence," *IEEE Trans. Ind. Informat.*, early access, Jun. 4, 2020, doi: 10.1109/TII.2020.2999901.

[90] W. Miao, G. Min, Y. Wu, H. Wang, and J. Hu, "Performance modelling and analysis of software-defined networking under bursty multimedia traffic," *ACM Trans. Multimedia Comput. Commun. Appl.*, vol. 12, no. 5s, p. 77, Sep. 2016. [Online]. Available: https://doi.org/10.1145/2983637

[91] H. Huang, H. Yin, G. Min, H. Jiang, J. Zhang, and Y. Wu, "Data-driven information plane in software-defined networking," *IEEE Commun. Mag.*, vol. 55, no. 6, pp. 218–224, Mar. 2017.

[92] D. Miao, L. Liu, R. Xu, J. Panneerselvam, Y. Wu, and W. Xu, "An efficient indexing model for the fog layer of industrial Internet of Things," *IEEE Trans. Ind. Informat.*, vol. 14, no. 10, pp. 4487–4496, Jan. 2018.
[93] P. Zhou, W. Chen, S. Ji, H. Jiang, L. Yu, and D. Wu, "Privacy-preserving online task allocation in edge-computing-enabled massive crowdsensing," *IEEE Internet Things J.*, vol. 6, no. 5, pp. 7773–7787, Oct. 2019.
[94] P. Zhou, K. Wang, J. Xu, and D. Wu, "Differentially-private and trustworthy online social multimedia big data retrieval in edge computing," *IEEE Trans. Multimedia*, vol. 21, no. 3, pp. 539–554, Dec. 2019.
[95] N. Zhang, R. Wu, S. Yuan, C. Yuan, and D. Chen, "RAV: Relay aided vectorized secure transmission in physical layer security for Internet of Things under active attacks," *IEEE Internet Things J.*, vol. 6, no. 5, pp. 8496–8506, Mar. 2019.
[96] W. Sun, J. Liu, and Y. Yue, "AI-enhanced offloading in edge computing: When machine learning meets industrial IoT," *IEEE Netw.*, vol. 33, no. 5, pp. 68–74, Jun. 2019.
[97] M. Chen, Y. Hao, K. Lin, Z. Yuan, and L. Hu, "Label-less learning for traffic control in an edge network," *IEEE Netw.*, vol. 32, no. 6, pp. 8–14, Aug. 2018.
[98] X. Wang, X. Wei, and L. Wang, "A deep learning based energy-efficient computational offloading method in Internet of Vehicles," *China Commun.*, vol. 16, no. 3, pp. 81–91, 2019.
[99] M. Huang, W. Liu, T. Wang, A. Liu, and S. Zhang, "A cloud-MEC collaborative task offloading scheme with service orchestration," *IEEE Internet Things J.*, vol. 7, no. 7, pp. 5792–5805, Jul. 2020.
[100] T. K. Rodrigues, K. Suto, and N. Kato, "Edge cloud server deployment with transmission power control through machine learning for 6G Internet of Things," *IEEE Trans. Emerg. Topics Comput.*, early access, Dec. 31, 2016, doi: 10.1109/TETC.2019.2963091.
[101] K. Cao, L. Li, Y. Cui, T. Wei, and S. Hu, "Exploring placement of heterogeneous edge servers for response time minimization in mobile edge-cloud computing," *IEEE Trans. Ind. Informat.*, early access, Feb. 24, 2020, doi: 10.1109/TII.2020.2975897.
[102] N. C. Luong, Y. Jiao, P. Wang, D. Niyato, D. I. Kim, and Z. Han, "A machine-learning-based auction for resource trading in fog computing," *IEEE Commun. Mag.*, vol. 58, no. 3, pp. 82–88, Mar. 2020.
[103] A. Morshed *et al.*, "Deep osmosis: Holistic distributed deep learning in osmotic computing," *IEEE Cloud Comput.*, vol. 4, no. 6, pp. 22–32, Nov. 2017.
[104] E. G. Renart *et al.*, "Distributed operator placement for IoT data analytics across edge and cloud resources," in *Proc. 19th IEEE/ACM Int. Symp. Clust. Cloud Grid Comput. (CCGRID)*, 2019, pp. 459–468.
[105] R. Muñoz *et al.*, "Integration of IoT, transport SDN, and edge/cloud computing for dynamic distribution of IoT analytics and efficient use of network resources," *J. Lightw. Technol.*, vol. 36, no. 7, pp. 1420–1428, Apr. 1, 2018.
[106] Z. Zhou, Q. Wu, and X. Chen, "Online orchestration of cross-edge service function chaining for cost-efficient edge computing," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 8, pp. 1866–1880, Aug. 2019.
[107] W. Na, S. Jang, Y. Lee, L. Park, N. Dao, and S. Cho, "Frequency resource allocation and interference management in mobile edge computing for an Internet of Things system," *IEEE Internet Things J.*, vol. 6, no. 3, pp. 4910–4920, Jun. 2019.
[108] K. Gilly, A. Mishev, S. Filiposka, and S. Alcaraz, "Offloading edge vehicular services in realistic urban environments," *IEEE Access*, vol. 8, pp. 11491–11502, 2020.
[109] J. S. P. Roig, D. M. Gutierrez-Estevez, and D. Gündüz, "Management and orchestration of virtual network functions via deep reinforcement learning," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 2, pp. 304–317, Feb. 2020.
[110] W.-C. Chien *et al.*, "Intelligent architecture for mobile HetNet in B5G," *IEEE Netw.*, vol. 33, no. 3, pp. 34–41, May/Jun. 2019.
[111] L. Cui *et al.*, "Joint optimization of energy consumption and latency in mobile edge computing for Internet of Things," *IEEE Internet Things J.*, vol. 6, no. 3, pp. 4791–4803, Jun. 2019.
[112] L. Cui, D. Su, Y. Zhou, L. Zhang, Y. Wu, and S. Chen, "Edge learning for surveillance video uploading sharing in public transport systems," *IEEE Trans. Intell. Transp. Syst.*, early access, Jul. 31, 2020, doi: 10.1109/TITS.2020.3008420.
[113] Z. Zhao, G. Min, W. Gao, Y. Wu, H. Duan, and Q. Ni, "Deploying edge computing nodes for large-scale IoT: A diversity aware approach," *IEEE Internet Things J.*, vol. 5, no. 5, pp. 3606–3614, Oct. 2018.
[114] S. Lou, G. Srivastava, and S. Liu, "A node density control learning method for the Internet of Things," *Sensors*, vol. 19, no. 15, p. 3428, 2019.
[115] L. Chen, J. Xu, S. Ren, and P. Zhou, "Spatio–temporal edge service placement: A bandit learning approach," *IEEE Trans. Wireless Commun.*, vol. 17, no. 12, pp. 8388–8401, Oct. 2018.
[116] F. Xu, F. Yang, C. Zhao, and S. Wu, "Deep reinforcement learning based joint edge resource management in maritime network," *China Commun.*, vol. 17, no. 5, pp. 211–222, 2020.
[117] K. Cao, G. Xu, J. Zhou, T. Wei, M. Chen, and S. Hu, "QoS-adaptive approximate real-time computation for mobility-aware IoT lifetime optimization," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 38, no. 10, pp. 1799–1810, Oct. 2019.
[118] H. Gao, Y. Xu, Y. Yin, W. Zhang, R. Li, and X. Wang, "Context-aware QoS prediction with neural collaborative filtering for Internet-of-Things services," *IEEE Internet Things J.*, vol. 7, no. 5, pp. 4532–4542, May 2020.
[119] Y. Ma, Y. Wu, J. Li, and J. Ge, "APCN: A scalable architecture for balancing accountability and privacy in large-scale content-based networks," *Inf. Sci.*, vol. 527, pp. 511–532, Jul. 2020. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0020025519300659
[120] Z. Lv and W. Xiu, "Interaction of edge-cloud computing based on SDN and NFV for next generation IoT," *IEEE Internet Things J.*, vol. 7, no. 7, pp. 5706–5712, Jul. 2020.
[121] M. Kourtis *et al.*, "T-NoVa: An open-source MANO stack for NFV infrastructures," *IEEE Trans. Netw. Service Manag.*, vol. 14, no. 3, pp. 586–602, Jul. 2017.
[122] M. Bunyakitanon, X. Vasilakos, R. Nejabati, and D. Simeonidou, "End-to-end performance-based autonomous VNF placement with adopted reinforcement learning," *IEEE Trans. Cogn. Commun. Netw.*, vol. 6, no. 2, pp. 534–547, Jun. 2020.
[123] N. Zhao, J. Zhu, R. Liu, D. Liu, M. Zhang, and D. Pei, "Label-less: A semi-automatic labelling tool for KPI anomalies," in *Proc. IEEE Conf. Comput. Commun. (IEEE INFOCOM)*, 2019, pp. 1882–1890.
[124] N. Bostrom and E. Yudkowsky, "The ethics of artificial intelligence," *The Cambridge Handbook of Artificial Intelligence*, vol. 1. Cambridge, U.K.: Cambridge Univ. Press, 2014, pp. 316–334.



Yulei Wu (Senior Member, IEEE) received the B.Sc. degree (First Class Hons.) in computer science and the Ph.D. degree in computing and mathematics from the University of Bradford, Bradford, U.K., in 2006 and 2010, respectively.

He is a Senior Lecturer with the Department of Computer Science, College of Engineering, Mathematics and Physical Sciences, University of Exeter, Exeter, U.K. His expertise is on intelligent networking and his main research interests include computer networks, networked systems, software-defined networks and systems, network management, and network security and privacy.

Dr. Wu is an Editor of the IEEE TRANSACTIONS ON NETWORK AND SERVICE MANAGEMENT, the IEEE TRANSACTIONS ON NETWORK SCIENCE AND ENGINEERING, and *Computer Networks* (Elsevier).