We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900
Open access books available

185,000
International authors and editors

200M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



Chapter

Trends in Cloud Computing Paradigms: Fundamental Issues, Recent Advances, and Research Directions toward 6G Fog Networks

Isiaka A. Alimi, Romil K. Patel, Aziza Zaouga, Nelson J. Muga, Qin Xin, Armando N. Pinto and Paulo P. Monteiro

Abstract

There has been significant research interest in various computing-based paradigms such as cloud computing, Internet of Things, fog computing, and edge computing, due to their various associated advantages. In this chapter, we present a comprehensive review of these architectures and their associated concepts. Moreover, we consider different enable technologies that facilitate computing paradigm evolution. In this context, we focus mainly on fog computing considering its related fundamental issues and recent advances. Besides, we present further research directions toward the sixth generation fog computing paradigm.

Keywords: 5G, 6G, cloud computing, edge computing, fog computing, Internet of Things, mobile edge computing

1. Introduction

The need to achieve excellent Quality of Service (QoS) to facilitate effective Quality of Experience (QoE) is one of the notable factors that has brought about substantial evolution in the computing paradigms. For instance, the cloud computing paradigm has been presented to ensure an effective development and delivery of various innovative Internet services [1]. Also, the unprecedented development of various applications and growing smart mobile devices for supporting Internet-of-Things (IoT) have presented significant constraints regarding latency, bandwidth, and connectivity on the centralized-based paradigm of cloud computing [2–4]. To address the limitations, research interests have been shifting toward decentralized paradigms [2].

A good instance of a decentralized paradigm is edge computing. Conceptually, edge computing focuses on rendering several services at the network edge to

alleviate the associated limitations of cloud computing. Also, a number of such edge computing implementations such as cloudlet computing (CC), mobile cloud computing (MCC), and mobile edge computing (MEC) have been presented [2, 5–7]. Besides, another edge computing evolution is fog computing. It offers an efficient architecture that mainly focuses on both horizontal and vertical resource distribution in the Cloud-to-Things continuum [2]. In this light, it goes beyond mere cloud extension but serves as a merging platform for both cloud and IoT to facilitate and ensure effective interaction in the system. Nevertheless, these paradigms demand further research efforts due to the required resource management that is demanding and the massive traffic to be supported by the network. For instance, fog nodes are typically equipped with limited computing and storage resources which may prevent them from being a good solution for supporting and meeting requests of large-scale users. Conversely, cloud resources are usually deployed far away from users, which makes cloud servers unable to support services that demand low-latency. Based on this, there is a need for the integration of fog and other cloud-based computing platforms with an effective multiple access technique for efficient resource management across the fog-cloud platform. In this regard, the overall performance can be improved and effective computation offloading can be offered. One of such schemes for performance enhancement is non-orthogonal multiple access (NOMA) [8].

In addition, there have been significant research efforts toward the sixth generation (6G) networks. Also, it is envisaged that various technologies such as device-to-device communications, Big Data, cloud computing, edge caching, edge computing, and IoT will be well-supported by the 6G mobile networks [9]. Meanwhile, 6G is envisioned to be based on major innovative technologies such as super IoT, mobile ultra-broadband, and artificial intelligence (AI) [3, 10]. Besides, it is envisaged that terahertz (THz) communications should be a viable solution for supporting mobile ultra-broadband. Also, super IoT can be achieved with symbiotic radio and satellite-assisted communications. Besides, machine learning (ML) methods are expected to be promising solutions for AI networks [10]. Based on the innovative technologies, beyond 5G network is envisaged to offer a considerable improvement on the 5G by employing AI to automate and optimize the system operation [11].

This chapter presents various evolutions of computing paradigms and highlights their associated features. Also, different related technological implementations are comprehensively discussed. Besides, it presents different models that focus on effective resource allocation across an integrated computing platform for performance enhancement. Moreover, it presents AI as a resourceful technique for the achievement of high-level automation for efficient management and optimization of the 6G fog computing platform. This chapter is organized as follows. Section 2 presents a comprehensive discussion on the evolution of computing paradigms with related concepts and features. Section 3 focuses on the fog architectural model. We discuss the challenges of fog computing and its integration with other computing platforms in Section 4. In Section 5, we present some models for resource allocation in an integrated fog-cloud hierarchical architecture. Section 6 focuses on the trends toward intelligent integrated computing networks and concluding remarks are given in Section 7.

2. Evolution of computing paradigms

This section presents the evolution of computing paradigms. In this regards, related concepts, features, and architectural models are considered.

2.1 Cloud computing

As aforementioned, cloud computing has been in the mainstream of research and has been revolutionizing the information and communication technology (ICT) sector. Based on the National Institute of Standards and Technology (NIST) definition, cloud computing presents an enabling platform that offers ubiquitous and on-demand network access to a shared pool of computing resources such as storages, servers, networks, applications, and services. These interconnected resource pools can be conveniently configured and provisioned with minimal interaction. Besides cost-effectiveness regarding support for pay-per-use policy and expenditure savings, some of the key inducements for the adoption of the cloud computing paradigm are easy and ubiquitous access to applications and data [12].

It is noteworthy that with the cloud computing paradigm, network entities regarding control, computing, and data storage are centralized in the cloud. For instance, storage, computing, and network management functions have been moved to different network places such as backbone IP networks, centralized data centers, and cellular core networks [13]. However, it is challenging for the centralized cloud model to meet the stringent requirements of the emerging IoT. The IoT comprises varieties of computing devices that are connected through the Internet to support a variety of applications and services [2, 13]. In this context, things such as smart meters, tablets, smartphones, robots, wireless routers, sensors, actuators, smart vehicles, and radio-frequency identification (RFID) tags are Internet-connected to ensure a more convenient standard of living [2, 14–16]. Therefore, the centralized-based paradigm offered by cloud computing is insufficient to attend to the stringent requirement of the IoT. Some of the fundamental challenges of the IoT are presented in this section.

2.1.1 Latency requirements

One of the main challenges of the IoT is the associated stringent latency requirements. For instance, a lot of industrial control systems usually require end-to-end latencies of a small number of milliseconds between the control node and the sensor [17, 18]. Examples of such applications are oil and gas systems, manufacturing systems, goods packaging systems, and smart grids. On the other hand, end-to-end latencies below a few tens of milliseconds are required by some time-sensitive (high-reliability and low-latency) IoT applications like drone flight control applications, vehicle-to-roadside communications, gaming applications, virtual reality applications, and vehicle-to-vehicle communications, and other real-time applications. However, these requirements are beyond what a conventional cloud can effectively support [13].

2.1.2 Bandwidth constraints

The unprecedented increase in the number of connected IoT devices results in the generation of huge data traffic. The created traffic can range from tens of megabytes to a gigabyte of data per second. For instance, about one petabyte is been trafficked by Google per month while AT&T's network consumes about 200 petabytes in 2010. Besides, it is estimated that the U.S. smart grid will generate about 1000 petabytes per year. Consequently, for effective support of this traffic, relatively huge network bandwidth is demanded. Moreover, there are some data privacy concerns and regulations that prohibit excessive data transmission. For example, according to ABI Research, about 90% of the generated data by the

endpoints should not be processed in the cloud. In this context, it has to be stored and processed locally [13].

2.1.3 Resource-constrained devices

The IoT system comprises billions of objects and devices that have limited resources mainly regarding storage (memory), power, and computing capacity [15]. Based on these limitations, it is challenging for constrained devices to simultaneously execute the entire desired functionality [19]. Besides, it will be impractical to depend exclusively on their relatively limited resources to accomplish their entire computing demands. It will also be cost-prohibitive and unrealistic for the devices to interact directly with the cloud, owing to the associated complex protocols and resource-intensive processing [13]. For example, some constrained medical devices such as insulin pumps and blood glucose meters have to fulfill certain authentication and authorization tasks. Likewise, it has been observed that most of the resource-constrained IoT devices cannot partake in the blockchain consensus mechanisms such as Proof-of-Work (PoW) and Proof-of-Stake (PoS) protocols in which huge processing power is required for the mining process [15].

2.1.4 Intermittent connectivity

It will be challenging for the centralized-based cloud platforms to offer uninterrupted cloud services to systems and devices such as oil rigs, drones, and vehicles with intermittent network connectivity to the cloud resources. As a result, an intermediate layer of devices is required to address the challenges [2, 15].

2.1.5 Information and operational technologies convergence

The advent of Industry 4.0 facilitates the convergence of Information Technology and Operational Technology. In this context, new operational requirements and business priorities are presented. It is noteworthy that safe and incessant operation is of utmost importance in current cyber-physical systems. This is owing to the fact that service disruption can result in a significant loss or dissatisfaction. Consequently, software and hardware update in such sensitive systems is challenging. This calls for a novel architecture that is capable of reducing the system updates [2].

2.1.6 Context awareness

A lot of IoT applications like augmented reality and vehicular networks require access to be able to process local context information such as network conditions and user location. However, owing to the physical distance between central computing and IoT devices, the centralized-based cloud computing implementation is insufficient to support the requirement [2].

2.1.7 Geographical location

The IoT devices are huge in number and are widely distributed over broad geographical areas. These devices require computation and storage services for effectiveness. However, it is challenging to have a cloud infrastructure that can support the entire requirements of the IoT applications [2].

2.1.8 Security and privacy

The present Internet cybersecurity schemes are mainly designed for securing consumer electronics, data centers, and enterprise networks. The solutions target perimeter-based protection provisioning using firewalls, Intrusion Detection

	Computing		**Reference**
	Cloud	**Fog**	
Deployment	Centralized	Distributed[1]	[13, 21–23]
Planning	Demands complicated deployment planning	Demands cautious deployment planning[2]	[13]
Operation	It is controlled and maintained by the expert cloud personnel and operated in designated environments.	The environments are usually determined by customer demands and may require little or no human expert intervention.	[13, 22]
	Usually owned by large companies.	Depending on the size, it could be owned by large or small companies.	
Supported application	Mainly cyber-domain systems.	Cyber-domain and cyber-physical systems.	[13, 22]
	Few seconds round-trip delay-tolerant applications.	Time-critical applications that demand less than tens of milliseconds.	
Connectivity	Work effectively with consistent connectivity.	Can work with intermittent connectivity.	[13]
Latency	High	Low	[21–23]
Storage and computation capabilities	Strong	Weak	[22]
Energy consumption	High	Low	[22]
Bandwidth requirement	High[3]	Low[4]	[13, 22]
Location	Core network	Edge network	[13, 21]
Location awareness	Partially supported	Supported	[22]
Security aspect	Less secure	More secure	[24]
Attack on moving data	High probability	Very low probability	[24]
Client - server distance	Multiple hops	One hop	[25]
Mobility support	Limited	High	[25]

[1]*A distributed or centralized control system can be employed for distributed fog nodes.*
[2]*Some fog deployment is ad-hoc and demands either minimal or no planning.*
[3]*Bandwidth requirement increases with the aggregate volume of generated data by the entire clients.*
[4]*Bandwidth requirement increases with the aggregate volume of filtered data to be sent to the cloud.*

Table 1.
Comparison of main features of cloud and fog computing.

Systems (IDSs), and Intrusion Prevention Systems (IPSs). Besides, based on the associated advantages, certain resource-intensive security functions have been shifted to the cloud. In this regard, they are focusing on perimeter-based protection by requesting authentication and authorization through the clouds. However, the security paradigm is insufficient for IoT-based security challenges.

2.2 Fog computing

To address the centralized-based limitations and for effective support of the IoT devices, edge computing has been presented [2, 14]. Besides, a broader architecture known as fog computing that is based on a distributed scheme has been presented. In the fog paradigm, storage, control, communication, computation, and networking functions are distributed in close proximity to the end-user devices along the cloud-to-things continuum [13].

In addition, to complement the centralized-based cloud platforms in which data, computing functions, and control functions are stored and performed in the cellular core networks and remote data centers, fog stores a significant amount of data and performs considerable functions at or near the end-user. Likewise, instead of routing the entire network traffic over the backbone networks, a considerable amount of networking and communication are performed at or in close proximity to the end-user in fog computing [2, 8, 15, 20]. In this regard, when applications/tasks are offloaded to the neighboring fog nodes rather than a cloud center, fast-response and low-latency services can be offered by fog computing. Besides, the required enormous backhaul burden between the fog nodes and the remote cloud center is alleviated [8, 20].

Cloud and fog are complementing computing schemes. They establish a service continuum between the endpoints and the cloud. In this regard, they offer services that are jointly advantageous and symbiotic to ensure effective and ubiquitous control, communication, computing, and storage, along the established continuum [13]. In **Table 1**, we present the major features of the cloud and the fog to illustrate the advantage of their complements for effective and ubiquitous service delivery along the continuum.

3. Fog architectures and features

Fog computing can enhance the QoS and the efficiency of different use cases. In this context, it can offer noble technical support for cyber-physical system, Mobile Internet, and IoT. This section presents the fog architectural model and the related advantages of fog computing.

3.1 Three-layer architecture of fog

As aforementioned, one of the main features that differentiate fog from cloud computing is that in the former, resources regarding the storage, communication, control, and computation are deployed in proximity to the end-user devices. Moreover, fog architecture can be predominantly centralized, fully distributed, or somewhere amid the two former configurations. Furthermore, fog architecture and supported applications can be implemented in dedicated hardware and software. In addition, fog architecture can also be virtualized to exploit the associated advantages of network virtualization. This will facilitate the execution of the same application wherever it is demanded. In this context, the demand for dedicated applications will be reduced. It can also encourage an open platform in which

applications from different vendors can share a common network infrastructure with support for common lifecycle management. Based on this, different applications can be removed, added, updated, deactivated, activated, and configured, to ensure seamless end-to-end services across the continuum [13].

The fog computing architectural model is usually represented by a three-layer architecture that consists of the cloud, fog, and IoT layers [2, 22, 26]. Besides, a broader N-layer reference architecture has been defined by the OpenFog Consortium [2]. This architecture is an improvement on the three-layer architecture. This subsection focuses on three-layer fog architecture and related concepts.

Figure 1 illustrates hierarchical architecture of fog computing with three-layer. This architecture presents a significant extension to cloud computing. In this regard, to bridge the gap between the cloud infrastructure and the end/IoT devices, it offers a transitional layer that is known as *Fog layer*. We expatiate on the associated layers of the architecture in this subsection.

3.1.1 Terminal/IoT layer

The terminal/IoT layer is the layer that is close to the physical environment and end-user. It comprises numerous devices such as mobile phones, tablets, smart vehicles, smartphones, smart cards, drones, sensors, etc. Typically, these IoT devices are usually distributed geographically. Also, their major purpose is to sense feature data of physical events or objects for onward transfer to the upper layer for processing and/or storage. It is noteworthy that certain local processing can also be executed by a number of devices such as smart vehicles, smartphones, and mobile phones that have substantial computational capabilities. After local processing, the resulting data can then be forwarded to the upper layers [2, 22].

3.1.2 Fog layer

The fog layer is normally positioned on the network edge and is the fundamental layer of fog computing hierarchical architecture. The layer comprises a huge number of fog nodes such as fog servers, base stations, switches, routers, access points, and gateways, which are broadly distributed between the cloud and end-user devices [2, 22]. It should be noted that the fog nodes are not only physical network elements but are also logical ones that execute fog computing services [2].

Moreover, the fog nodes can be based on mobile implementation, when deployed on a nomadic carrier, or static, when fixed at a location. With these implementations, end-user devices can suitably connect with appropriate nodes to have access to the required services. Besides, the nodes are connected to the cloud

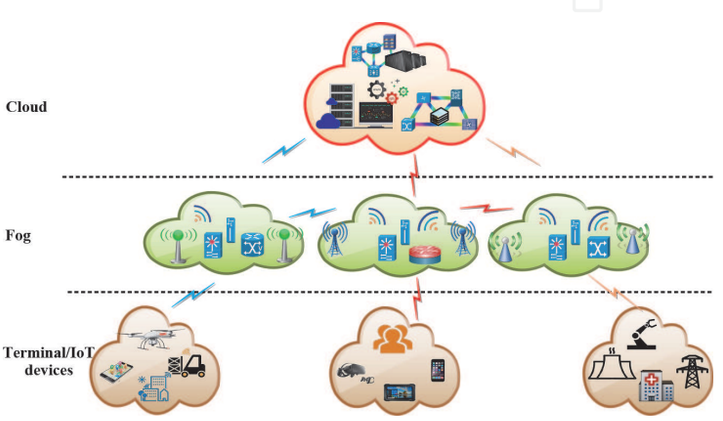


Figure 1.
Fog computing hierarchical network model: a three-tier architecture.

infrastructure through an IP core network for service provision [22]. As aforementioned, fog nodes are capable of computing and transmitting sensed data. Likewise, the received sensed data can be temporarily stored by the nodes. Due to their access to specific network resources; latency-sensitive applications, as well as real-time analysis, can be realized in the fog layer. Furthermore, certain applications demand more powerful storage and computing capabilities. In this context, fog nodes have to interact with the cloud to obtain the required network resources [2, 22].

3.1.3 Cloud layer

The main components of the cloud layer are high-performance servers (with high memory and powerful computational capabilities) and storage devices. Consequently, a huge amount of data can be processed and permanently stored in this layer. Based on this, the layer normally supports various intensive services such as smart factories, smart transportations, and smart homes [2, 22].

Furthermore, unlike the traditional cloud computing architecture, not the entire storage and computing tasks traverse the cloud. Therefore, in the fog architecture, certain services and/or resources can be moved (offloaded) from the cloud to the fog layer through a number of control schemes to optimize resource utilization and increase efficiency [2, 22].

3.2 Features of fog computing

Fog offers various advantages that facilitate new business models and services that can help in expenditure reduction or product rollouts acceleration. In the following subsection, we discuss some of the main advantages of fog network.

3.2.1 Client-centric objective cognizance

Fog architecture consists of widely distributed nodes that support mainly short-range communication and are capable of tracking and getting the end device locations to enable mobility. This feature can facilitate enhanced location-based services and improved potentials for real-time decision making [22]. For example, as fog applications are close to the end-user devices, they can be designed for efficient awareness of the customer requirements. Cognizance of customer requirements helps the fog architecture in establishing the appropriate place to perform storage, computing, and control functions across the cloud-to-thing continuum [13].

3.2.2 Resource pooling and bandwidth efficiency

In fog computing, there can be ubiquitous distributions of resources between the endpoints and the cloud. This helps in the efficient exploitation of the available resources. Besides, with the fog architecture, various applications can leverage the available resources that are abundant but idle on the end-user devices and network edge [13]. For instance, certain computation tasks such as data filtering, data cleaning, data preprocessing, valuable information extraction, redundancy removing, and decision making, are locally performed. In this context, a certain portion of useful data is conveyed to the cloud. Consequently, there is no need to transmit the majority of the data over the Internet. Based on this, fog computing is capable of reducing the network traffic, consequently, the bandwidth is effectively saved [22]. Similarly, the proximity of the fog system to the endpoints facilitate effective integration with the end-user systems. This helps in enhancing the performance and efficiency of the entire system [13].

3.2.3 Scalable and cost-effective architecture

As aforementioned, fog architecture is relatively simple and encourages prompt innovation. Besides, it offers a platform that supports economical scaling. For instance, it is much more cost-effective and faster to use the edge (or client) devices for innovation experimentations instead of that of large operators and vendors networks. In this context, fog encourages an open-market (interoperable) that can support open-application programming interfaces. This is of utmost importance for the proliferation of mobile devices to facilitate innovation, development, deployment, and operation of advanced services [13].

3.2.4 Low-latency and real-time applications

In a fog network, data analytics are allowed at the network edge [13]. The generated data by devices and sensors is acquired locally by the fog nodes at the network edge. The acquired (high priority) data is then processed and stored by edge devices in the local area network. Based on this, traffic across the Internet can be considerably reduced and swift localized services with high-quality can be supported. Hence, time−/latency-sensitive applications for real-time interactions can be supported [22]. For instance, time-sensitive functions can be well-supported for local cyber-physical systems. Besides, this feature is crucial for stability in control systems. Likewise, to support embedded AI applications, the feature is also important for the tactile Internet vision [13]. On the other hand, low priority data that is delay-insensitive can be conveyed to certain aggregation nodes where it will be further processed and analyzed [26].

Furthermore, fog computing focuses on allowing ubiquitous local access to centralized computing resource pools that can be swiftly and flexible provisioned on-demand basis. So, to alleviate communication latency and support delay/jitter sensitive applications, resource-limited end-user devices that are close to the fog nodes can access resource pools. In general, the key native features of fog computing are context awareness and edge location. Besides, it is based on pervasive spatial

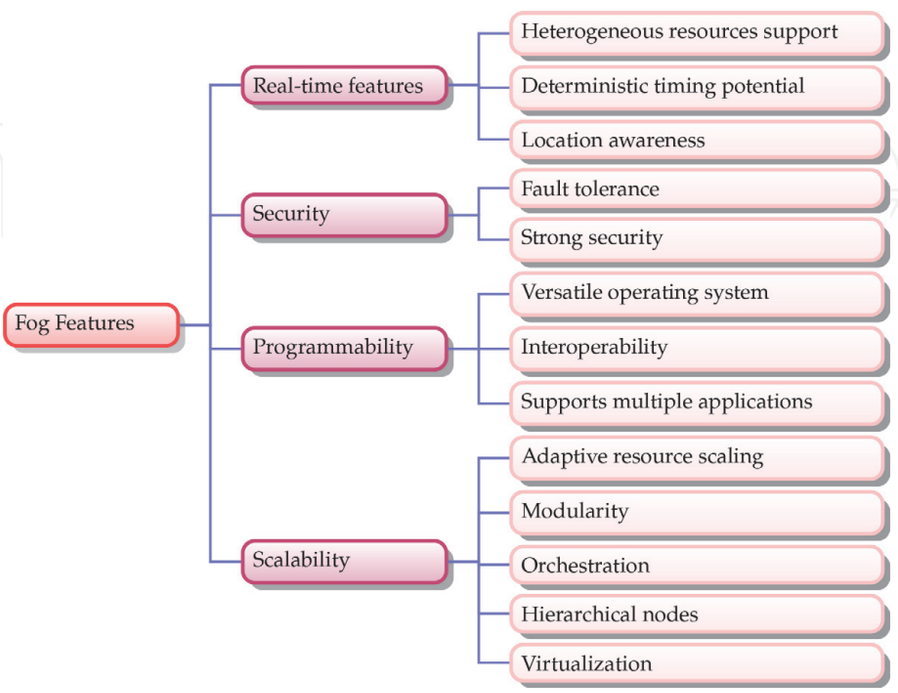


Figure 2.
Fog paradigm features.

deployment to support various devices [27]. In **Figure 2**, we present some of the major features of the Fog paradigm. Also, the following section focuses on the resource allocation challenges in fog computing.

4. Challenges of fog computing and its integration with other computing platforms

As aforementioned, fog computing offers a complementary platform that allows users to offload computational tasks to the network edge. So, resource-limited fog nodes with processing and storage functionalities are deployed at the edge of the network to enhance network performance. Nevertheless, the huge users' data from various applications will be wirelessly transmitted to the cloud center or fog nodes, as the case may be. Based on this, massive communication bandwidth is demanded. It is noteworthy that this is an expensive and constrained resource in the communication systems [8, 28, 29].

Moreover, there are a number of associated challenges of fog-cloud-based computing integrated architecture. One of the main associated challenges is on efficient management of fog infrastructure and allocation of accessible resources to the IoT devices. It is noteworthy that a huge amount of services can be demanded simultaneously by the IoT devices while on the other hand, the respective fog service node is bestowed with limited storage and computing capabilities. Based on this, the entire fog nodes have to be managed optimally. In this context, for the efficient provision of the requested services, they have to be optimally allocated to the required IoT devices. Besides, fog computing resource management is another notable challenge that calls for effective control among fog nodes [26].

Furthermore, it is highly imperative to consider various factors such as energy consumption, service availability, and associated expenses, when fog nodes are deployed to deliver services [30]. In this context, to meet IoT application requirements, optimal mapping of the fog service nodes to the IoT devices is a challenging task. Besides, privacy and security issues such as access control, trust management, intrusion detection, access authentication, etc. [31] in integrated fog computing and IoT device setups are challenging [26].

As aforementioned, one of the main challenges in an integrated computing platform is the multiple resource allocation issue. In this context, efficient resource management across the fog-cloud platform for an effective computation offloading is of paramount importance. To address this challenge, an effective multiple access technique such as NOMA is required [8].

NOMA is an attractive radio multiple access techniques that mainly targets next-generation wireless communications. In NOMA, the power domain for multiple access is leveraged. It presents a number of potential advantages such as high reliability, reduced transmission latency, enhanced spectrum efficiency, and massive connectivity [32]. The main concept of NOMA is to serve multiple users through the same resource regarding time, code/space, and frequency, by exploiting different power levels. Afterward, at the fog nodes, a cancelation technique like successive interference cancelation (SIC) can be implemented to separate and decode the superimposed signals [8, 33]. In the following section, we present some models for resource allocation in an integrated fog-cloud architecture with NOMA implementation.

5. System model

This section presents some related models for the fog computing hierarchical network model illustrated in **Figure 1**. Assume that M fog nodes with various

computing and storage capabilities are deployed to offer offloading services given by a set of $\mathcal{M} = \{1, 2, \ldots, M\}$. Besides, assume N users denoted by $\mathcal{N} = \{1, 2, \ldots, N\}$ with $\mathcal{J} = \{1, 2, \ldots, J\}$ independent computation tasks to be executed. The respective task can be expressed as [8]

$$F_{nj} = \left\{A_{in}(nj), Q_{req}(nj), T_{max}(nj), n \in \mathcal{N}, j \in \mathcal{J}\right\}, \tag{1}$$

where $A_{in}(nj)$ is the size of computation input data of the j-th task demanded by the n-th user, $T_{max}(nj)$ represents the maximum tolerable latency of the j-th task required by the n-th user, and $Q_{req}(nj)$ is the total number of central processing unit (CPU) cycles needed to execute the task.

In addition, to express different related models, we assume a quasi-static scenario in which the users are unchanged in the course of computation offloading, however, they can change over different periods. Besides, we assume a perfect instantaneous channel that remains unchanged during the packet transmission. Based on this, we present the following models for an integrated fog-cloud architecture with NOMA implementation.

5.1 Communication model

When an n-th user with a number of offload tasks and the transmission power, p_{mn}, transmits signal, x_{mn}, to the m-th fog node, the received signals, y_{mn}, can be expressed as

$$y_{mn} = \underbrace{\sqrt{p_{mn}} h_{mn} x_{mn}}_{\text{Desired signal}} + \underbrace{\sum_{i \neq n, i \in \mathcal{N}} \sqrt{p_{mi}} h_{mi} x_{mi}}_{\text{Intra-cell interference}} + \underbrace{z_{mn}}_{\text{Noise}}, \tag{2}$$

where the first term represents the desired signal from the n-th user, the second term is the intra-cell interference suffered by the n-th user from other users being served by the m-th fog node on the same frequency band, the third term, z_{mn}, denotes the additive white Gaussian noise (AWGN) with zero mean and variance δ^2, and h_{mn} denotes the channel gain for the n-th user that connects to the m-th fog node.

It is noteworthy that the transmitted signals from various users to each fog node are the desired signals. However, they bring about interference with each other. Also, as individual users that are connected to a specified fog node suffer different channel conditions, the interference can be alleviated and the superimposed signals can be decoded sequentially by each fog node using SIC [8, 33].

In the linear interference cancelation techniques, the desired signal is detected, but other signals are regarded as interference. So, the SIC concept is based on the fact that the signal that has the highest signal-to-interference-plus-noise-ratio (SINR) can be detected first. In this regard, its interference is canceled from other streams [34]. Furthermore, regarding the integrated computing platform, the received signal by a specified fog node from the user that has the highest channel gain is the potential strongest signal, so it is decoded first at the fog node. Afterward, the strongest signal will be removed from the streams. The same approach is then applied to the user with the second-highest channel gain and so on. Consequently, the users' signals on the same frequency band can be sorted in relation to the channel gains. In this context, the users served by the m-th fog node can be arranged in descending as [8]

$$|h_{m1}|^2 \geq |h_{m2}|^2 \geq \cdots \geq |h_{mN}|^2 \ \forall n \in \mathcal{N} \tag{3}$$

Using Eq. (3), every single fog node can subtract and decode the desired signals. Besides, the received SINR, γ, of the n-th user being served through the m-th fog node can be defined as [8, 35, 36]

$$\gamma_{mn}(p_{mn}) = \frac{p_{mn}|h_{mn}|^2}{\delta^2 + \sum_{i=n+1}^{N} p_{mi}|h_{mi}|^2} \tag{4}$$

Furthermore, the resultant transmitted data rate of the n-th user at the m-th fog node can be expressed as [8, 35, 37]

$$R_{mn}(w_{mn}, p_{mn}) = w_{mn}\log_2(1 + \gamma_{mn}(p_{mn})), \tag{5}$$

where w_{mn} represents the occupied frequency band of the n-th user that is served by the m-th fog node and $\mathcal{W}$ denotes the total frequency band.

Moreover, as a result of the limited resources in the fog node, it is challenging to concurrently fulfill the entire services demanded by the end users. So, to acquire the demanded services, each end-user should have a satisfaction function for the evaluation of the allocated resources, ξ. The associated satisfaction function, χ, can be defined as [26]

$$\chi(\xi) = \begin{cases} \log(\xi + 1), & 0 \leqslant \xi < \xi_{\min} \\ \log(\xi_{\max} + 1), & \xi \geqslant \xi_{\min} \end{cases}, \tag{6}$$

where $\xi_{\max}$ denotes the maximum resource that is required to offer the demanded service.

Moreover, based on the satisfaction function, the major objective of the fog node is to offer a global satisfaction maximization for the entire end users. This can be expressed as [26]

$$\textbf{Objective}.\ \max\left\{\chi_g\right\} \tag{7}$$

$$\textbf{S.t.}$$

$$\begin{cases} \chi_g = \sum_{i=1}^{n}\{\tau_i \cdot \chi_i(\xi_i)\} \\ \xi_1 + \xi_2 + \cdots + \xi_n \leqslant \Xi \\ \tau_1 + \tau_2 + \cdots + \tau_n = 1 \\ \xi_1, \xi_2, \ldots, \xi_n \geqslant 0 \end{cases}, \tag{8}$$

where χ_g denotes the overall satisfaction of the entire end users, ξ_i denotes the allocated resource to the n-th end-user, Ξ represents the possessed resource by the fog node, and τ represents the associated priority level for the n-th end-user.

Furthermore, using Eqs. (7) and (8), resources of the fog node can be allocated to the entire end-device while the overall maximum satisfaction is achieved. Moreover, the fog nodes are connected and are capable of sharing their resources to deliver the requested service by the end-users. Assume a scenario in which a fog node does not possess sufficient resources to offer services that are locally requested, then it can shift certain requested services with low priority level to the neighboring fog nodes with spare resources for processing. The spare resources, R_s^f, of the m-th fog node can be defined as

$$\Xi_{\text{spare}}^f = \Xi^f - \sum_{i=1}^{n} \xi_i^{\max}, \tag{9}$$

where, $\xi_i^{\max}$ denotes the maximum resource required by the n-th end-user and Ξ^f represents the resource m-th fog node.

5.2 Fog computing model

The fog computing model is based on the required tasks and the associated overhead. For instance, each fog node will receive task offloading requests from the users. Based on its resource capabilities, the respective node is expected to process the requested computational tasks. As a result of this, certain overhead regarding time and energy will be incurred to transmit and process at the fog nodes [8, 37]. The associated overheads are discussed in the following subsections.

5.2.1 Task processing latency

Based on the communication model presented in subsection 5.1, the transmission latency regarding computation offloading can be evaluated. Assume that the m-th fog node receives computation task F_{nj} from the n-th user, the incurred transmission latency when the n-th user send data to offload the j-th task using Eq. (5) can be expressed as

$$T^f_{mnj,t} = \frac{A_{in}(nj)}{R_{mn}} \tag{10}$$

As aforementioned, each fog node possesses limited computation capabilities. Assume that the m-th fog node with computing capability, C^f_{mn}, is assigned to the n-th user, the related computation execution time $T^f_{mnj,e}$ can be defined as

$$T^f_{mnj,e} = \frac{Q_{req}(nj)}{C^f_{mn}} \tag{11}$$

Furthermore, consider a scenario in which each fog node is equipped with a CPU that is based on non-preemptive allocation. Also, assume that computing resource is assigned to an individual user each time until its required tasks are accomplished. Moreover, assume the process sequence, $\mathbf{q}_m = \{q_{ms} | q_{ms} \in \{1, 2, \ldots, N\}, q_{ms} \neq q_{mn}, s, n \in \mathcal{N}\}$ in the m-th fog node in which the tasks are executed in the ascending order, q_m. In this scenario, for task j, the queuing delay time can be defined as

$$T^f_{mnj,q} = \sum_{s, q_{ms} < q_{mn}}^{N} b_{ms} T^f_{msj,e} \tag{12}$$

where b_{ms} represents the outcome of user scheduling for the fog nodes that specifies selection of the m-th fog node by the s-th user for offloading. The selection criteria of the m-th fog node by the s-th user can be defined as [8, 37]

$$b_{ms} = \begin{cases} 1 & \text{if } s\text{-th user selects (associated with) the } m\text{-th fog node} \\ 0 & \text{otherwise} \end{cases} \tag{13}$$

The aggregate latency incurred by the fog computing when the n-th user offloads the j-th task to the m-th fog node can be defined as [8, 36, 37]

$$T^f_{mnj} = T^f_{mnj,t} + T^f_{mnj,e} + T^f_{mnj,q} \tag{14}$$

5.2.2 Energy consumption

The energy consumptions for transmitting and processing tasks are the main offloading energy consumptions [30]. When an n-th user offloads j-th task to an m-th fog node, the associated energy consumption can be defined as

$$E^f_{mnj} = \underbrace{\overbrace{E^f_{mnj,t}}^{\text{Transmission energy consumption}} + \overbrace{E^f_{mnj,e}}^{\text{Computing energy consumption}}}_{\text{Off loading energy consumption}} = T^f_{mnj,t} p_{mn} + \eta_m C^f_{mn} T^f_{mnj,e}, \tag{15}$$

where η_m represents the coefficient that signifies the energy consumption per CPU cycle of an m-th fog node. The first and the second terms are the transmission and computing energy consumptions on the m-th fog node.

5.3 Cloud computing model

As aforementioned, the fog nodes have relatively limited resources regarding memory (storage), power, and computing capacity [15]. Therefore, when resource-limited fog nodes are not capable of accomplishing the requested computational tasks due to their constrained resources, the tasks have to be sent to the cloud center via the backhaul links [8].

Furthermore, the tasks can be efficiently executed at the cloud centers owing to their sufficiently high resource capabilities. It should be noted that additional overhead regarding energy and time will be incurred in the process of forwarding tasks to the remote cloud center [8]. The related overheads are considered in the following subsections.

5.3.1 Task processing latency

Suppose the backhaul links data rate, R^b_m, is available between the m-th fog node and the remote cloud center with computation capability given by C_c. Besides, based on sufficient resources and powerful computation capabilities of the cloud center, the tasks from various users can be executed instantly. In this context, the queuing delay can be omitted in the processing latency analysis of cloud computing. Following the fog computing model analysis presented in subsection 5.2, the aggregate latency that will be incurred in forwarding the j-th task of the n-th user from the m-th fog node to the remote cloud center can be expressed as [8, 35]

$$\begin{aligned} T^c_{mnj} &= T^f_{mnj,t} + T^c_{mnj,t} + T^c_{mnj,e} \\ &= \frac{A_{in}(nj)}{R_{mn}} + \frac{A_{in}(nj)}{R^b_m} + \frac{Q_{req}(nj)}{C_c} \end{aligned} \tag{16}$$

5.3.2 Energy consumption

The aggregate energy consumption of an m-th fog node that offloads a j-th task of an n-th user to a remote cloud center can be defined as [8, 36]

$$\begin{aligned} E^{c}_{mnj} &= E^{f}_{mnj,t} + E^{c}_{mnj,t} + E^{c}_{mnj,e} \\ &= T^{f}_{mnj,t} p_{mn} + T^{c}_{mnj,t} p^{b}_{m} + \eta_c T^{c}_{mnj,e} C_c \end{aligned} \tag{17}$$

where η_c represents a coefficient that signifies the energy consumed by the CPU of the cloud per cycle, p^{b}_{m} denotes the allocated power for tasks forwarding to the cloud center by an m-th fog node.

6. Trends toward intelligent-based fog computing

This section focuses on the current trends toward intelligent integrated computing networks. As aforementioned, different challenges as regards scalability, management, and optimization have been presented in the fog-cloud-based computing integrated architecture. For efficient management of resource-limited fog nodes and optimization of the cloud computing platform, the trend is toward the adoption of AI-enabled techniques in the *Network 2030* (6G and Beyond) [4, 10]. For instance, apart from intelligent driving that the 6G network is anticipated to support, it will also offer a promising path toward the industrial revolutions where the future intelligent factories are anticipated to support densely concentrated intelligent mobile robots. Based on this, a number of new service classes like ultra-high data density (uHDD), ubiquitous mobile ultra-broadband (uMUB), and ultrahigh-speed-with-low-latency communications (uHSLLC) have been defined [38, 39]. A typical instance of an integrated hierarchical computing platform supported by the AI techniques is illustrated in **Figure 3**.

Furthermore, in the considered integrated hierarchical computing AI-enabled platform, we consider the fog radio access networks (F-RANs) as a case study. In this context, we discuss how F-RANs can facilitate the deployment of hierarchical AI in wireless networks. Besides, consideration is given to the influences of AI in making F-RANs smarter in rendering better services to mobile devices.

In addition, regarding the influences of F-RANs on the deployment of AI (F-RAN-Enabled AI), the F-RANs present hierarchical layers (cloud, fog, and IoT) that can be exploited. So, F-RAN offers heterogeneous processing capability that can be leveraged for hierarchical intelligence across the integrated layers through centralized, distributed, and federated learning. Besides, to significantly alleviate the memory issue of mobile devices, cross-layer learning can also be employed. Besides, concerning the influences of AI on the F-RANs (AI-Enabled F-RAN), AI presents F-RANs with techniques and technologies for effective support of the huge traffic. Likewise, it helps in making intelligent decisions in the networks. These features can be harnessed through the implementation of ML tools such as reinforcement learning (RL) algorithms and deep neural networks (DNNs) [28, 29, 40, 41]. For instance, DNNs can be adopted for data processing. Besides, RL algorithms can be employed for optimizations and decisions [40]. We expatiate on the relationship between the F-RANs and AI in the following subsections.

6.1 F-RAN-enabled AI

The F-RAN heterogeneous platforms with varied memory and computational resources offer hierarchical application scenarios for the AI. Based on this, hierarchical intelligence such as cloud intelligence, fog intelligence, and on-device intelligence, can be achieved across the layers [40]. In this part, we present learning-based intelligence schemes for the F-RAN-Enabled AI.

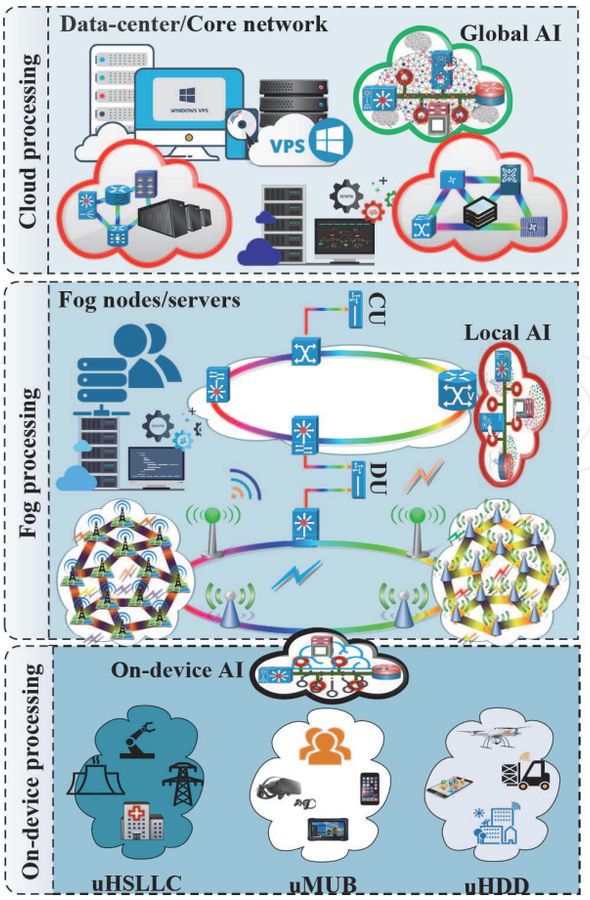


Figure 3.
A typical AI/fog-enabled computing architecture. uHSLLC: Ultrahigh-speed-with-low-latency communications; uMUB: ubiquitous mobile ultra-broadband; and uHDD: ultra-high data density.

6.1.1 Centralized learning-based cloud intelligence

The centralized cloud is not only endowed with considerable access to a global dataset but also has a significant amount of storage and computing power. Consequently, with sufficient data samples, training of the centralized DNN algorithms can be used to leverage the powerful cloud intelligence. Owing to its flexibility and pay-as-you-go capability, cloud intelligence-based services can be on-demand. In this regard, it can scale in accordance with the subscribers' requirements [42].

Moreover, in centralized-based learning, it is assumed that the mobile devices transmit data to the central cloud. However, this is at the expense of communication overhead, regarding bandwidth and energy. Usually, it is challenging to meet the real-time application demands because of the incurred latency. Besides, due to privacy concerns of the mobile devices, attention should also be paid to the transmission of the generated data. Therefore, these concerns demand alternative solutions. A viable approach is based on the exploitation of the distributed architecture and processing capabilities of the mobile devices and/or fog nodes in the development of distributed ML techniques [40, 42].t

6.1.2 Distributed learning-based fog intelligence

In edge computing, cloud resources are leveraged by the fog nodes. Also, the service latency is significantly reduced because the fog nodes are in proximity to the devices. Also, the proximity helps in enhancing privacy. Moreover, based on the distributed natures of the fog nodes and mobile devices, the edge ML algorithms are

normally implemented distributedly. In this regard, the training samples are distributed randomly over a considerable number of mobile devices and fog nodes. Each fog node executes training tasks using the gathered local data samples from the associated mobile devices. In addition, when the local model state information (MSI) is aggregated, a global model can be acquired from the fog nodes. Nevertheless, devices in this scheme (distributed ML) have to send their respective data to the fog nodes. This procedure can also violate data privacy. In view of this, federated learning can be employed to further enhance privacy [11, 40, 43].

6.1.3 Federated learning-based on-device intelligence

There have been unprecedented improvements in the processing capability of mobile devices, making joint training and inference more viable. The learning in this layer can be achieved using federated learning architecture [44]. This approach is contingent on periodic computation and exchange of updated MSI versions of the individual mobile devices. Consequently, rather than sending the raw data, the MSI computed using their datasets are exchanged. At the fog node, the distributed MSI updates of the associated mobile devices are aggregated. Based on this, the results of low-latency inference can be acquired. These results can be used in delay-sensitive applications to make a fast response to local events. Apart from being a swifter solution, this approach also helps in enhancing data privacy and is a promising solution for privacy-sensitive applications. In addition, it is also possible to aggregate the distributed MSI updates of the fog nodes to achieve a global model in the cloud. It is noteworthy that this learning process is appropriate for training a low-weight AI model in which there are fewer parameters on mobile devices. Nevertheless, for a remarkably large number of mobile devices and an AI model with huge parameters, wireless data aggregation of MSI updates is challenging [11, 40, 43].

6.1.4 Cross-layer learning-based hierarchal intelligence

In a scenario where the AI model size is more than the memory size of the mobile devices, the mobile devices will be unable to complete the entire model training on themselves. In such a case, the model should be partitioned into sections and distributed over the network entities. So, the lower layers' (mobile devices) outputs will be aggregated prior to transmission to the fog nodes (intermediate layers). Likewise, the intermediate layers' output will be aggregated ahead of transmission to the top cloud layers. One of the advantages of cross-layer learning is that it aids system scalability. Besides, the demand for mobile devices' memory size can also be reduced. Nevertheless, cross-layer DNNs demand stringent training algorithms [40].

6.2 AI-enabled F-RAN

The significant growth in the bandwidth demands by the radio access networks is mainly due to the proliferation of mobile devices and various supported bandwidth-intensive multimedia applications. This results in a traffic explosion that is challenging for the current mobile networks. To address the issue, various network architectures that exploit different types of resources have been presented. However, resource management in such emerging architectures is very demanding. Based on this, innovative techniques for excellent data processing and efficient network optimization are required [20, 40]. The following subsections present AI techniques as viable tools for attending to the associated network challenges.

6.2.1 Intelligent data processing

Based on the growing increase in the diversity of F-RAN applications, the envisaged multimedia data to be supported will be heterogeneous, huge, and high-dimensional. Therefore, direct raw data transmission to the cloud and fog node will bring about high communication overhead. Besides, direct utilization of raw data for network optimization can cause high-computing overhead and low-efficiency issues. Moreover, there has been considerable advancement in the DNNs that facilitate data processing. For instance, convolutional operations have been exploited by convolutional neural networks (CNNs) for spatial feature extraction from input signals [40].

6.2.2 Intelligent network optimization

There are a number of ML techniques such as unsupervised, supervised, and RL algorithms that can be employed for efficient network optimization. For instance, as supervised learning focuses on mapping inputs to outputs in accordance with the training samples, the DNN-based supervised learning is an attractive scheme for beamforming design and power control of fog nodes. On the other hand, unsupervised learning is based on inferring the underlying data structure without any label, so it is appropriate for empirical analysis such as computation offloading, clustering, and resource allocation, in the F-RAN. Besides, in the RL, to maximize predicted cumulative return, sequential actions are taken by actor/agent based on the environment observations [42, 45].

7. Conclusion

In this chapter, we have presented a comprehensive overview of the evolution of computing paradigms and have highlighted their associated features. Moreover, different models that focus on effective resource allocation across an integrated computing platform have been presented. Besides, a comprehensive discussion on efficient resource management and optimization of the 6G fog computing platform to meet strict on-device constraints, reliability, end-to-end latency, bit-rate, and security requirements have been presented. In this context, we have presented AI as a resourceful technique for the achievement of high-level automation in the integrated computing heterogeneous platform.

Acknowledgements

This work is supported by the European Regional Development Fund (FEDER), and Internationalization Operational Programme (COMPETE 2020) of the Portugal 2020 (P2020) framework, under the projects DSPMetroNet (POCI-01-0145-FEDER-029405) and UIDB/50008/2020-UIDP/50008/2020 (DigCORE). It is also supported by the Project 5G (POCI-01-0247-FEDER-024539), SOCA (CENTRO-01-0145-FEDER-000010), ORCIP (CENTRO-01-0145-FEDER-022141), and RETIOT (POCI-01-0145-FEDER-016432).

DOI: http://dx.doi.org/10.5772/intechopen.98315

IntechOpen

Author details

Isiaka A. Alimi[1*], Romil K. Patel[1,2], Aziza Zaouga[1], Nelson J. Muga[1], Qin Xin[3], Armando N. Pinto[1,2] and Paulo P. Monteiro[1,2]

1 Instituto de Telecomunicações and University of Aveiro, Portugal

2 Department of Electronics, Telecommunications and Informatics, University of Aveiro, Portugal

3 Department of Science and Technology, University of the Faroe Islands, Tórshavn, Faroe Islands

*Address all correspondence to: iaalimi@ua.pt

IntechOpen

© 2021 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/3.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. CC BY

References

[1] Michael Armbrust, Armando Fox, Rean Griffith, Anthony D. Joseph, Randy Katz, Andy Konwinski, Gunho Lee, David Patterson, Ariel Rabkin, Ion Stoica, and Matei Zaharia. A View of Cloud Computing. *Commun. ACM*, 53 (4):50–58, April 2010.

[2] M. De Donno, K. Tange, and N. Dragoni. Foundations and Evolution of Modern Computing Paradigms: Cloud, IoT, Edge, and Fog. *IEEE Access*, 7: 150936–150948, 2019.

[3] P. M. Shakeel, S. Baskar, H. Fouad, Gunasekaran Manogaran, V. Saravanan, and Q. Xin. Creating Collision-Free Communication in IoT with 6G Using Multiple Machine Access Learning Collision Avoidance Protocol. *Mobile Networks and Applications*, pages 1–12, 2020.

[4] A. Marahatta, Q. Xin, *C. chi*, F. Zhang, and Z. Liu. PEFS: AI-driven Prediction based Energy-aware Fault-tolerant Scheduling Scheme for Cloud Data Center. *IEEE Transactions on Sustainable Computing*, pages 1–1, 2020.

[5] Y. Liu, J. E. Fieldsend, and G. Min. A Framework of Fog Computing: Architecture, Challenges, and Optimization. *IEEE Access*, 5:25445–25454, 2017.

[6] K. Dolui and S. K. Datta. Comparison of edge computing implementations: Fog computing, cloudlet and mobile edge computing. In *2017 Global Internet of Things Summit (GIoTS)*, pages 1–6, 2017.

[7] Isiaka Ajewale Alimi, Nelson Jesus Muga, Abdelgader M. Abdalla, Cátia Pinho, Jonathan Rodriguez, Paulo Pereira Monteiro, and Antonio Luís Teixeira. *Towards a Converged Optical-Wireless Fronthaul/Backhaul Solution for 5G Networks and Beyond*, chapter 1, pages 1–29. John Wiley & Sons, Ltd, 2019.

[8] Y. Liu, F. R. Yu, X. Li, H. Ji, and V. C. M. Leung. Distributed Resource Allocation and Computation Offloading in Fog and Cloud Networks With Non-Orthogonal Multiple Access. *IEEE Transactions on Vehicular Technology*, 67 (12):12137–12151, 2018.

[9] C. L. Stergiou, K. E. Psannis, and B. B. Gupta. IoT-based Big Data secure management in the Fog over a 6G Wireless Network. *IEEE Internet of Things Journal*, pages 1–1, 2020.

[10] L. Zhang, Y. Liang, and D. Niyato. 6G Visions: Mobile ultra-broadband, super internet-of-things, and artificial intelligence. *China Communications*, 16 (8):1–14, 2019.

[11] I. Tomkos, D. Klonidis, E. Pikasis, and S. Theodoridis. Toward the 6G Network Era: Opportunities and Challenges. *IT Professional*, 22(1):34–38, Jan 2020.

[12] S. U. Khan. Elements of Cloud Adoption. *IEEE Cloud Computing*, 1(1): 71–73, 2014.

[13] M. Chiang and T. Zhang. Fog and IoT: An Overview of Research Opportunities. *IEEE Internet of Things Journal*, 3(6):854–864, 2016.

[14] I. A. Alimi, A. Tavares, C. Pinho, A. M. Abdalla, P. P. Monteiro, and A. L. Teixeira. *Enabling Optical Wired and Wireless Technologies for 5G and Beyond Networks*, chapter 8, pages 177–199. IntechOpen, London, 2019.

[15] Hany F. Atlam and Gary B. Wills. Chapter Three - Intersections between IoT and distributed ledger. In Shiho Kim, Ganesh Chandra Deka, and Peng Zhang, editors, *Role of Blockchain Technology in IoT Applications*, volume 115 of *Advances in Computers*, pages 73 – 113. Elsevier, 2019.

[16] I. Alimi and A. Shahpari and A. Sousa and R. Ferreira and P. Monteiro and A. Teixeira. *Challenges and Opportunities of Optical Wireless Communication Technologies*, chapter 2. IntechOpen, London, 2017.

[17] Z. Ma, M. Xiao, Y. Xiao, Z. Pang, H. V. Poor, and B. Vucetic. High-Reliability and Low-Latency Wireless Communication for Internet of Things: Challenges, Fundamentals, and Enabling Technologies. *IEEE Internet of Things Journal*, 6(5):7946–7970, 2019.

[18] M. Weiner, M. Jorgovanovic, A. Sahai, and B. Nikolié. Design of a low-latency, high-reliability wireless communication system for control applications. In *2014 IEEE International Conference on Communications (ICC)*, pages 3829–3835, 2014.

[19] Corinna Schmitt, Claudio Anliker, and Burkhard Stiller. Chapter 8 - Efficient and Secure Pull Requests for Emergency Cases Using a Mobile Access Framework. In Quan Z. Sheng, Yongrui Qin, Lina Yao, and Boualem Benatallah, editors, *Managing the Web of Things*, pages 229 – 247. Morgan Kaufmann, Boston, 2017.

[20] I. A. Alimi, A. L. Teixeira, and P. P. Monteiro. Toward an Efficient C-RAN Optical Fronthaul for the Future Networks: A Tutorial on Technologies, Requirements, Challenges, and Solutions. *IEEE Communications Surveys Tutorials*, 20(1):708–769, 2018.

[21] Z. Liu, J. Zhang, Y. Li, L. Bai, and Y. Ji. Joint jobs scheduling and lightpath provisioning in fog computing micro datacenter networks. *IEEE/OSA Journal of Optical Communications and Networking*, 10(7):152–163, 2018.

[22] Pengfei Hu, Sahraoui Dhelim, Huansheng Ning, and Tie Qiu. Survey on fog computing: architecture, key technologies, applications and open issues. *Journal of Network and Computer Applications*, 98:27 – 42, 2017.

[23] E. Baccarelli, P. G. V. Naranjo, M. Scarpiniti, M. Shojafar, and J. H. Abawajy. Fog of Everything: Energy-Efficient Networked Computing Architectures, Research Challenges, and a Case Study. *IEEE Access*, 5:9882–9910, 2017.

[24] C. Nandyala and Haeng-Kon Kim. From Cloud to Fog and IoT-Based Real-Time U-Healthcare Monitoring for Smart Homes and Hospitals. *International Journal of Smart Home*, 10: 187–196, 2016.

[25] Syed Noorulhassan Shirazi, Antonios Gouglidis, Arsham Farshad, and David Hutchison. The Extended Cloud: Review and Analysis of Mobile Edge Computing and Fog From a Security and Resilience Perspective. *IEEE Journal on Selected Areas in Communications*, 35(11):2586–2595, 2017.

[26] J. Lin, W. Yu, N. Zhang, X. Yang, H. Zhang, and W. Zhao. A Survey on Internet of Things: Architecture, Enabling Technologies, Security and Privacy, and Applications. *IEEE Internet of Things Journal*, 4(5):1125–1142, 2017.

[27] E. Baccarelli, M. Scarpiniti, and A. Momenzadeh. Fog-Supported Delay-Constrained Energy-Saving Live Migration of VMs Over MultiPath TCP/IP 5G Connections. *IEEE Access*, 6: 42327–42354, 2018.

[28] Isiaka A. Alimi, Paulo P. Monteiro, and António L. Teixeira. Analysis of multiuser mixed RF/FSO relay networks for performance improvements in Cloud Computing-Based Radio Access Networks (CC-RANs). *Optics Communications*, 402:653 – 661, 2017.

[29] I. Alimi, P. Monteiro, and A. Teixeira. Outage probability of multiuser mixed rf/fso relay schemes for heterogeneous cloud radio access

networks (h-crans). *Wireless Personal Communications*, 95:27–41, 2017.

[30] Isiaka Ajewale Alimi, Abdelgader M. Abdalla, Akeem Olapade Mufutau, Fernando Pereira Guiomar, Ifiok Otung, Jonathan Rodriguez, Paulo Pereira Monteiro, and Antonio Luís Teixeira. *Energy Efficiency in the Cloud Radio Access Network (C-RAN) for 5G Mobile Networks*, chapter 11, pages 225–248. John Wiley & Sons, Ltd, 2019.

[31] Zhijiang Chen, Guobin Xu, Vivek Mahalingam, Linqiang Ge, James Nguyen, Wei Yu, and Chao Lu. A cloud computing based network monitoring and threat detection system for critical infrastructures. *Big Data Research*, 3:10 – 23, 2016. Special Issue on Big Data from Networking Perspective.

[32] L. Dai, B. Wang, Y. Yuan, S. Han, I. Chih-lin, and Z. Wang. Non-orthogonal multiple access for 5G: solutions, challenges, opportunities, and future research trends. *IEEE Communications Magazine*, 53(9):74–81, 2015.

[33] Z. Yang, Z. Ding, P. Fan, and N. Al-Dhahir. A General Power Allocation Scheme to Guarantee Quality of Service in Downlink and Uplink NOMA Systems. *IEEE Transactions on Wireless Communications*, 15(11):7244–7257, 2016.

[34] I. Alimi and O. Aboderin. Adaptive Interference Reduction in the Mobile Communication Systems. volume 1, pages 12–19, 2015.

[35] Y. Gu, Z. Chang, M. Pan, *L. song*, and Z. Han. Joint Radio and Computational Resource Allocation in IoT Fog Computing. *IEEE Transactions on Vehicular Technology*, 67(8):7475–7484, 2018.

[36] Y. Lan, X. Wang, D. Wang, Z. Liu, and Y. Zhang. Task Caching, Offloading, and Resource Allocation in D2D-Aided Fog Computing Networks. *IEEE Access*, 7:104876–104891, 2019.

[37] S. Tong, Y. Liu, M. Cheriet, M. Kadoch, and B. Shen. UCAA: User-Centric User Association and Resource Allocation in Fog Computing Networks. *IEEE Access*, 8:10671–10685, 2020.

[38] T. Huang, W. Yang, J. Wu, J. Ma, X. Zhang, and D. Zhang. A Survey on Green 6G Network: Architecture and Technologies. *IEEE Access*, 7:175758–175768, 2019.

[39] B. Zong, C. Fan, X. Wang, X. Duan, B. Wang, and J. Wang. 6G Technologies: Key Drivers, Core Requirements, System Architectures, and Enabling Technologies. *IEEE Vehicular Technology Magazine*, 14(3):18–27, Sep. 2019.

[40] W. Xia, X. Zhang, G. Zheng, J. Zhang, S. Jin, and H. Zhu. The interplay between artificial intelligence and fog radio access networks. *China Communications*, 17(8):1–13, 2020.

[41] J. Liu, B. Zhao, Q. Xin, J. Su, and W. Ou. DRL-ER: An Intelligent Energy-aware Routing Protocol with Guaranteed Delay Bounds in Satellite Mega-constellations. *IEEE Transactions on Network Science and Engineering*, pages 1–1, 2020.

[42] K. M. Sim. Agent-Based Approaches for Intelligent Intercloud Resource Allocation. *IEEE Transactions on Cloud Computing*, 7(2):442–455, April 2019.

[43] J. Park, S. Samarakoon, M. Bennis, and M. Debbah. Wireless Network Intelligence at the Edge. *Proceedings of the IEEE*, 107(11):2204–2239, Nov 2019.

[44] S. Zhang, J. Liu, H. Guo, M. Qi, and N. Kato. Envisioning Device-to-Device Communications in 6G. *IEEE Network*, 34(3):86–91, May 2020.

[45] K. M. Sim. Agent-based Cloud commerce. In *2009 IEEE International Conference on Industrial Engineering and Engineering Management*, pages 717–721, Dec 2009.