Republic of Macedonia

Ss CYRIL AND METHODIUS UNIVERSITY

FACULTY OF ELECTRICAL ENGINEERING AND INFORMATION TECHNOLOGIES

Institute of Telecommunications

Skopje

Stojan Kitanov

QUALITY OF CLOUD ORCHESTRATED SERVICES IN 5G MOBILE NETWORKS

- DOCTORAL DISSERTATION –

-ABSTRACT-

Supervisor: Prof. Toni Janevski, PhD

Skopje, October 2017

TABLE OF CONTENTS

Abstract

The rapid increase of the mobile devices resulted the mobile users to demand more and more ubiquitous mobile broadband services comparable to the fixed broadband Internet. In that direction on a global level the number of research initiatives on ***5G mobile networks*** *has rapidly increased. 5G networks would act as a nervous system of the digital society, on the economy and people's everyday lives and would create new paradigms of Internet services such as "Anything as a Service", where the devices, terminals, machines, smart things and the robots would become innovative tools that would create applications, services, and data and at the same time would use them. In order to satisfy the requests for big data processing and more intelligent networking demands in 5G networks, the main trend in the last decade was to push the processing, computing and the storage of data in the* ***Cloud Computing*** *environment. However, the cloud alone encounters growing limitations for reduced latency, high degree of mobility, high scalability, and real-time execution. A new paradigm called* ***Fog Computing*** *has emerged to overcome the limitations in the Cloud Computing technology. The fog computing extends the cloud computing at the edge of the network, that distributes computing, data processing, and networking services closer to end users. It is an architecture where distributed edge and user devices collaborate with each other and with the clouds to carry out computing, control, networking, and data management tasks. The application of fog computing in 5G mobile networks would contribute significantly to improve the network performances in terms of spectral and energy efficiency, enable direct device-to-device wireless communications, and support the growing trend of network function virtualization and the separation of network control intelligence from radio network hardware. This doctoral dissertation performs an evaluation of the quality of cloud and fog orchestrated services in 5G mobile networks through the delay (latency), user throughput, and the energy efficiency for the power consumed. In addition, 5 possible algorithms are proposed for an optimal selection of the best 5G radio access network depending from the types of the services that have certain requirements in terms of throughput, latency, and energy efficiency of the power consumed. The performed analysis of the results of this doctoral dissertation demonstrate that 5G networks would have a great benefit of implementing the fog computing technology, because the fog computing technology possess mechanisms that would handle with the new up-coming services, which would require low latency, high level of mobility, high scalability and real-time execution.*

Keywords: 5G, Cloud Computing (CC), Edge Computing, Fog Computing, Fog Networking, Internet of Everything (IoE), Internet of Things (IoT), Mobile Cloud Computing (MCC), Mobile Edge Computing (MEC), Web of Things (WoT).

Part One

1 Introduction

The rapid development of the telecommunications which was intensified in the last three decades, as well as, the necessity for generalizing the telecommunication networks in the direction of service integration, resulted with continuous advancement of the network architectures and the mechanisms in order to satisfy the increasing user demands. Until now, the development in the field of telecommunications has been marked by extreme progress, and something that was thought to be impossible 20 years ago, today it becomes a reality, through the telecommunications technology itself. For example: using different types of services that require data rates from the order of Gbps from the mobile devices, while the users are in a motion by a train, car or simply they walk; then availability of the services from any place, any time and any device, online payment of the bills from a personal computer, or a smartphone, the existence of 3D Hologram capable devices and services; usage of different sensors that measure temperature, air humidity, light intensity under different circumstances, data transfer with the speed of light to any place on the earth, and many other examples that are enabled by the modern telecommunication networks. Today, 3G wireless and mobile networks provide IP connectivity for the real time and non-real-time services with satisfactory quality of service, and also already have been implemented 4G wireless and mobile networks that have much better performances than 3G.

On the other hand, today other heterogeneous opportunities exist for cooperation and inter-working with 3G and 4G wireless and mobile networks such as the standards 802.11 (WLAN), LTE, LTE-Advanced (LTE-A), LTE-Advanced Pro (LTE-A Pro), IEEE 802.16 Wireless Metropolitan Area Network (WMAN) and the wireless networks for digital TV broadcasting (with resolution). Also, recently are quite important not only in the scientific areas, but also in the industrial fields the Personal Ad-hoc Networks (PANs) and Mobile Ad-hoc Networks (MANETs).

1.1 The Fifth Generation of Mobile Networks – 5G

In this context and under the continuous pressure by the users and their service requirements is initiated the development of the next **Fifth generations of mobile and wireless networks – 5G**, whose architecture should support the existing services, as well as to provide fast and easy development of new services. The 5G is currently still under development phase [1 – 12]. At the moment, only the main concepts and suggestions for 5G exist, and also some standards that have very high probability to be included in 5G have already appeared. It is expected the 5G standard to contain the standards IEEE 802.11ac, IEEE 802.11ad and IEEE 802.11af, 802.16 and 802.21 and many similar standards to 4G like LTE, LTE-Advanced (Release 10, 11, 12, etc.), and LTE-Advanced Pro (LTE-A Pro) and the mobile WiMAX 2.0 (IEEE 802.16m). When it is talked about next generation of wireless networks, then it is meant of a system composed of wireless system that provide service continuity to the user.

5G would require new more smart and intelligent devices capable to provide a broad range of multimedia services (speech, audio and video applications) to the mobile users with extended mobility, support of broadband connections, high processing power of the mobile device, support of machine-to-machine types of communications, advanced network technologies, better utilization of the wireless media, load balancing, advanced QoS support for any type of service (lower probability for network outage, lower packet bit error rate probability, minimum delay of few milliseconds, higher data rates per user, multi-homing, multi-streaming, etc.), as well as, longer battery life, higher memory, that will provide enough capacity for control information and many other advanced capabilities. 5G definitely would have user-centric oriented standard, where the smart mobile devices can simultaneously be connected to one, or more available wireless mobile networks and can combine the flows from different networks. Of course, the whole intelligence and data processing should not entirely be left to the mobile device, but it is necessary to advance and upgrade the remaining part of the network with additional intelligence for control and additional nodes that would support the above-mentioned capabilities in 5G mobile devices.

1.2 Cloud Computing

In parallel with the development of the networks and the request for new services by the end users of the mobile networks appeared the concept of **Cloud Computing in fixed, or mobile environment (Mobile Cloud Computing)** [13 – 21]. The cloud computing technology already changed the way of applications being developed and the applications being accessed. Cloud computing enables data processing based on Internet, where multiple connected servers provide resources, software and data for the computers and other devices that have sent request to the servers, an approach which is very similar to the electric power distribution network. At the moment the analysis and processing of data for example video, pictures, audio, animations (which is present in many fields for example in the business, ecology, education, health, social networks, etc.), requires more processing powerful resources than the mobile device can provide. Considering the limitations of the mobile devices such as memory, processing power and bandwidth, the mobile cloud can be used for data processing and delivery of the final results to the end users.

1.3 Fog Computing

However the future Internet would exacerbate the need for improved QoS/QoE, supported by services that are orchestrated on-demand and are capable to adapt at runtime, depending on the contextual conditions, to allow reduced latency, high mobility, high scalability, and real-time execution. These demands can only be partially fulfilled by existing cloud computing solutions.

The conventional cloud computing environment itself is completely centralized by nature and the network is not capable to provide real time services with low latency to billion IoT devices at the same time. At the virtualization process, cloud computing servers offer their services from different geographic locations. This would result with high level of latency during the service provisioning [22 – 23].

A new paradigm called **Fog Computing**, or briefly **Fog** has emerged to meet these requirements [24 – 27]. Fog extends cloud computing and services to the edge of the network, and provides data, computing, storage, and application services to end-users that can be hosted at the network edge, or even end devices such as set-top-boxes, or access points. It reduces service latency, and improves QoS/QoE, that results in superior user experience.

Fog Computing would provide support of present and future IoT applications (industrial automatics, transport, sensor networks and actuators) that require mobility and predictable low latency in real time [28 – 29]. Therefore, fog computing should be considered as a serious candidate for 5G networks, where the cloud would be diffused among the user devices. A potential application of fog computing in 5G network can be found in [30], where are pointed several important and crucial cases and scenarios such as network selection in heterogeneous environment, borrowing bandwidth from the neighbour devices at M2M, or D2D communications, etc.

Fog and Cloud computing are two technologies that are complementary to each other. These complementary functions of fog and cloud would enable the 5G users to use a new technology which would respond to their low latency and real-time service requirements for the applications that are being executed at the edge of the network [31], and at the same time would provide support of complex data analysis and data processing, as well as their long-term storage in 5G core.

1.4 Motivation for the Doctoral Dissertaion

The move from cloud to fog in 5G brings out several key challenges, including the need for supporting the on-demand orchestration and runtime adaptation of resilient and trustworthy Fog Services.

Inspired by such extreme and rapid development of telecommunications, and ICT in general, especially in the field of mobile and wireless communications in the last decade, this doctoral dissertation presents novel scientific researches, frameworks, analysis, comparisons for the quality of cloud and fog orchestrated services in 5G mobile networks.

From the current literature survey and review very rarely can be found a reference about the quality of cloud orchestrated services in 5G mobile networks. That's why the main idea in this dissertation is to contribute in providing more details about defining the cloud and fog orchestrated services in 5G mobile networks, which would contribute in the creation of the 5G standard in its final form (for whom only concepts still exist).

1.5 Research Subject and Hypothesis

The main Research Subject in this doctoral dissertation is the analysis and comparison of the quality of cloud and fog orchestrated services in 5G mobile networks. Here is presented a model and architecture of 5G in a fog computing environment, and also the service orchestration mechanisms are considered. In addition, this dissertation presents results and analysis of simulations and practical experiments for different scenarios under different circumstances and different network technologies in which are considered the quality of fog orchestrated services and their comparison to the quality of cloud orchestrated services in 5G mobile networks. As input data are used different reference QoS and KPI parameters, such as peak data rate, bandwidth, delay, etc., and as output data are presented the latency, user throughput and energy efficiency.

In the researches of this dissertation, as the main (general) **hypothesis** appears the following:

The introduction of Fog Computing Service Orchestration Mechanisms contributes to a higher satisfactory level of all QoS parameters and obtains higher estimations of QoS parameters of any multimedia services.

The first **special thesis** is the following:

1) *The introduction of Fog Computing Service Orchestration Mechanisms in 4G homogeneous mobile and wireless IP networks contributes to a higher satisfactory level of all QoS parameters for any given multimedia service.*

The second **special thesis** is the following:

2) *The introduction of Fog Computing Service Orchestration Mechanisms in 5G heterogeneous mobile and wireless IP networks contributes to a higher satisfactory level of QoS parameters for any given multimedia service.*

1.6 Organization of the Doctoral Dissertation

This doctoral dissertation is organized in three parts.

The first part contains 6 chapters. Chapter 1 is an introduction, and provides the motivation, the research subject, and the list of publications that came out from this dissertation.

Chapter 2 is about 5G mobile networks, its service and performance requirements, its architectures, 5G smart mobile devices, applications and services.

Chapter 3 is about Internet of Everything (IoE), which represents a natural evolution of Internet of Things (IoT).

Chapter 4 discusses the Cloud Computing in fixed and mobile environment. Also different deployment models of the cloud, service oriented architecture and the benefits of cloud computing are considered.

Chapter 5 considers the fog computing, which is extension of the cloud computing at the edge of the network. In addition a comparison between cloud and fog computing is performed, and the main fog computing features are presented.

At the end of Part One, Chapter 6 considers the cloud in 5G mobile networks. Firstly, it is discussed why telco operators should implement the cloud in their network, then it is considered the possible cloud computing architecture in 5G RAN and 5G core. Then the concepts Mobile Edge Computing (MEC) and Fog Computing are considered for 5G and both of them are compared. Finally, some possible cases and scenarios of applying the fog computing in 5G mobile networks are considered.

Part two of this dissertation contains the next three chapters.

A novel model and architecture for the fog computing service orchestration mechanisms in 5G mobile networks are proposed in Chapter 7.

The evaluation of the quality of cloud and fog orchestrated services in 5G mobile networks is performed in Chapter 8. The following parameters are considered: latency (end-to-end delay), user throughput in downlink and uplink direction, and energy efficiency in downlink and uplink direction. These parameters are compared in cloud and fog computing environment for each 3G, 4G, or 5G mobile network separately, and then in all three networks together. The results clearly demonstrate that fog computing environment provides better performances than cloud computing environment in any mobile network. And these results actually provide a proof for the statements in the main hypothesis and the special thesis given in the previous section. In addition, 5G network has much better performances than 3G and 4G mobile networks.

Finally, Chapter 9 provides 5 possible algorithms for an optimal network selection depending from user service requirements. Each algorithm selects the best network according to some parameter (latency, bit rate or energy efficiency) under some given limits and constraints. A comparison is made what results are obtained when the algorithm is applied and when the algorithm is not applied. It is concluded that each algorithm offers better performances compared to the case when the algorithm is not applied.

Finally, Part 3 (Chapter 10) gives a conclusion of this doctoral dissertation.

1.7 List of Publications from the Doctoral Dissertation

1.7.1 Publications in Book Chapters

- Stojan Kitanov, and Toni Janevski. **Fog Computing Mechanisms in 5G Mobile Networks**. *Book Chapter in **Networks of the Future : Architectures, Technologies and Implementations***. CRC Press Taylor and Francis Group, Florida, USA, 2017.
- Stojan Kitanov, and Toni Janevski. **Fog Networking for 5G and IoT**. *Book Chapter in **5G Mobile: From Research and Innovations to Deployment Aspects, Architectures, Technologies and Implementations***. Nova Publishers, Inc., 2017.

1.7.2 Publications in International Journals

- Stojan Kitanov, Toni Janevski. **Fog Computing Service Orchestration Mechanisms for 5G Networks**. *Journal of Internet Technology (JIT), ISSN 1607-9264*, **Taiwan,** to be published in January, 2018. (**2016 IMPACT FACTOR: 1.930**).
- Stojan Kitanov, and Toni Janevski. **Quality Evaluation of Cloud and Fog Orchestrated Services in 5G Network**. *European Journal of Information Science and Technology (EJIST), Vo l. 1, No. 1*, published by University of Information Science and Technology "St. Paul the Apostle, " Ohrid,

Macedonia, 2016.

- Stojan Kitanov, and Toni Janevski. **Fog Computing as a Support for 5G Network**. *Journal of Emerging Research and Solutions in ICT (ERSICT), Vol. 1, No. 2*, published by Faculty of Information and Communication Technologies in Bitola, Macedonia, 2016.

1.7.3 Publications in International Conferences

- Stojan Kitanov, Toni Janevski. **Energy Efficiency of Fog Computing and Networking Services in 5G Networks**. *Proceedings of the XVII IEEE International Conference on Smart Technologies EUROCON 2017*, Ohrid, Macedonia, 2017.
- Stojan Kitanov, and Toni Janevski. **State of the Art: Fog Computing for 5G Networks**. *Proceedings of the XXIV International Telecommunication Forum TELFOR 2016*, Belgrade, Serbia, 2016.
- Stojan Kitanov, and Toni Janevski. **Energy Efficiency of 5G Mobile Networks in Hybrid Fog and Cloud Computing Environment**. *Proceedings of the II International Conference Recent Trends and Applications in Computer Science and Information Technology RTA-CSIT 2016*, Tirana, Albania, 2016.
- Stojan Kitanov, Toni Janevski. **5G Networks in the Fog Computing Environment**. *Proceedings of the XII International Conference of Society for Electronics, Telecommunications, Automation and Informatics ETAI 2016*, Struga, Macedonia, 2016.
- Stojan Kitanov, Toni Janevski. **Fog Computing for 5G Networks**. *Accepted paper as a poster presentation on 8th ICT -Innovations International Conference*, Ohrid, Macedonia, 2016.
- Stojan Kitanov, Toni Janevski. **Hybrid Environment Service Orchestration for Fog Computing**. *Proceedings of the 51st International Scientific Conference on Information, Communication and Energy Systems and Technologies ICEST 2016*, Ohrid, Macedonia, 2016.
- Stojan Kitanov, Edmundo Monteiro, Toni Janevski. **5G and the Fog – Survey of Related Technologies and Research Directions**. *Proceedings of the XVIII Mediterranean IEEE Electrotechnical Conference MELECON 2016*, **Limassol, Cyprus**, 2016.
- Stojan Kitanov, Toni Janevski. **Mobile Cloud Computing in 5G Networks.** *Proceedings of the XI International Conference of Society for Electronics, Telecommunications, Automation and Informatics ETAI 2015*, Ohrid, Macedonia, 2015.
- Stojan Kitanov, Toni Janevski. **State of the Art: Mobile Cloud Computing**. *Proceedings of the Sixth IEEE International Conference on Computational Intelligence, Communication Systems and Networks*, pp. 153-158, Tetovo, Macedonia, 2014.
- Stojan Kitanov, Toni Janevski. **Performance Evaluation of Scheduling Strategies for LTE Networks in Downlink Direction**. *Proceedings of the XI International Conference of Society for Electronics, Telecommunications, Automation and Informatics ETAI 2013*, Ohrid, Macedonia, 2013.
- Stojan Kitanov, Toni Janevski. **Modeling and Analyzing LTE Networks with EstiNet Network Simulator and Emulator**, *Proceedings of the XVIII International Scientific Conference on Information, Communication and Energy Systems and Technologies (ICEST)*, **Ohrid, Macedonia**, 2013.

2 Fifth Generation of Mobile Networks – 5G

Today we are witnesses of the enormously increased volume of mobile traffic, the fast development of the broadband mobile networks and the appearance of intelligent mobile devices. According to some predictions, it is expected the global mobile traffic from 2010 to 2020 to be increased for 1000 times. This is due to the widespread of smart mobile devices, and the increased demand of advanced multimedia services such as UHD (Ultra-High Definition) and 3D video, as well as extended reality and experience [32 – 33], as well as the social networks [34].

2.1 Migration to 5G

The Fifth Generation of mobile and wireless networks or 5G is a name which is used in some research papers and projects to denote the next major phase of mobile telecommunications standard. 5G network would include support for large number of connected devices and flexible air interfaces, different interworking technologies that are energy efficient, and possess always on-line capabilities. This requires not only upgrade of existing systems, but also innovation of new protocols and new access technologies altogether.

There are three possible migration paths to 5G network (see Figure 2.1) [52 – 53]:

- a step-by-step evolutionary path focusing on further enhancements of existing technologies;
- a revolutionary path using brand new innovative technologies; or
- a symbiotic integration and convergence of existing, or new technologies such as communication, information systems and electronics, multi radio access technologies, computing techniques, device-to-device communications, bands, links, layers, services, multiplexing, etc.

As it is shown in Figure 1, 5G is a multi-layered heterogeneous network that consists of existing 2G, 3G, 4G and future Radio Access Technologies (RATs). It may also converge many other radio technologies like Mobile Satellite System (MSS), Digital Video Broadcasting (DVB), Wireless Local Access Network (WLAN), Wireless Personal Access network (WPAN), Worldwide Interoperability for Microwave Access network (WiMAX), etc., with multi-tiers coverage by macro, pico, femto, relay and other types of small cells. 5G would support a wide range of applications and services to satisfy the requirements of the information society by the year 2020 and beyond [5], [9].

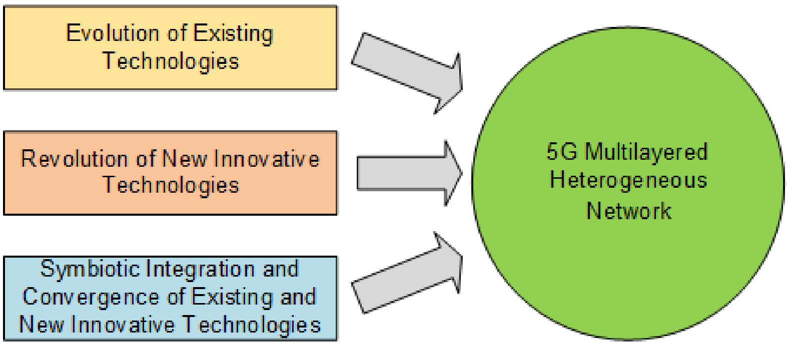


Figure 2.1 Technology Routes to 5G [54]

2.2 5G Performances and Service Requirements

In order to fulfill these demands 5G systems would be required to deliver a higher order of magnitude cell capacities and per-user data rate compared to 4G. 5G would be a set of telecommunication technologies and services that support 1000 times more data capacity than today, and should provide ultra-low latency response of less than few milliseconds. The network should provide a capacity of 50 Gbps per cell, and guarantee anywhere more than 1 Gbps per user through super dense networking, regardless of the user location, including the cell edge. Compared to 4G, the cell spectral efficiency would be increased by 3 to 5 times, and the latency response in control plane would be reduced to one half, i.e. to 50 ms. In addition, 5G would support ultralow latency response of 1 ms in data plane, that equals to one tenth of 4G network [12], [52 – 53], [55].

Table 2.1 A Comparison QoS parameters between 4G and 5G network

Parameter	4G	5G
Air Link User Plane Latency	10 ms	1 ms
Air Link Control Plane Latency	100 ms	50 ms
Jitter	20 ms	0.02 ms
Density of simulatenous connections per unit km^2	10^5	10^6
Density of traffic volume per unit km^2	0.5 Tbps/ km^2	20 Tbps/ km^2
Mobility	300 km/h	500 km/h
Uplink Cell Spectral Efficiency	1.8 bps/Hz	5 bps/Hz
Downlink Cell Spectral Efficiency	2.6 bps/Hz	10 bps/Hz
Peak Throughput (Downlink) per Connection	100 Mbps to 1 Gbps	10 Gbps to 50 Gbps
Downlink Cell Edge Data Rate	0.075 bps/Hz/cell	Anywhere 1 Gbps
Uplink Cell Edge Data Rate	0.05 bps/Hz/cell	Anywhere 0.5 Gbps
Cost Efficiency	10 times	100 times
Packet delay budget without quality assurance	100 to 300 ms	undefined
Packet delay budget with guaranteed quality	50 to 300 ms	1 ms
Average Packet Loss Ratio for Video Broadcast	10^{-8} (4k UHD)	10^{-9} (8k UHD)
Average Packet Loss Ratio for M2M services (without quality assurance)	10^{-3}	10^{-4}
Average Packet Loss Ratio for M2M services (with guaranteed quality)	10^{-6}	10^{-7}

In order to enable the forthcoming Internet of Everything (IoE) [28], 5G should provide 4A (Anytime, Anywhere, Anyone, Anything) massive and simultaneous connectivity which would accommodate one million different mobile devices per unit square kilometre. It would have flexible and intelligent network architecture with software based structure capable to analyse data in real time and to provide intelligent and personalized services [10].

Table 2.1 summarizes the QoS comparison between 4G and 5G networks [12], [52 – 53], [55].

2.3 5G Applications and Services

5G would have user-centric approach, where telecom operators would invest in developing new applications in order to provide ubiquitous, pervasive, seamless, continual and versatile mobile experience to the end-user [5]. The applications would become more personalized, and more context-aware capable to recognize user identity, user location, and user preferences [56].

Therefore, context-rich support services such as context extraction service, recommendation service and group privacy service should be supported in 5G. Particularly important is the context extraction service that performs data mining analysis of mobile data combined with other forms of data such as social networking data, and sensor network data in order to extract contextual clues relevant to the user. Data mining services should be able to scale and analyse large group of people and large quantities of data (big data) in order to extract collective trends among the population of users in real time. Additionally, recommendation services based on collective group context rather than individual context need to be created and scaled. By using these clues, a layer of recommendation services can be built that creates output which is adjusted to a user, or set of users with those contextual characteristics [56].

2.4 5G Network Architectural Levels

5G network would be an all IP Network that would provide a continued evolution and optimization of the system concept in order to provide a competitive edge in terms of both performance and cost [5], [9], [11], [56]. This would lessen the burden on aggregation point and traffic would directly move from base station to media gateways. The network architecture in 5G consists of several levels that are shown in Figure 2.1.

The innovative service and content providers are located at the top level, and they should accommodate 5G requirements and would provide new ubiquitous and pervasive user experience. New applications and services such as augmented and virtual reality, hologram, mobile ultra-high definition, cloud computing, etc. would be provided.

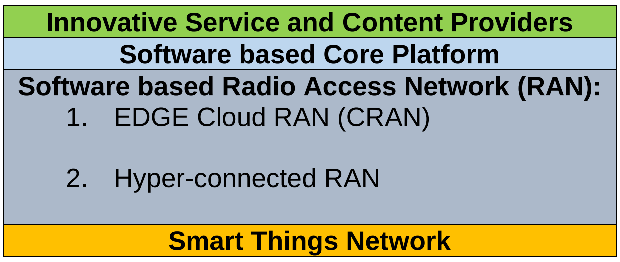


Figure 2.2 5G Network Architectural Levels [54]

One level below is the enabling software-based core platform. At this level flexible and reliable flat IP architecture would converge into different technologies to form a single 5G core, where a range of intelligent complex telecommunication network functions can be efficiently implemented. 5G core consists of reconfigurable multi-technologies combined in single core called super-core, nano-core, or master-core [11]. All network operators would be connected to one single network core with massive capacity which would eliminate all interconnecting charges and complexities, in which right now the network operator is facing, and would reduce the number of network entities in an end-to-end connection, thus reducing latency significantly. The main functionality in 5G core would be Network as a Service (NaaS) platform that allows configuration of all telecommunication and service functions with virtualized software on a programmable hardware [56].

The third level is the Radio Access Network (RAN) that consists of EDGE Cloud RAN (EDGE CRAN) service and Hyper-connected RAN. EDGE CRAN moves the cloud from the innovative service content provider to the radio access network. By placing storage and computing resources at, or close to, the cell site, operators can improve response times, making services feel “snappier” and, uniquely, more responsive to prevailing radio conditions.

Hyper-connected Radio Access Network (RAN) infrastructure acts as a data pipe supporting the massive and ultra-high-speed connectivity over ultra-dense networks (UDN) of wireless access with heterogeneous cells arrangement. It is a multi-layered heterogeneous network that contains the existing 2G, 3G, 4G and future Radio Access Technologies (RATs). It may also converge many other radio technologies like Mobile Satellite System (MSS), Digital Video Broadcasting (DVB), Wireless Local Access Network (WLAN), Wireless Personal Access network (WPAN), Worldwide Interoperability for Microwave Access network (WiMAX), etc., with multi-tiers coverage by macro, pico, femto, relay and other types of small cells. Cognitive radio and software defined networking together with existing, or new modulation and transmission methods can be used for deploying new applications and services which would improve the utilization of the congested RF spectrum. The RAN infrastructure is responsible for cellular content caching, radio and resource allocation, mobility, interference control, etc.

Finally, the lowest level of the 5G network architecture consists of smart things network with different smart mobile, Internet of Things (IoT) and Web of Things (WoT) end user devices. This includes various devices from laptops, smartphones, tablets, phablets, and wearable devices to connectivity embedded devices and sensors in cars, trucks, bikes, etc.

The architectural levels in 5G would be defined primarily by the service requirements, and the following technologies: nanotechnology, quantum cryptography, and (mobile) cloud computing.

2.5 5G Smart Mobile Device

5G network would provide ubiquitous and pervasive connectivity with smart capabilities to deal with the communication requirements of smartphones, smart TV and smart small devices, and to include efficient and effective management of resources. 5G smart mobile device would be fully autonomously reconfigurable device attached to several mobile, or wireless networks at the same time, and would use the flows of mobile accesses (e.g., 2G, 3G, or 4G) and/or wireless accesses (e.g., WiFi and WiMAX) in a separate and/or integrated manner (i.e., the multiple connection capability) [5]. Each network access technology would be responsible for handling user-mobility, while the terminal would make the final choice among different wireless/mobile networks. An overview of the 5G mobile smartphone device protocol stack layers compared to the OSI layers is given in Figure 2.2.

The physical and data link OSI layers would be based on Open Network Access Layer, i.e. any existing or new mobile and wireless access network, such as LTE/LTE-A, LTE-A Pro, WiFi, WiMAX, etc. would be used [5].

The network layer (see Figure 2.3) shall consist of two layers: upper network layer and lower network layer [5].

The upper network layer shall contain the fixed IPv6 address for the mobile device that would be implemented in the mobile phone by 5G manufacturers. At lower network layer Mobile IP (MIP) protocol would be used. Because the mobile device would be simultaneously connected to several mobile and wireless networks at this level the mobile device should maintain different IP addresses for each of the radio interfaces. Each of these IP addresses would be Care of Address (CoA) for the Foreign Agent (FA) placed in the mobile phone. This

FA would perform the CoA mapping between the fixed IPv6 address and the CoA for the current mobile or wireless network. Like that the 5G mobile device shall maintain multi-wireless network environment. The network address translation is performed by the middleware between the upper and lower network layers. In addition, MIP protocol can also be used to perform service continuity while the device performs roaming, or moves from one wireless access network to another.

At the transport layer Open Transport Protocol would be used [5]. This would enable to download and install protocol such as TCP, RTP, etc., or new transport protocol that is targeted to a specific mobile, or wireless network.

The application layer would be used for QoS and QoE management. The QoS and QoE parameters would be stored in a database in the 5G mobile device with aim to be used by intelligent algorithms running in the mobile terminal as system processes, which at the end shall provide the best wireless connection upon QoS and QoE requirements and personal cost constraints [5].

OSI Layers	**5G Smart Mobile Device Layers**
Application Layer	Application Layer (Applications, Services, and QoS/QoE Management)
Presentation Layer	
Session Layer	Open Transport Layer (TCP, RTP, or new protocol version for download)
Transport Layer	
Network Layer	Network Layer
Data Link Layer	Open Network Access Layer (LTE-A, WiFi, IEEE 802.16m, etc.)
Physical Layer	

Figure 2.3 5G Smartphone Device Protocol Stack Layers [5]

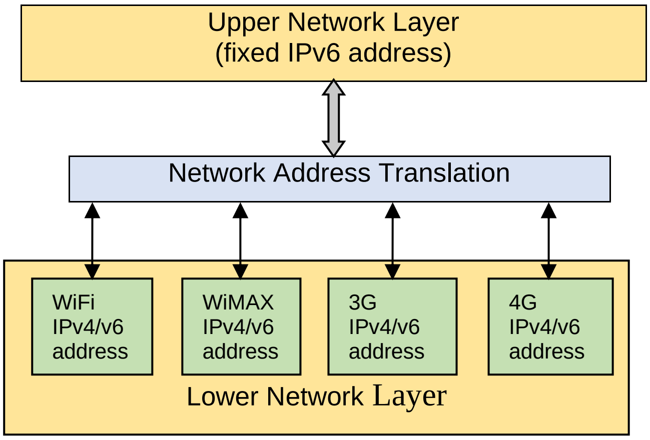


Figure 2.4 5G Smartphone Network Layer [5]

3 Internet of Everything (IoE)

Because of ICT developments, many end-user devices, networks and services have been acquiring more complicated features and capabilities. At the moment the Internet progressively evolves from a network of interconnected computers (Internet of Computers – IoC) to a network of interconnected objects, Internet of Things (IoT) and moreover any-thing communications, called Web of Things (WoT) [59]. The next step in the development of future Internet would be a clear evolution of the Internet of Things (IoT) into Internet of Everything (IoE), which would introduce high mobility, high scalability, real-time, and low latency requirements that raise new challenges on the services being provided to the users [28].

3.1 Internet of Things (IoT)

The **Internet of Things (IoT)** is a system of internetworking smart physical objects or "things" embedded with electronics, software, sensors, actuators, and network connectivity that enable these objects to collect and exchange data over various networking interfaces and Internet [28 – 29], [60 – 61]. In 2013 the Global Standards Initiative on Internet of Things (IoT-GSI) defined IoT as *the infrastructure of the information society*.

IoT allows objects to be sensed, and/or controlled remotely across existing network infrastructure. The smart things may have their own IP addresses, can be embedded in complex systems, use sensors to obtain information from their environment, and/or use actuators to interact with it. When IoT is augmented with sensors and actuators, the technology becomes an instance of more general class of cyber-physical systems, which also encompasses technologies such as smart grids, smart homes, intelligent transportation, and smart cities.

Through the exploitation of identification, data capture, processing and communication capabilities, the IoT makes full use of things to offer services to all kinds of applications, whilst maintaining the required privacy. The Internet of Things would create opportunities for more direct integration between the physical world and computer-based systems, that result in improved efficiency, accuracy and economic benefit.

Experts estimate that IoT would consist of almost 50 billion objects by 2020 [62], which would generate large amounts of data from diverse locations. In addition, there would be consequent necessity for quick aggregation of such big data, as well as to index, store, and process the big data more effectively.

3.2 Web of Things (WoT)

IoT has limited capabilities in the integration of the devices from various manufacturers into a single application or system, because many incompatible IoT protocols exist. As a result, the integration of data and services from various devices is extremely complex and costly. In addition, no unique and universal application layer protocol for IoT exists, that can work across many available networking interfaces.

Rather than re-inventing completely new standards, this issue can be solved by using the existing and well-known Web application layer protocols, standards and blueprints for connecting heterogeneous devices. This is known as **Web of Things (WoT)** [59].

The WoT is considered an IoT by integrating smart things not only into the Internet (the network), but into the Web (the application layer). WoT provides an Application Layer that simplifies the creation of IoT applications.

The physical devices can be accessed as web resources by using standard Web protocols, where the services/applications can be provided either upon a web-based service environment, or legacy telecommunications. Therefore, the integration across the systems and applications would be much easier. The services and data offered by smart web objects would be available to a larger pool of web developers that enables them to build new scalable and interactive and beneficial applications for everyone.

3.3 Internet of Everything (IoE)

In future everyday objects would interact and communicate each other by using the network intelligence that allows convergence, orchestration and visibility across disparate systems. Large amounts of data would circulate between the objects in order to create smart and proactive environments which would significantly enhance the user experience. Smart interacting objects should adapt to the current situation with or without any human involvement. This would cause IoT and WoT to evolve into **Internet of Everything (IoE)** [28 – 29], [60 – 61].

IoE is the intelligent connection of people, process, data and things. It brings together people, process, data, and things to make networked connections more relevant and valuable than ever before [28 – 29], [60 – 61]. It also turns information into actions that create new capabilities, richer experiences and unprecedented economic opportunity for businesses, individuals and countries.

The main drivers which would enable IoE to become reality are: the development of IP devices, the global availability of broadband services and the advent of IPv6. The network plays a critical role in the Internet of Everything – it must provide an intelligent, manageable, secure infrastructure that can scale to support billions of context-aware devices. The main elements in IoE are: *people, process, data,* and *things* (See Figure 3.1).

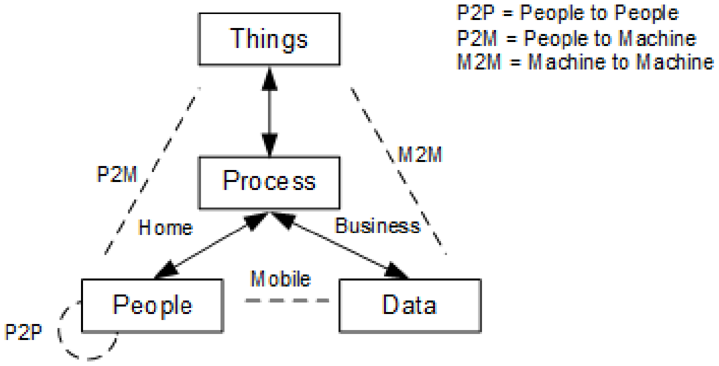


Figure 3.1 Internet of Everything Connectivity [29]

People, or **humans** would be able to connect to the Internet with different devices, like smartphones, PCs and tablets, as well as through social networks, such as Facebook and LinkedIn. In addition, they themselves would become nodes on the Internet, with both static information and a constantly emitting activity system (for example the health status) [61].

Data shall be collected by the smart things, where it would be transformed into useful rich information which would be transmitted throughout the Internet to a central source (machines, computers and people) for further processing, analysis, evaluation and decision making. The data transform into information is very important because it would enable us to make faster, and more intelligent decisions, as well as control our environment more effectively [61].

The **smart things**, or objects consists of sensors, meters, actuators, consumer devices, and enterprise assets that are connected to the Internet and each other. In IoE, these smart things would sense more data, would become context-aware, and would provide more experiential information to help people and machines to make more relevant and valuable decisions [61].

The **process** manages the way people, data, and things work together. It plays an important role in the way of communication and cooperation among people, data, and things in order to deliver an economic value and social benefits in the connected world of IoE. With the correct process, connections become relevant and add value because the right information is delivered to the right person at the right time in the most appropriate way [61].

4 Cloud Computing

The idea of cloud computing is based on a very fundamental principal of reusability of IT capabilities [13], [63]. **Cloud computing** is a computing paradigm, where a large pool of systems are connected in private or public networks, in order to provide dynamically scalable infrastructure for application, data and file storage [64]. At the same time, the shared cloud resources (networks, servers, data warehouses, applications and services) can be rapidly provisioned and managed with minimal interaction by service providers.

The cloud computing users may use these resources for development, hosting and running of services and applications on demand in a flexible way at any device, at any time and at any place in the cloud. With the advent of this technology, the cost of computation, application hosting, content storage and delivery is reduced significantly.

The framework for cloud computing is defined in the recommendations ITU Y.3501 [14] and Y.3510 [15] as well as the NIST standards for cloud computing [16].

4.1 Mobile Cloud Computing

Today the mobile devices (smartphones, tablets, etc.) became essential part of the life, as well as, the most effective and convenient tools for communication without any limitations for time and space. The mobile user device accumulates rich experience of the mobile applications (iPhone applications, Google applications etc.), that are executed either on the mobile devices, or on the distant servers through the wireless networks. The fast development of the Mobile Computing (MC) resulted to become a powerful trend in the information technology. But on the other hand the mobile devices face with many challenges with their resources (battery life, memory, storage, bandwidth, processing power), environment (heterogeneity, availability and scalability) and security (reliability and privacy). The constrained resources additionally worsen the improvement of the quality of the services.

With the rapid increase of the mobile applications and the support of the cloud computing for different types of services for the mobile device users is introduced the **Mobile Cloud Computing (MCC)** as an integration of the cloud computing in the mobile environment [17], [65 – 66]. The mobile cloud computing brings new types of services for the mobile device that would use the benefits of the cloud computing.

MCC is an infrastructure used by mobile applications where both the data storage and data processing are moved away from the mobile device to powerful and centralized computing platforms located in the clouds [20]. The access to this platform is enabled through the wireless network by using a thin web client, or browser. Mobile cloud computing is usually referred to the following two perspectives [65]:

- *infrastructure based* – where the hardware infrastructure is static and provides cloud services to mobile users; and
- *ad-hoc mobile cloud* – where a group of mobile devices acts as a cloud and provides cloud services to other mobile devices

4.2 Cloud Computing Model and Architecture

(Mobile) cloud computing is made up of complex network and relationships of and in between infrastructure Providers, Application/Services Providers, End-Users and Developers, all producing and/or consuming applications and/or services on web [18]. Such MCC model is given on Figure 4.1.

The Infrastructure Providers provide hardware and software infrastructure, or services and applications, and/or all the above. The Application/Services Providers are 1st tier consumer of Cloud Computing. They are typically business consumers of cloud computing infrastructure and providers of applications and/or services. The Developers are 2nd tier consumer of Cloud Computing, and they develop applications and services that are typically hosted on the Cloud.

The End Users also known as 3rd tier consumer of Cloud Computing, are typical end users of applications. They consume applications that in turn consume services on the cloud, and they care whether the application works well when needed with the necessary availability level and the security.

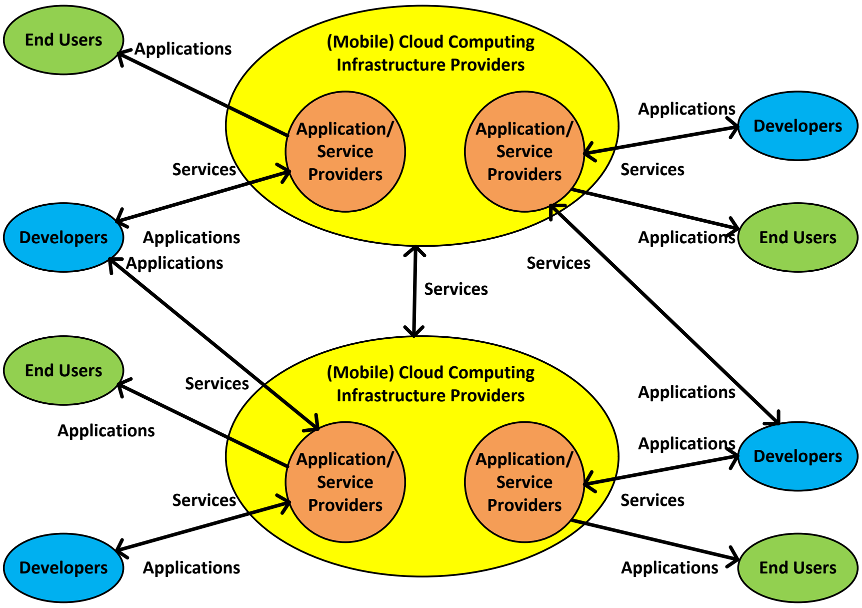


Figure 4.1 (Mobile) Cloud Computing Model [18]

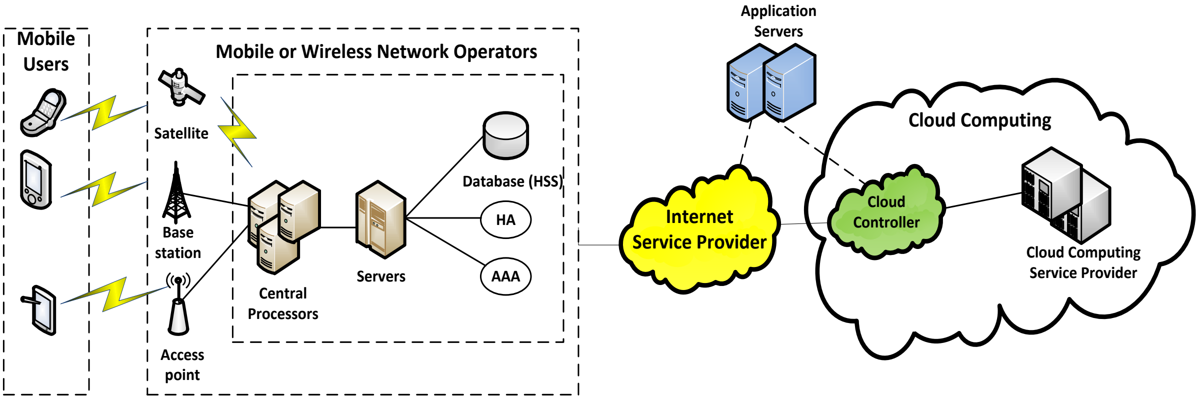


Figure 4.2 Mobile Cloud Computing Architecture [17]

A general mobile cloud computing architecture [17] is given on Figure 4.2. Mobile devices are connected to the mobile, or wireless network (GSM, GPRS, UMTS, HSPA, LTE, LTE-Advanced, LTE-A Pro, etc.) through a base station (BTS, UTRAN, nodeB, enodeB), satellite link or access point (WiFi or WiMAX). The mobile or wireless network provides internet connectivity to the users. Therefore, the users can access cloud based services by Internet, if they have mobile devices that support network connectivity.

Mobile users' requests and users' profiles are transmitted to the central processors that are connected to servers providing mobile and wireless network services. Mobile and wireless network operators can provide services Authentication, Authorization and Accounting (AAA) for mobile users based on the Home Agent (HA) and subscribers' data stored in databases. After that, the subscribers' requests are delivered to cloud through Internet. In the cloud, the cloud controllers process the users' requests and provide the corresponding services from the Cloud Computing Service Provider to mobile users.

Recently mobile applications have begun to adapt to cloud computing environment. However, these applications are often linked with server instances running in the cloud. Because of this, the MCC users may face some problems such as congestion due to the limited bandwidth, network disconnection, and the signal attenuation caused by mobile users' mobility. This would cause delays when MCC users want to communicate with the cloud, so QoS and QoE are significantly reduced.

CloneClouds and Cloudlets are some of the possible solutions which would reduce the network delay [17]. *CloneCloud* uses nearby computers, or data centers to increase the speed of running smart phone applications, by cloning the entire set of data and applications from the smartphone onto the cloud and selectively executing some operations on the clones, reintegrating the results back into the smartphone. *Cloudlet* is a trusted, resource-rich computer, or cluster of computers that has good Internet connection and it is available for use by nearby mobile devices. The MCC users may use a cloudlet if it is available, and if they do not want to offload the content and the request to the cloud (maybe due to the delay, cost, etc).

4.3 Cloud Computing Service Oriented Architecture

Once a cloud is established, how its cloud computing services are deployed in terms of business models can differ depending on requirements [17], [67]. A layered architecture of cloud computing that demonstrates the effectiveness in terms of meeting the user's requirements is given on Figure 3.

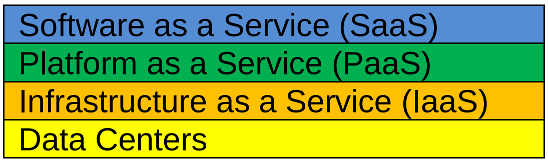


Figure 4.3 Cloud Computing Service Oriented Architecture

The layer of **Data centers** provides the hardware facility and infrastructure for clouds. A number of servers are linked with high-speed networks to provide services for customers. Usually, data centers are built in less populated places, with high power supply stability, and a low risk of disaster.

The **Infrastructure as a Service (IaaS)** layer is built on top of Data centers layer. The consumer is not able to manage the underlying cloud infrastructure. However, IaaS provides provision processing, storage, networks, and other fundamental computing resources where the consumer is able to deploy and run arbitrary software, which can include operating systems and applications.

The **Platform as a Service (PaaS)** layer provides an advanced integrated environment for building, testing and deploying custom applications. The consumer does not control the underlying cloud infrastructure such as network, servers, operating systems, or storage, but has control over the deployed applications and possibly application hosting environment configurations.

The **Software as a Service (SaaS)** layer provides applications running on a cloud infrastructure. These applications can be accessed by various client devices by using a thin client interface such as a web browser. The consumer is not able to manage, or control the underlying cloud infrastructure including network, servers, operating systems, storage, or even individual application capabilities. However, the consumer may control and manage only limited user-specific application configuration settings.

Also, there are subsets of these service layers that relate to a particular industry or market. Such models are for example Network as a Service and Communication as a Service [64], [68].

Network as a Service (NaaS) is a category of the cloud computing services where the user has an opportunity to use the transport connection services, and/or network connection services between the clouds. NaaS services include flexible and extended VPN, bandwidth on demand, etc.

Communications as a Service (CaaS) is a category of cloud computing services where the user has an opportunity to use real time communication and collaborative services. CaaS services include VoIP, instant messaging service and video conference.

4.4 Cloud Computing Deployment Models

Depending on the requirements different deployment models for (mobile) cloud computing exist. The deployment models can be: private cloud, community cloud, public cloud, hybrid cloud and inter-cloud computing [67 – 69].

The **Private Cloud,** also known as **Corporate Cloud**, or **Internal Cloud** is a proprietary computing architecture for an organization that provides hosted services on private networks. The deployment, the maintenance, and the operations of the cloud infrastructure are performed by the organization itself. The operation may be in-house, or with a third party on the premises. However, this model has its own disadvantages, since the organizations still need to purchase, set up, and manage their own clouds.

The **Community Cloud** is a private cloud infrastructure that is shared among a number of organizations with similar interests and requirements. This may help limit the capital expenditure costs for its establishment as the costs are shared among the organizations. The operation may be in-house, or with a third party on the premises.

The **Public Cloud** or **External Cloud** infrastructure is available to the public on a commercial basis by a cloud service provider. This enables a consumer to develop and deploy a service in the cloud with very little financial outlay compared to the capital expenditure requirements normally associated with other deployment options. The users are connected to the cloud data centers that provide the cloud services via the public Internet.

The **Hybrid Cloud** can be a combination of private and public clouds that support the requirement to retain some data within the organization, and also the need to offer services in the cloud. The clouds in this model have the ability through their interfaces to allow data and/or applications to be moved from one cloud to another. The hybrid cloud offers a suitable environment for the needs of the enterprises, but it also introduces the complexity of determining which services and applications should be distributed across the private, public, or both clouds. In this model the Cloud Management System (CMS) is responsible for the administration of hybrid clouds. The CMS should contain some functionality such as security management, resource scheduling (e.g., immediate, on-demand, or for later use), resource allocator, and monitoring the university activities and performance. This information can be used to determine the required resources that need to be allocated in the future.

The **Inter-cloud computing** is a model that enables on demand to provide cloud computing resources, such as processing, storage and network, as well as, a distribution of the workload through the internetworking of the clouds.

5 Fog Computing

The future IoE would become the linkage between extremely complex networked organizations (e.g. telecoms, transportation, financial, health and government services, commodities, etc.), which would provide the basic ICT infrastructure that supports the business processes and the activities of the whole society in general [28], [61]. Frequently, these processes and activities should be supported by orchestrated cloud services, where a number of services work together to achieve a business objective [13].

Although mobile cloud computing is a promising solution for 5G to cope with the future Internet, still it cannot deal with all future Internet services and applications. This is because the future Internet would exacerbate the need for improved QoS/QoE, supported by services that are orchestrated on-demand and are capable to adapt at runtime, depending on the contextual conditions, to allow reduced latency, high mobility, high scalability, and real-time execution. The emerging wave of Internet of Things (IoTs) would require seamless mobility support and geo-distribution in addition to location awareness and low latency. These demands can only be partially fulfilled by existing cloud computing solutions [10]. Also, the existing cloud computing security mechanisms such as sophisticated access control and encryption have not been able to prevent unauthorized and illegitimate access to data.

Recently a new paradigm called **Fog Computing**, or briefly **Fog** has emerged to meet these requirements [24], [70]. Fog Computing extends cloud computing and services to the edge of the network. Fog would combine the study of mobile communications, micro-clouds, distributed systems, and consumer big data. It is a scenario where a huge number of heterogeneous (wireless and sometimes autonomous) ubiquitous and decentralized devices communicate, and potentially cooperate among them and with the network to perform storage and processing tasks without the intervention of third parties [25]. These tasks support basic network functions, or new services and applications that run in a sand-boxed environment. Users leasing part of their devices to host these services get incentives for doing so. The distinguishing Fog characteristics are its proximity to end-users, its dense geographical distribution, and its support for mobility. Therefore, Fog paradigm is well positioned for real-time big data analytics. Services are hosted at the network edge or even end devices such as set-top-boxes, or access points [26]. By doing so, fog reduces service latency, and improves QoS, resulting in superior user-experience. It supports emerging IoE applications that demand real-time/predictable latency (industrial automation, transportation, networks of sensors and actuators).

5.1 Fog Computing Architecture

The fog computing architecture uses one, or a collaborative multitude of end-user clients, or near-user edge devices to carry out a substantial amount of storage, communication and management [30]. An overview of Fog Computing architecture is given in Figure 5.1. It consists of centralized cloud computing center, IP/MPLS core network, RAN network with distributed fog computing intelligence and smart things network. Each smart thing device is attached to one of fog devices in the RAN network. The fog devices could be interconnected to each other, and each of them is linked to the centralized cloud computing center via the IP/MPLS core the network.

The RAN network is actually an intermediate fog layer that consists of geo-distributed intelligent fog computing servers which are deployed at the edge of networks, e.g., parks, bus terminals, shopping centers, etc. Each fog server is a highly virtualized computing system and is equipped with the on-board large volume data storage, computing, and wireless communication facility [26].

The role of fog servers is to bridge the smart mobile device things and the cloud [25]. The geo-distributed intelligent fog servers directly communicate with the mobile users through single-hop wireless connections using the off-the-shelf wireless interfaces, such as, LTE, WiFi, Bluetooth, etc. They can independently provide pre-defined service applications to mobile users without assistances from cloud, or Internet. In addition, the fog servers are connected to the cloud in order to leverage the rich functions and application tools of the cloud.

Fog computing and networking contains both data plane and control plane that enable different applications with different communication protocols over all layers in the OSI system [30]. This is illustrated in Figure 5.2. Fog data plane is focused on 5G mobile network, IoT, and the future IoE. Fog control plane is mainly about cyber physical system control and real-time data analytics.

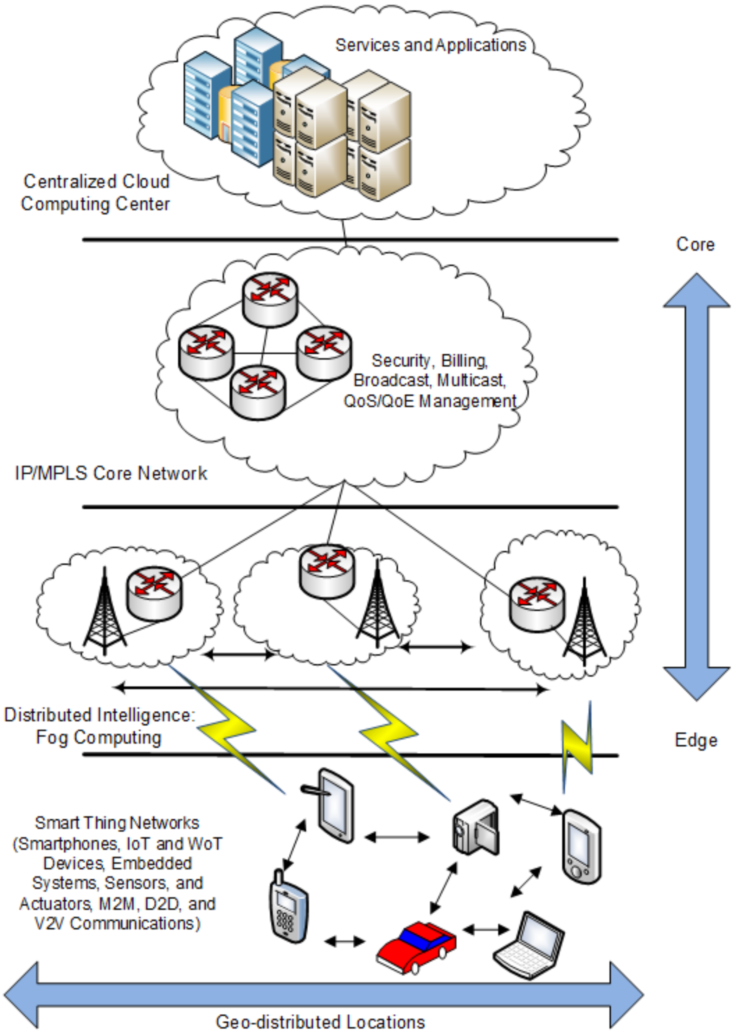


Figure 5.1 Fog Computing Architecture [72]

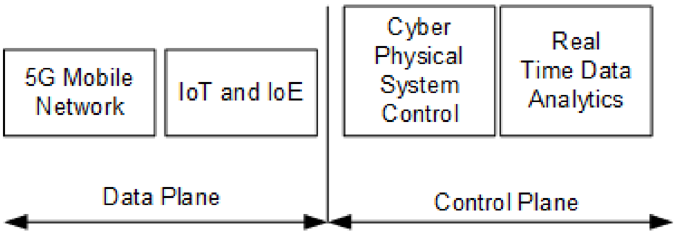


Figure 5.2 Data and Control Plane of Fog Computing [30]

5.2 A Comparison between Cloud and Fog Computing

The existence of Fog would be enabled by the emerging trends on technology usage patterns on one side, and the advances on enabling technologies on other side. A comparison between Fog Computing and Cloud Computing is given in [26] which is summarized in Table 5.1.

Cloud and fog form a mutually beneficial, inter-dependent continuum [30]. They are inter-dependent, e.g., coordination among devices in a Fog may rely on the Cloud. They are also mutually beneficial: certain functions are naturally more advantageous to carry out in Fog while others in Cloud.

Some of the main fog computing features are described in the following sections.

Table 5.1 A Comparison between Cloud Computing and Fog Computing

	Cloud Computing	**Fog Computing**
Target Type	General Internet Users	Mobile Users
Service Type	Global Information Collected from Worldwide	Limited localized information services related to specific deployment locations
Service Location	Within the Internet	At the edge of the local network
Distance between client and server	Multiple hops	Single hop
Number of server nodes	Few	Very large
Latency	High	Low
Delay jitter	High	Very low
Geo-distribution	Centralized	Distributed
Security	Undefined	Can be defined
Hardware	Ample and scalable storage, processing and computing power	Limited storage, processing and computing power, wireless interface
Deployment	Centralized and maintained by OTT service providers	Distributed in regional areas and maintained by local businesses

5.3 Ubiquity of Devices

The main factor which would bring fog into reality is the ubiquity of the devices, whose increase is driven by the user devices and sensors/actuators, and enabled by the presence and usage of devices everywhere around us for different services and applications. By decreasing the size of the device, the device portability is increased. However, the power consumption is also reduced, which may be crucial in some context application. This can be solved with packaging and power management technologies such as System on Chip and System in Package Technologies, 3D Microbatteries, RF powered Computing, etc. that aim to create smaller and more autonomous mobile devices which would run longer at a minimum price [25].

5.4 Network Management

The configuration and the maintenance of many different types of services running on many heterogeneous devices would only exacerbate the current management problems. Therefore, in fog computing environment heterogeneous devices and their running services need to be managed in a more homogeneous manner with the following technologies: Network Function Virtualization (NFV); small edge clouds to host services close to, or at the endpoints; and peer-to-peer (P2P)- and sensor network-like approaches for auto-coordination of applications [25].

NFV is the reaction of telecom operators to their lack of agility and constant need for reliable infrastructures. It is capable of dynamically deployment for on-demand network services, or user-services. Software Defined Networks (SDNs) are one of the main enablers for NFV. For example, the router can be seen as an SDN-enabled virtual infrastructure where NFV and application services are deployed close to the place where they are actually going to be used, which would result in cheaper and more agile operations. However, NFV capabilities still do not reach end user devices and sensors.

Telecom operators had already started to deploy clouds in their Long-Term Evolution (LTE) Networks closer to the edge (to the user), because the Evolved Packet Core (EPC) more efficiently delivers services close to users (at the edge) and confines traffic there while reducing the traffic overload with the help of SDNs. The fog would enable the devices to become virtual platform, that can lease some computing/storage capacity for applications to run on them.

The peer-to-peer (P2P) and sensor network-like approaches exploit the locality and allow endpoints to cooperate in order to achieve similar results, but can scale better, and can be implemented in a fog. A fog application can be seen as a Content Distribution Network (CDN) where a data is exchanged between peers. As a result, the applications and data are no longer required to stay in centralized data centers.

A part of a network and some user devices/sensor can act as mini-clouds in a fog computing environment. The mini clouds can be implemented by using droplets, or small pieces of code that run in a secured manner in devices at the edge with minimum interaction with central coordinating elements, and thus reducing the unnecessary and undesired uploads of data to central servers in corporate data centers. The users are able to retain control and ownership of their own data and applications, and the scalability is improved.

5.5 Fog Connectivity

The presence of many mobile devices that consume and produce big data at the edge of the network may cause a huge bottleneck in the fog [25]. On physical level the following technologies can cope with this: LTE-Advanced, WiFi ac, Bluetooth Low Energy, ZigBee, etc.

On the network level, each node must be able to act as a router for its neighbours and must be resilient to node churn (nodes entering and leaving the network) and mobility. Mobile Ad-hoc Networks (MANETs) and Wireless Mesh Networks (WMNs) can provide these functionalities [25]. MANET would enable the formation of densely populated networks without requiring fixed and costly infrastructures to be available beforehand.

WMNs on the other side use mesh routers at its core. Nodes can use the mesh routers to get connectivity, or other nodes if no direct link with the routers can be established. Routers would grant access to other networks such as cellular, Wi-Fi, etc.

On higher levels some protocols already exist for IoT, such as Message Queue Telemetry Transport (MQTT) and Constrained Application Protocol (CoAP), that provide low resource consumption and resilience to failure [25]. The network and IoT protocols can benefit from data locality, since they no longer need to send all the data around the world all the time, except for potential congestion problems at the edge of the network. Data locality has also a very positive impact on privacy.

5.6 Fog Privacy and Security

The greatest concern of fog users is data ownership, i.e. data security and privacy [27], [80]. One method to maintain the privacy is by storing encrypted sensitive data in traditional clouds.

The existing cloud computing data protection mechanisms such as encryption have failed in preventing data theft attacks, especially those perpetrated by a malicious insider.

However, the value of stolen information can be decreased. This can be achieved through a preventive disinformation attack, by using the following additional security features: user behaviour profiling, decoys and combination of both [80].

User behaviour profiling is used to model the normal user behaviour, i.e. how, when, and how much a user accesses his/her information in the cloud [80]. Such profiles contain volumetric information, i.e. how many files are accessed and how often. The occurrence of abnormal access to a user information in the cloud can be determined by monitoring this normal user behaviour, based partially upon the scale and the scope of data being transferred.

Decoys are any bogus information that can be generated on demand. They are used to:

- validate whether data access is authorized when abnormal information access is detected; and
- confuse the attacker with bogus information [80].

The serving decoys would confuse the malicious attacker into believing that he/she has exfiltrated useful information, but he/she has not. The attacks can be prevented by deploying decoys within the fog by the service customer and within personal on-line social networking profiles by the individual users.

A combination of decoys with user behaviour profiling would provide unprecedented levels of security for the fog, and would improve detection accuracy. When the access to user information is correctly identified as an unauthorized access, the fog security system would deliver unbounded amounts of decoy information to the attacker. Thus, the true user data is protected from unauthorized disclosure. When abnormal access to the fog service is not recognized, decoy information may be returned by the fog and delivered in such a way as to appear completely legitimate and normal. The true owner of the information, would identify when decoys are returned by the fog. Hence the legitimate user could alter the fog responses through a variety of means, such as challenge questions, to inform the fog security system that has inaccurately detected an unauthorized access. At the moment the existing security mechanisms do not provide this level of security [80].

6 The Cloud in 5G Mobile Networks

In general, the big Over The Top (OTT) players have a leading role in the cloud computing market. Dropbox, Apple, and Google are the main personal cloud providers, while Salesforce.com, Google, and Microsoft the leaders in SaaS market. Amazon is mostly used for the PaaS and IaaS services.

One part of the cloud computing market would also belong to the Telecom operators. Compared to OTT players, the telco operators have recently entered in the cloud computing market [68].

6.1 Several Reasons that Telecom Operators Should Use the Cloud

There are two main reasons why telecom operators should include the cloud computing. The first reason is to use the benefits of cloud computing for IT optimization (lower costs, bigger elasticity and speed). The second reason is to use the new business opportunities.

The telco operators have a competitive advantage on a local level, where they can benefit from the local commercial presence, compared to many OTT players.

Due to their relative close proximity to the users, the telecom operators may offer cloud services with very low latency. Additionally, they can offer services to the users that are flexible and elastic in terms of the frequency bandwidth. Finally, the users have more trust and confidence in telco operators rather than in many Internet brands, or OTT players.

6.2 The Cloud in Telecom Operators

Telecom operators are investigating how they can make the best use of their assets, and how future network investment can be aligned with a cloud model [81]. One option is to make use of network equipment to host server modules to create telecom-grade clouds. This could include a mix of centralized cloud and a distributed cloud using access network and RAN elements, as it is shown on Figure 6.1.

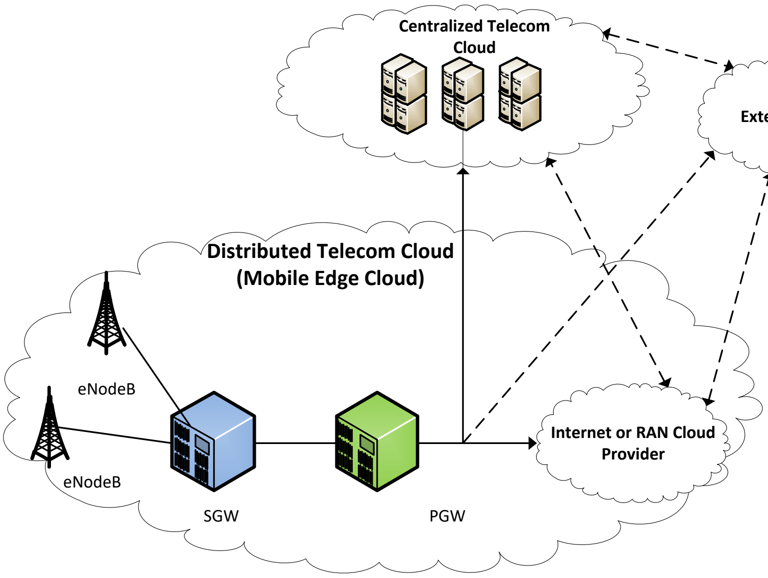


Figure 6.1 Cloud Integration with Telecommunication Network [20]

Currently hosting content and applications in the RAN has become popular. By placing storage and computing resources at, or close to, the cell site, operators can improve response times for the services requested by the users in the prevailing radio conditions. This might be useful for congestion control, or rate adaptation for video streams. On the network side, cell site caching can reduce demand on the backhaul network, and potentially play a role in limiting signaling to the core network. The two primary advantages of placing content close to the radio and close to the user are application performance and network efficiency. One emerging application where both application performance and network efficiency could work together in a distributed cloud model is LTE Broadcast that uses evolved Multimedia Broadcast Multicast (eMBMS) technology.

6.3 The Conventional Cloud Computing in 5G Mobile Networks

The 5G network architecture in a mobile cloud computing environment [82] is given on Figure 6.3, and is based on the architectural levels explained in Chapter 2.4.

The top of the architecture consists of innovative and content service providers that provide Over-The-Top (OTT) applications and services such as augmented and virtual reality, hologram, mobile ultra-high definition, cloud computing, etc.

The software based single core is enhanced Evolved Packet Core (EPC enhanced) that provides the core functionalities in the network such as billing and charging, mobility and handover, location and geo-management, home subscriber server, Authorization Authentication and Accounting (AAA) security, and Network as a Service (NaaS) that allow configuration of all telecommunication and service functions with virtualized software on a programmable hardware.

The 5G RAN consists of consists of EDGE Cloud RAN (EDGE CRAN) service and Hyper-connected RAN [82].

Because the smart mobile devices have limited capabilities for storing and processing of data, this is solved by moving the storing and processing of data from the smart mobile device to the cloud computing centers. However, this requires high bandwidth and low latency.

In order to solve this issue, the EDGE CRAN moves the cloud computing functionalities to the radio access network [83 – 84]. However, the application storing and all radio signal processing functions are still centralized at the cloud computing server in 5G core. Because billions of smart user devices need to transmit and exchange their data fast enough with the Base Band Unit (BBU) pool, there is a requirement for high bandwidth and low latency.

To overcome this, Heterogeneous Cloud Radio Access Networks (H-CRANs) have been proposed in which the user and control planes are decoupled [83], [85], as it is shown in Figure 6.3. The centralized control function is shifted from the BBU pool in CRANs to the High-Power Nodes (HPNs) in HCRANs. HPNs are also used to provide seamless coverage and execute the functions of control plane. The high-speed data packet transmission in the user plane is enabled with the radio heads (RRHs). HPNs are connected to the BBU pool via the backhaul links for interference coordination.

However, HCRANs still have its own drawbacks. The data traffic data over the fronthaul between RRHs and the centralized BBU pool surges a lot of redundant information, which worsens the fronthaul constraints. In addition, HCRANs do not take fully utilize the processing and storage capabilities in edge devices, such as RRHs and smart user devices, which is a promising approach to successfully alleviate the burden of the fronthaul and BBU pool. Finally, operators must deploy a huge number of fixed RRHs and HPNs in H-CRANs in order to meet the requirements of peak capacity, which makes a serious waste when the volume of delivery traffic is not sufficiently large.

The Hyper-connected RAN is a multi-layered heterogeneous network that consists of existing 2G, 3G, 4G and future Radio Access Technologies (RATs) that provides cellular content caching, radio and resource allocation, mobility, interference control, etc.

Finally, the smart things network consists of smart mobile, Internet of Things (IoT) and Web of Things (WoT) end user devices.

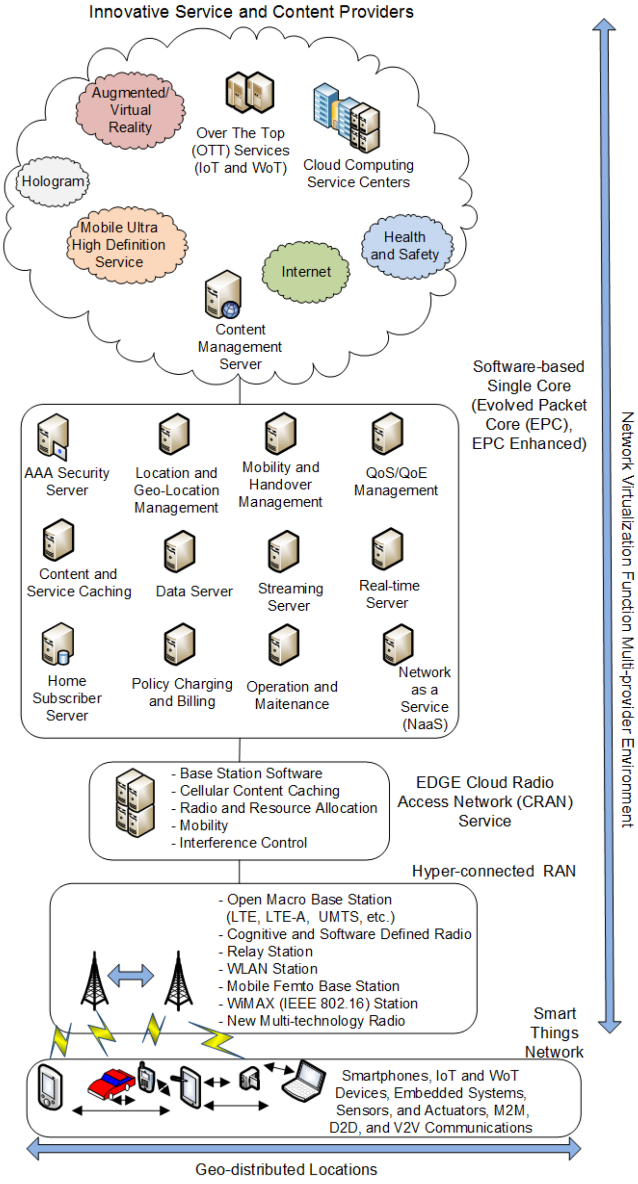


Figure 6.2 5G Network Architecture in a Mobile Cloud Computing Environment [82]

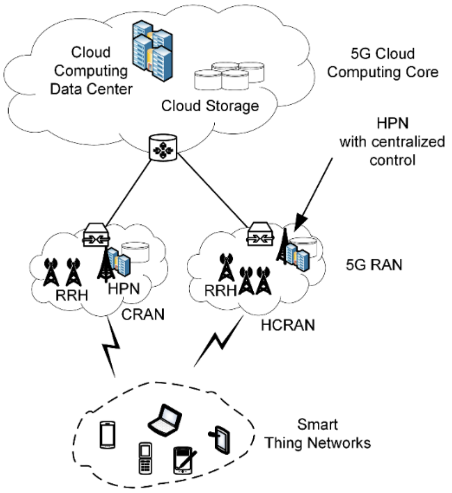


Figure 6.3 5G Network with CRAN and HCRAN

Full Network Function Virtualization (NFV) would take place in 5G, in order satisfy the service requirements. Network virtualization pools the underlying physical resources or logical elements in a network, by using the current technologies such as cognitive and software defined radios in 5G RAN, software defined networking and cloud services in 5G core [86].

The NFV functions should cover the control and management of QoS, the service policy and prioritization of traffic. The network functions like signal processing and path computation would be virtualized and offloaded to network management clouds. This would make network operating easier, reduce end-to-end energy consumption and open the way to new flexible communication services. The main cloud computing functionality in 5G core would be Network as a Service (NaaS) platform which would allow configuration of all telecommunication and service functions with virtualized software on a programmable hardware [82].

5G would include a mix of centralized cloud and a distributed cloud using access network and RAN elements [82]. By moving the cloud in radio access network, known as Cloud RAN (CRAN), i.e. by moving the base station from the cell site into cloud, i.e. by placing storage and computing resources at, or close to, the cell site, operators would guarantee the necessary Service Level Agreement (SLA) and can better support delay-sensitive applications such as virtual desktops or electronic programming guides. They would also improve response times, making services feel “snappier” and, uniquely, more responsive to prevailing radio conditions. This might be useful for congestion control, or rate adaptation for video streams. On the network side, cell site caching can reduce demand on the backhaul network, and potentially play a role in limiting signalling to the core network.

Because 5G RAN would consist of a dense deployment of micro, pico and femto cells, mobile cloud computing would significantly reduce the latency over the wireless communication channel, as well as the transmit power necessary for computation offloading. Moreover, mobile device can easily locate a cloud access point. Millimetre wave links, massive MIMO, and multi-cell cooperation can be used to improve the spectral efficiency, by reducing the time necessary to transfer the users’ offloading requests to the cloud [58].

Mobile cloud computing in 5G network should support context-rich support services such as context extraction service, recommendation service and group privacy service [20], [82], [87] where a layer of cloud recommendation services can be built that creates output which is adjusted to a user, or set of users with those contextual characteristics.

6.4 Mobile Edge Computing in 5G Mobile Networks

Recently the mobile devices became an important tool for learning, entertainment, social networking, and acquiring new information from the news and the business [88]. However, due to the constrained resources of the mobile devices (the processing power, battery life, and the capacity for data storing) the mobile users do not receive the same quality as the conventional desktop users [17]. With the appeareance of mobile cloud computing, the limited resources in the mobile devices of data storage and data processing was resolved with the transfer of these resources to the cloud. Many cloud computing services such as m-health-care [89 – 90], mobile learning (m-learning) [91], mobile video games (m-gaming) [92] and m-governance [93] are already directly available from the mobile devices [94]

However, because the data need to be transferred from the mobile devise to the cloud computing centers, there is a need for increased bandwidth, and lower latency, which causes a network overload to occur. It is expected the need for the frequency bandwidth to be doubled every year [35], [96]. Moreover, the Internet of Things and the future Internet of Everything, would provide the devices with constrained resources to be connected on Internet [28], [96].

In order to resolve this issue, one possible proposed solution is the Cloudlet, where the mobile devices transmit data via a WiFi network for further processing to a server with less resources than the cloud, but is nearby to the mobile devices [97]. However, this approach is less effective due to the following two reasons. Firstly, the access to the Cloudlet is only possible through a WiFi access point. And secondly the Cloudlet possess less resources than the cloud, which means it is not scalable in resource and service provisioning.

Therefore, a better proposed solution is the Mobile Edge Computing (MEC) [98 – 99]. MEC is a model that enables business oriented cloud platform to be implemented within the radio access network, close the mobile users in order to serve applications that are delay sensitive and context aware.

This approach is initiated by ETSI, where data processing and storage happens at the edge in the radio access network in the base station, rather than in cloud computing centers, in order to create new application and service possibilities and opportunities. In this way, the mobile edge computing manages to reduce latency, alleviates the network congestion in the core, enables efficient traffic management, and opens possibility other industrial services to be delivered for the critical applications through the mobile network. This approach provides innovative network architecture where the possibilities for cloud computing and IT services are converged with the mobile network.

The MEC architecture is given on Figure 6.4. The mobile edge computing entity is positioned next to the radio access network. This entity works with downlink bitstreams from the cloud computing servers to the mobile device and uplink bitstreams from the mobile device to the cloud computing servers. The MEC platform contains standard IT servers and network devices inside, or outside of the base station [100]. The external applications are executed in the Virtual Machines (VMs) [101] connected with network devices. Also, there is a possibility the MEC platform to be implemented with standard IT servers, where the network device is implemented as a software entity, such as Open Virtual Switch, or Open vSwtch (OVS) [102].

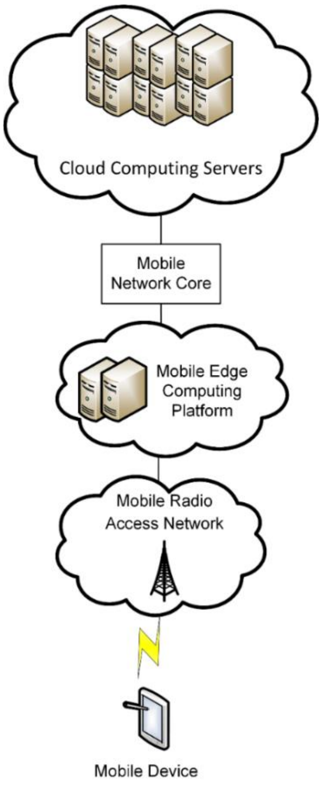


Figure 6.4 Mobile Edge Computing Architecture

The MEC platform contains the following fundamental modular functionalities: routing, exposure of network capabilities and management. The routing module is responsible for forwarding the packets between the MEC platform and the radio access network and the core part of the mobile network, as well as within the MEC platform itself. The module network exposure provides authorized information to the Radio Network Information Service (RNIS) and the Radio Resource Management (RRM). The management module supports the AAA security function, as well as the management of the external applications within the MEC platform itself. Here is included the orchestration of setting up the application and the authorization for the exposure of the network capabilities [100].

MEC enables the mobile network to evolve in an innovative way to handle the increased traffic volume, that arise from various domains in 5G era, which are grouped into enhanced Mobile BroadBand (eMBB) [103], massive Machine Type Communication (mMTC) [104] and Ultra-Reliable and Low Latency Communication (URLLC) [105].

6.5 Mobile Edge Computing vs Fog Computing in 5G Mobile Networks

Although the Mobile Edge Computing provides many benefits, still it has certain constraints and limits compared to the fog computing. In order to see these constraints and limits it is necessary to make a comparison between these two approaches.

Table 6.1 The Main Features of Mobile Edge Computing and Fog Computing [106]

	Mobile Edge Computing	**Fog Computing**
Application hosting	Limited	Yes
Data Service at Edge	Yes	Yes
Device and Application Management	Yes	Yes
Security and Safety	Partially solved with VPN or FW	A Complete E2E solution, data protection, on session level and hardware
Elastic compute / resource pooling	No	Yes
Modular Hardware	No	Yes
Virtualization with Windows support	To be defined	Yes
Real time control high level availability	No	Yes

Table 6.1 summarizes the main characteristics and features of mobile edge computing and fog computing [106].

The MEC is developed by ETSI and the vendors for telecommunication equipment in order to deliver standardized MEC architecture and industrial standardized application programming interface for external applications. This concept represents a control and management of independent end user device individually, or through a set of software functions in a cloud domain in the radio access network. The devices and the entities that perform computing at the edge of the network in the cloud domain are either independent or mutually connected through proprietary networks with usual security and low degree of interoperability. Recently this approach attempts to redefine the data computing range at the edge by including some functionalities for fog computing, such as interoperability, local security and safety, etc. However, this is not extended to the cloud or through the domains [106].

On the other hand fog computing is a concept proposed by CISCO and other manufacturers for Internet equipment and it represents a completely distributed network, multi-layered cloud computing architecture for data processing, where exists billions of devices as part of IoT, multiple local clouds at the edge of the network – fog and main central hyper scalable cloud computing data center. A single application in the fog is distributed through the devices through the cloud components embedded in the nodes in the different network levels, for example in the radio access network, multi-service edge, the core of the network (in the IP/MPLS routers and switches, the gateways of the mobile packet core, etc.). In this way the cloud is closer to the users of mobile devices and is capable to offer ultra-low latency, quicker response, high bandwidth, as well as real time access to the radio information that will be used by the applications in order to offer context related services.

Table 6.2 Differences between mobile edge computing and fog computing

No	Mobile Edge Computing	Fog Computing
1	Device aware and few services aware, unaware of the entire domain	Device independent, intelligent, and aware of the entire fog domain
2	Limited control in the edge domain	Controls all devices in the domain
3	Not aware for the cloud	Extends the cloud to a fog level in a continuum
4	Limited network scope	Complete network scope
5	No IoT vertical awareness	Support and enabler for multiple IoT verticals
6	No IoT vertical integration	Integrates multiple verticals
7	Uses Edge Controllers that are focused on edge device command and control	Uses fog nodes that are very versatile and capable of performing a variety of functions like Real time Control, application hosting and management.
8	Security scope is limited to devices	End-to-End security through data ownership
9	Not designed with virtualization	Enables rich virtualization
10	Analytics scoped to a single device	Fog Analytics enables collection, processing and analysis of data from multiple devices in the edge for analysis, machine learning, anomaly detection and system optimization.
11	Edge Computing typically is embedded in and controls the edge. Certain devices require hard real time control and others require non real time control, and the Edge Computing performs these functions.	Fog Computing uses the devices and the embedded edge control. However, Fog Computing also enables virtual machines that host soft PLC used in real time control.

Compared to Mobile Edge Computing, the edge devices in the Fog Computing may independently self-optimize among themselves and collectively to perform measurement and management to the remaining of the network. The fog computing moves the operational functions of the digital objects in Information Centric Networks (ICNs) and virtualized functions in the Software Defined Networks (SDNs) at the edge of the network [30].

Table 6.2 provides the main differences between mobile edge computing and fog computing. Although these two concepts are different from each other, still they share a similar vision, which is guided by the future predictions that will be characterized with the Internet of Things, Internet of Everything, Tactile Internet and the existence of appropriate mobile and wireless connectivity.

It can be concluded that fog computing is a better solution to be applied in 5G mobile networks. However better results can be achieved if both concepts are combined [107].

6.6 Several Reasons for Fog Computing to be applied in 5G Networks

There are four main reasons why fog computing is a suitable solution for 5G [30]:

1. **Real time processing and cyber-physical system control**. Edge data analytics, as well as the actions it enables through control loops, often have stringent time requirement in the order of few milliseconds that can be carried out only at the edge of the network. This is particularly essential for Tactile Internet, that enables virtual-reality-type interfaces between humans and devices.
2. **Cognition or awareness of client-centric objectives**. The applications can be enabled by knowing the requirements and the preferences of the clients. This is particularly true when privacy and reliability cannot be trusted in the Cloud, or when security is enhanced by shortening the extent over which the communication is carried out.
3. **Increased efficiency by pooling of idle and unused local resources**. The idle and unused gigabytes on many devices, the idle processing power, the sensing ability and the wireless connectivity within the edge may be pooled within a fog network.
4. **Agility, or rapid innovation and affordable scaling**. It is usually much faster and cheaper to experiment with client and edge devices. Rather than waiting for vendors of large boxes inside the network to adopt an innovation, in the fog computing world a small team may take advantages of smart phone Application Programming Interface (API) and Software Development Kit (SDK), proliferation of mobile applications, and offer a networking service through its own API.

6.7 Possible Scenarios of Applying Fog Computing in 5G Mobile Networks

Below are described some scenario cases where fog computing can be applied in 5G [30].

Case 1: *Crowd-sensing the states in 5G base station*. Through a collection of 5G end user client devices may be able to infer the states of a 5G base station such as the number of Resource Blocks used, by using a combination of passive received signal strength measurement (e.g., RSRQ), active probing (e.g., packet train), application throughput correlation and historical data mining [78].

Case 2: *Over-The-Top (OTT) network provisioning and content management*. The traditional approach to innovate in the networks is to introduce another box inside the network. The fog directly leverages the “things” and phones instead, and removes the dependence on boxes-in-the-network altogether. By using end user client devices for the tasks such as Universal Resource Locator (URL) wrapping, content tagging, location tracking, behaviour monitoring, network services can be innovated much faster [30].

Case 3: *Network selection in a heterogeneous environment*. The coexistence of heterogeneous networks (e.g., LTE, femto, WiFi) is a key feature in 5G. Instead of network operator control, 5G would have a user-centric approach where each client can observe its local conditions and make decision on which network to join [5]. Through randomization and hysteresis, such local actions may emerge globally to converge to a desirable configuration [76].

Case 4: *Borrowing bandwidth from neighbors in M2M, D2D or V2V communications*. When multiple devices are next to each other, one device may request the other devices to share their bandwidth by downloading other parts of the same file and transmitting, via WiFi Direct, client to client, etc. [73].

Case 5: *Distributed beam-forming*. Fog can also be applied in the physical layer, by exploiting multi-user MIMO to improve throughput and reliability when a client can communicate with multiple access points, or base stations. For uplink, multi-user beam-forming can be used so that the client can send multiple data streams to multiple Access Points (APs), or Base Stations (BSs) simultaneously. For downlink, interference nulling can be used in order the client to be able to decode parallel packets from multiple APs or BSs. These can be done entirely on the client side [75].

As a conclusion the cloud in 5G network would be diffused among the client devices often with mobility too, i.e. the cloud would become fog [56]. More and more virtual network functionality would be executed in a fog computing environment, which would provide ***mobiquitous*** service to the users. This would enable new services paradigms such as Anything as a Service (AaaS) where devices, terminals, machines, and also smart things and robots would become innovative tools that would produce and use applications, services and data.

Part Two

7 Fog Computing Service Orchestration Mechanisms in 5G Mobile Networks

5G Network with its cloud computing mechanisms cannot deal with the future Internet services and applications. This is because the traditional service orchestration approaches that have been applied to cloud services are not adequate to the forthcoming large-scale and dynamic fog services, since they cannot effectively cope with reduced latency, high mobility, high scalability, and real-time execution [13]. In addition, in the previous chapter it was pointed out the importance of the applying the fog computing environment in 5G mobile network. Therefore, it is time to think of a new approach so called 5G network in the fog.

7.1 Hybrid Environment Service Orchestrator

In order fog computing environment to become a reality in 5G mobile networks a new **Hybrid Environment Services Orchestrator (HESO)** is needed, that would ensure resilience and trustworthiness of open, large scale, dynamic services [56], [108]. The HESO Orchestrator would be responsible for the composition of service elements available in the fog environment (e.g. sensing, connectivity, storage, processing, platform services, and software services) into more complex fog services (e.g. traffic crowd sensing and trip planning services) that can be offered to the users. The execution of the fog services may involve multiple different components and entities spread in a wide area, increasing the complexity in terms of decision making process in what regards the resource allocation to achieve acceptable QoS/QoE levels. To coordinate the execution of the fog services, the orchestration mechanisms need to synchronize and combine the operation of the different service elements in order to meet the specifications of the composed fog services, including low latency, scalability and resilience.

The HESO in Fog should operate in a loosely coupled mode, resulting in a solution with several levels: Regional Service Orchestrator (RSO), Domain Service Orchestrator (DSO), and Federated Service Orchestrator (FSO), as it is shown in Figure 7.1.

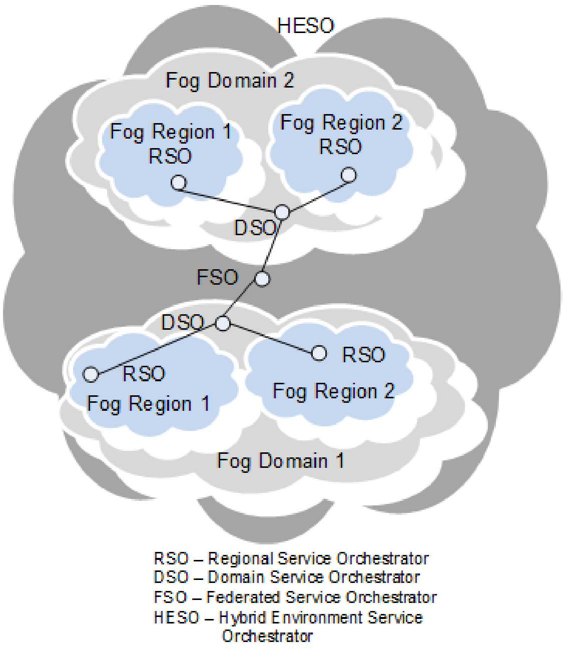


Figure 7.1 Fog Computing Hybrid Environment Service Orchestration Model in 5G Mobile Networks [54]

The RSOs are located at the edges of the Fog environment and enable semi-autonomous operation of different Fog Regions. This allows the distribution of the load which provides scalability and much higher proximity to the end users with lower latencies.

The DSOs is responsible for the Fog domains and supervises the RSOs below. This level supports mechanisms that enable intra-domain cooperation between different regions.

The FSO allows a fruitful interaction between different Fog domains. It is responsible for the management between different Fog domains and, similarly to the DSOs, it should be properly adapted to operate in a federate Cloud environment. The FSOs support federation mechanisms to enable cooperation among different Fog Domains (e.g. belonging to different entities or under the administration of different authorities) and the creation of a Multi-Domain Fog Environment able to support service ubiquity.

HESO model is flexible and scalable and can be implemented in any network technology standard. In particular, its application is important for critical usage cases of IoT devices and Tactile Internet [109] that requires 1 ms end-to-end latency in order to provide virtual-reality-type interfaces between humans and machines, and big data analytics that requires real time processing with stringent time requirement that can only be carried out in the fog.

7.2 5G Network Architecture in a Fog Computing Environment

On the basis of the fog computing hybrid environment service orchestration model in 5G mobile networks, Figure 7.2 presents the 5G mobile network architecture in a fog computing environment. The architecture consists of several levels.

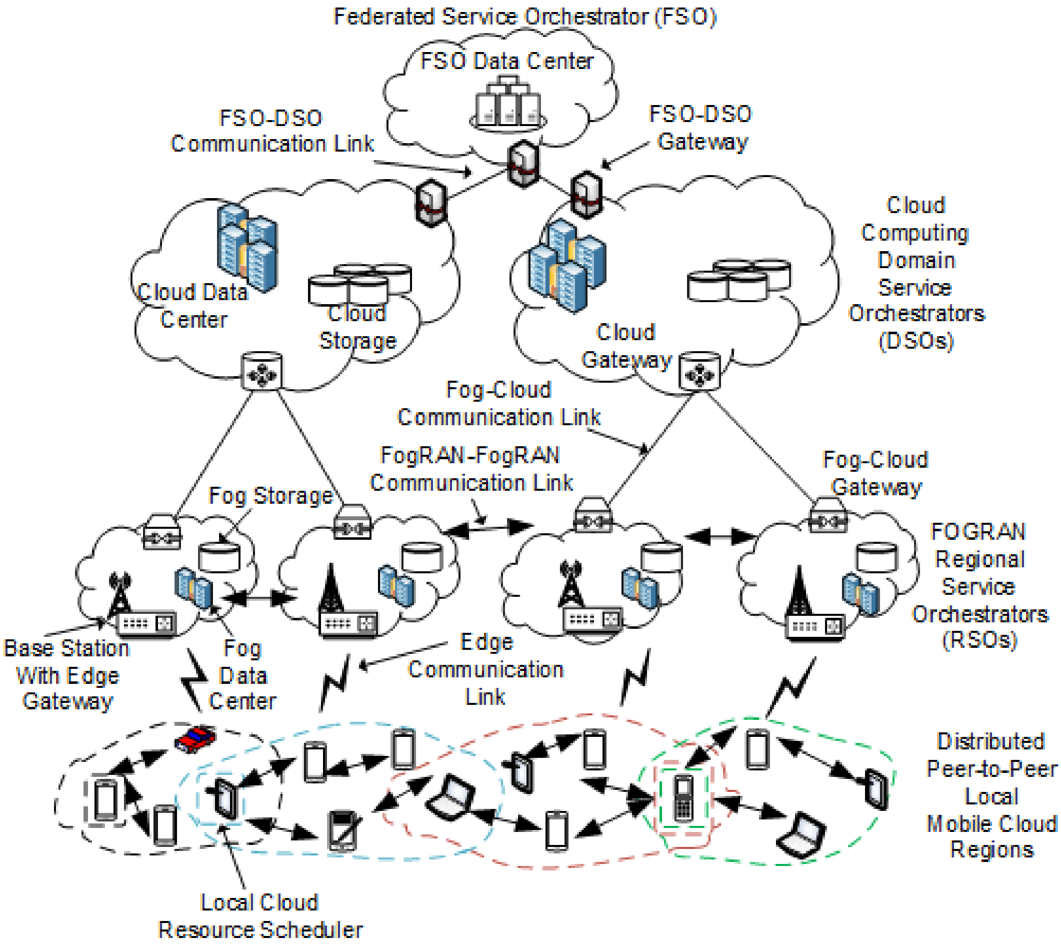


Figure 7.2 Fog Computing Architecture in 5G Mobile Networks [72]

The highest two levels are located in 5G core.

The top level is the Federated Service Orchestrator (FSO) which allows a fruitful interaction between different cloud computing Domain Service Orchestrators (DSOs). The FSO with its federation mechanisms and with its FSO Data Center would enable cooperation and exchange of data among communication links. The domains may belong to different entities, and can be administered by different authorities. Like that a multi-domain fog environment would be created, that would support service ubiquity.

The next level contains the cloud computing centers or cloud computing Domain Service Orchestrators (DSOs). Each DSO is responsible for a single domain and supervises several FogRAN Regional Service Orchestrators (RSOs) below. The DSO would support federation mechanisms to enable intra-domain cooperation and exchange of data between different fog regions. This is enabled through the cloud gateways, fog-cloud communication links and fog-cloud gateway. Each cloud computing center contains multiple high-end high performance computing cloud data centers, servers, and cloud storage that are capable of processing and storing an enormous amount of data.

The third level is the fog computing and networking layer, that consists of the 5G Radio Access Network (5G RAN), or FogRAN Regional Service Orchestrators (RSOs). Each RSO enables semi-autonomous operation a particular local cloud region. Each RSO could be interconnected with other RSOs and each of them is linked to the cloud. This allows the distribution of the load which provides scalability and a much higher proximity to the end users, i.e. lower latencies.

The fog computing and networking layer comprises of geo-distributed fog devices, deployed at the edge of the network, such as Fog Data Center and Fog Storage, that are intelligent enough to process, compute, and temporarily store the received information. The fog devices directly communicate with the mobile users through edge gateway and single-hop wireless connections using the off-the-shelf wireless interfaces, such as, LTE, WiFi, Bluetooth, etc. They can independently provide pre-defined service applications to mobile users without assistances from cloud or Internet. In addition, the fog servers are connected to the cloud in order to leverage the rich functions and application tools of the cloud.

Finally, the lowest level consists of local cloud regions, formed by a group of smart devices, such as smartphones, IoT, sensors, which sense multitude of events and transmit the sensed data to the upper fog computing and networking layer, for further processing if necessary. Within this region the smart devices can be either resource users, or resource providers. The devices form so called locally distributed peer-to-peer mobile cloud, where each device shares the resources with other devices in the same local cloud. The devices in each local cloud elect a Local Cloud Resource Scheduler, that performs management on the resource requests and allocates tasks to the devices in the local cloud or Fog Data Center if necessary. The decision about the election of the Local Cloud Resource Scheduler is done according to the connectivity to the local network, CPU performance battery life time, etc. One device can be served by several FOGRANs, and one device can be elected as local cloud resource scheduler for several local cloud regions.

The locally distributed peer-to-peer mobile cloud has its own strong capacities such as storage space, computational power, online time, and bandwidth. The workload of the application is managed in a distributed fashion without any point of centralization. The lack of centralization provides scalability, while exploitation of user resources reduces the service cost. The local cloud has ability to adapt to network failures and dynamically changing network topology with a transient population of devices, while ensuring acceptable connectivity and performance.

The scheduling of cloud computing resources to the smart user device are explained in the following two sections.

7.3 Schedulling of the Resources from the Locally Distributed Peer-to-Peer Mobile Cloud Devices

The scheduling of the cloud resources in the distributed peer-to-peer mobile cloud region is given Figure 7.3. Smart device 1 request cloud resources from the Local Cloud Resource Coordinator for data processing. The request includes information what type of task the device has to perform, how many resources the device has, and how many additional resources are needed to perform the task.

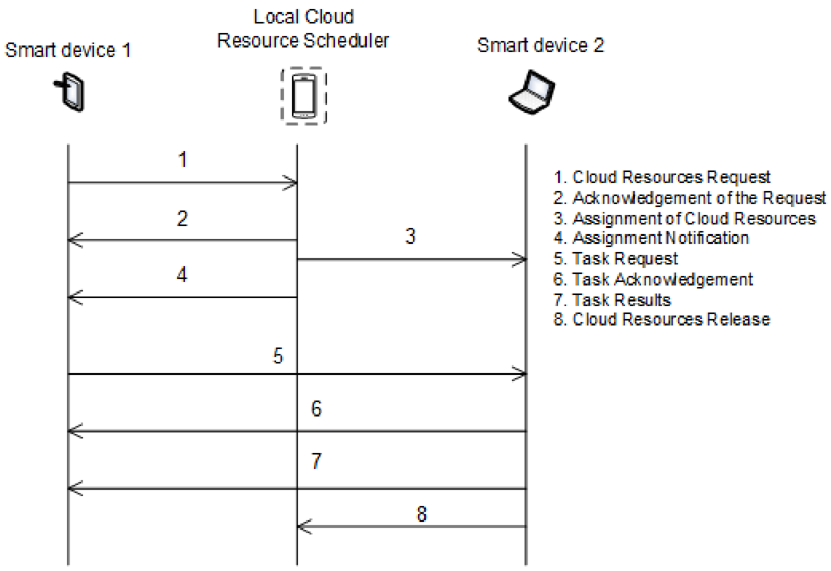


Figure 7.3 Schedulling of Local Distributed Peer-to-Peer Cloud Resources to the Mobile Device [72]

The Local Cloud Resource Coordinator acknowledges this request and starts to search for available cloud resources from the other smart user devices. When the available cloud resources are found for example in smart device 2, the local cloud resource coordinator allocates them to the smart device 1, and notifies him about this allocation. The local cloud resource coordinator may allocate resources from several smart devices.

When the local cloud resources are assigned to smart user device 1, it starts to transmit the task requests to the cloud resource provider (smart device 2). Smart device 2 acknowledges the task request, and starts to process the tasks, according to the instructions in the task requests. The smart device 2 forwards the results to the smart device 1, and also informs the cloud resource coordinator that its cloud resources are released.

7.4 Allocation of Cloud and Fog Computing Resources to the Smart Mobile Device

If the cloud resources provided to the user from the local cloud are insufficient, then the user would request cloud resources from the serving FOGRAN, or the cloud computing center. This is illustrated in Figure 7.4.

The smart device requests cloud resources from the serving RSO. The request includes information about what type of task the device has to perform, and how many resources are needed to perform the task.

Upon the request being received, the serving RSO acknowledges this request and checks whether with it has a capability to process this request. If there are sufficient cloud resources, the serving RSO would send an assignment notification of the cloud resources to the smart user device.

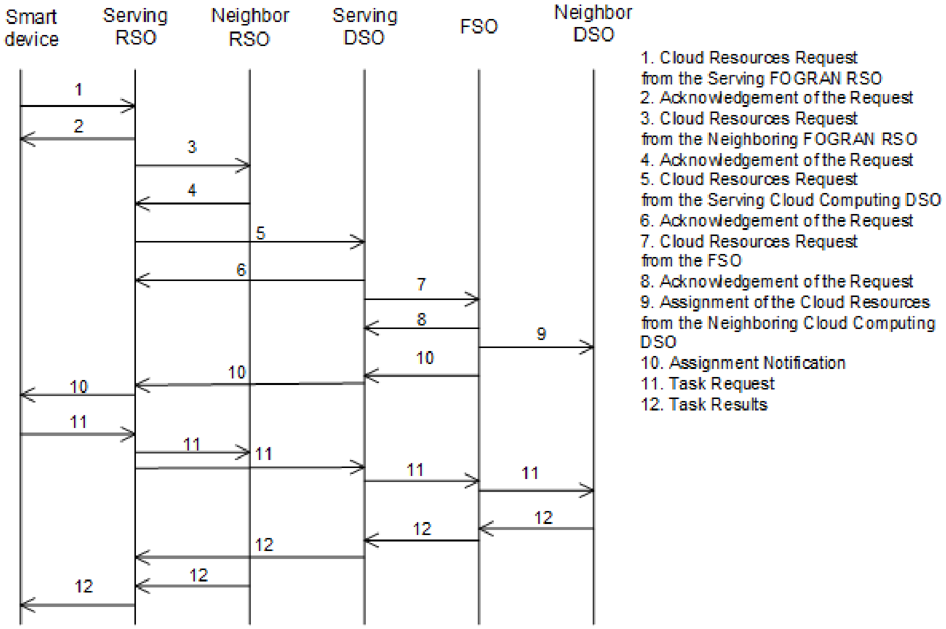


Figure 7.4 Scheduling of Cloud and FogRAN Resources to the Smart Mobile Device [72]

If the serving RSO doesn't possess enough resources to process the request, it shall communicate the neighboring RSO. Several neighboring RSOs can be contacted. The neighboring RSO sends either positive or negative acknowledgement to the request depending whether it has the necessary resources. If the neighbor RSO sends a positive acknowledgement it shall send assignment notification about the cloud resources to the smart user device via the serving RSO.

If the neighboring RSO sends a negative response, then the serving RSO would forward the request to the serving DSO. The serving DSO acknowledges the request and checks whether it contains the necessary information requested by the user device. If the serving DSO possess the relevant cloud resources, it shall notify the serving smart user device via the serving RSO.

If the serving DSO does not contain the necessary information, the request would be forwarded to the FSO. The FSO acknowledges this request and starts to look for DSOs that are capable to deal with this request. When such DSO is found, the FSO assigns the cloud resources to the correspondent FSO, and notifies the serving DSO about the resource assignment. Several neighbor DSOs can be assigned. The serving DSO forwards this notification to to the smart user device, via the serving RSO.

When the notification is received by the smart user, it starts to transmit the task request to the relevant cloud resource providers. The cloud resource providers process the task and forward the task results to the smart user device.

8 Quality Evaluation of Cloud and Fog Orchestrated Services in 5G Mobile Network

There are many ways to evaluate the quality of cloud and fog orchestrated services in 5G mobile networks. The evaluation can be performed in terms of Round Trip Time (RTT) latency, or end-to-end delay, throughput and energy efficiency for the used bits per power consumption for the user device, or vice versa [110 – 112]. Up till now there is no similar research performed for the quality of the cloud orchestrated services in 5G mobile networks.

8.1 Simulation Scenario

The simulation scenario is given in Figure 8.1. There is a single region in which are located a group of smart user devices, which are simultaneously served by 3G, 4G and 5G RAN network. Each RAN is connected to ten clouds. The base stations of these radio access networks are assumed to have equal regional coverage and are positioned in the same location.

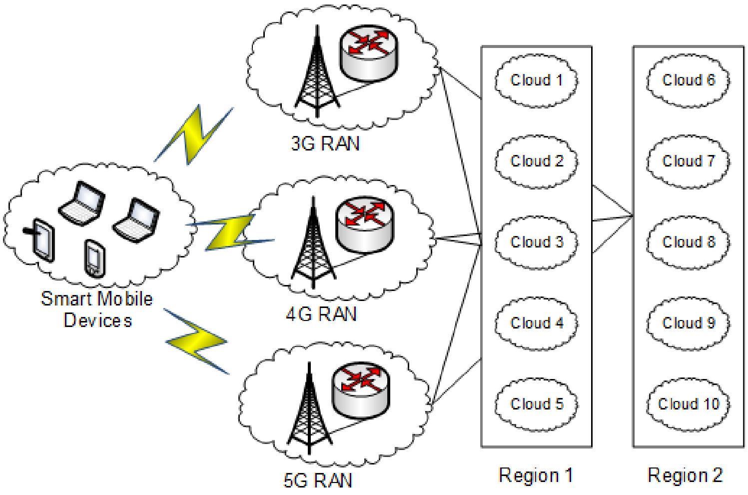


Figure 8.1 Simulation Scenario [110]

Every radio access network is connected to 10 clouds. First five clouds are in the same region with the RANs, and the other 5 clouds are in a different region with the RANs. The first cloud (cloud 1) has the nearest distance to the radio access network, and the last cloud (cloud 10) is the most distant from the radio access networks.

The smart user devices are assumed to be equally capable and are simultaneously served by the RANs and the clouds. Here a single user equipment will be considered.

The simulation scenario can be easily constructed with MATLAB. The obtained results represent an average value of 10 simulations.

8.2 RTT Latency or End-to-End Delay

RTT latency, or end-to-end delay is the time needed to happen a single transaction. It is the time it takes for the packet of data to travel to and from the source to the destination, and back to the source. The RTT latency between the user equipment towards any cloud is equal to:

$$RTT = RTT_{RAN} + RTT_{RAN-CLOUD} \tag{1}$$

where RTT_{RAN} represents the RTT RAN latency for 3G, 4G and 5G fog radio access networks, and $RTT_{RAN-CLOUD}$ represents the RTT latency between the radio access network and the cloud.

If there is a fog computing node in the radio access network (FogRAN) and if the information required by the user is contained in the fog computing node, then the RTT latency for any user device is equal to:

$$RTT = RTT_{RAN} \tag{2}$$

Table 8.1 RTT Latency for 3G, 4G and 5G RAN

Parameter	**RAN Type**		
	3G	***4G***	***5G***
RTT_RAN Latency [ms]	70	10	0.5 - 1
RTT_RAN-CLOUD Latency [ms]	40 - 250	30-100	1 - 5

The average values for RTT_{RAN} and $RTT_{RAN-CLOUD}$ are given in Table 8.1, which are taken from the references [52 – 53], [113]. For the simulation purposes random generated values are used for RTT_{RAN} and $RTT_{RAN-CLOUD}$. Here is considered the cloud 1 to be the least distant from the radio access networks and therefore it has the lowest $RTT_{RAN-CLOUD}$ latency, and the cloud 10 is the most distant from the radio access networks and because the large number of hops has the highest $RTT_{RAN-CLOUD}$ latency.

The simulation results for the RTT latency in a cloud and fog computing environment for 3G RAN, 4G RAN and 5G RAN are displayed on Figure 8.2. It can be noticed that RTT latency is increased from the cloud 1 to the cloud 10, for any radio access network, because the cloud 1 has the least distance to the base stations of the radio access networks, and cloud 10 has the greatest distance to the base stations of the radio access networks. Also, the RTT latency in 5G RAN is significantly lower than the RTT latency in 3G RAN and 4G RAN [114 – 115], especially in the fog computing environment.

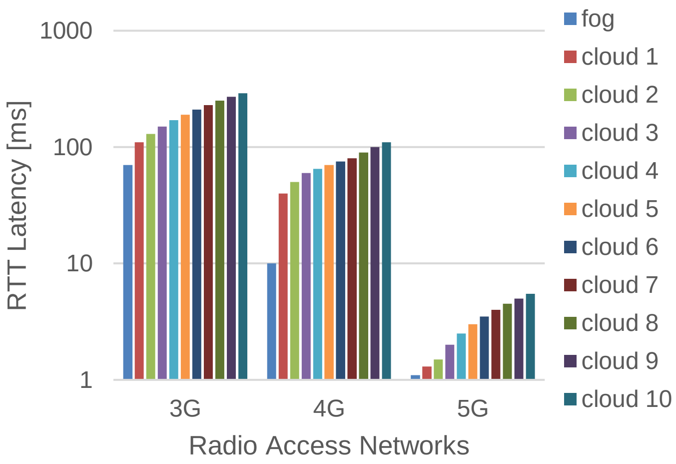


Figure 8.2 A Comparison between the RTT Latency in a Cloud and Fog Computing Environment in 3G, 4G and 5G RAN

8.3 Throughput

Throughput is the quantity of data that can pass from source to destination in a specific time [116]. The user throughput for any radio access network R can be calculated as a ratio between the peak data rate $R_{\max}$ of the radio access network and the number of the user devices N, and proportional to some weight coefficient μ:

$$R = \mu \cdot \frac{R_{\max}}{N}. \tag{3}$$

Here μ is a weight coefficient that models the bottleneck problem for the data that carry services from the cloud computing data centers. Due to the increased number of flows for different service requirements, the user throughput given in (3) is decreased for a certain factor.

The weight coefficient μ may receive values between 0.8 and 1, and its value depends how much the cloud is far away from the radio access network. If the cloud is closer to the base station of the radio access network then the coefficient μ has higher value, and if the cloud is at a greater distance from the base station of the radio access network then the coefficient μ would have lower value. If the mobile device uses a service that is located in the fog computing environment, i.e. in the radio access network, then the weight coefficient μ is equal to 1.

The number of the user device is taken to vary from 100 to 1000, with an increment step of 100.

Table 8.2 Modulation Coding Scheme (MCS) and Peak Data Rate $R_{\max}$ for 3G, 4G and 5G RAN in Downlink Direction [119]

3G RAN		**4G RAN**		**5G RAN**	
MCS	***Peak Data Rate (Mbps)***	***MCS***	***Peak Data Rate (Mbps)***	***MCS***	***Peak Data Rate (Mbps)***
64 QAM / MIMO 4x4	672	64 QAM / MIMO 8x8	3900	256 QAM / MIMO 16x16	50000
64 QAM / MIMO 2x2	84	64 QAM / MIMO 4x4	1000	256 QAM / MIMO 4x4	40000
64 QAM / MIMO 1x4	42	64 QAM / MIMO 2x2	600	128 QAM / MIMO 8x8	30000
16 QAM / MIMO 4x4	21	64 QAM	450	128 QAM / MIMO 4x4	20000
16 QAM / 15 codes	14.4	16 QAM / MIMO 4x4	390	128 QAM	15000
16 QAM / 10 codes	7.2	16 QAM / MIMO 2x2	300	64 QAM / MIMO 8x8	10000
16 QAM 5 codes	3.6	16 QAM/	150	64 QAM / MIMO 4x4	5000
QPSK 5 codes	1.8	QPSK / MIMO 4x4	100	64 QAM /	2500
8PSK	0.72	QPSK / MIMO 2x2	50	16 QAM / MIMO 4x4	1500
4PSK	0.384	QPSK	10	16 QAM	1000

The peak data rate $R_{\max}$ for any radio access network (3G, 4G or 5G) depends primary from the Adaptive Modulation Coding Scheme (AMSC), that makes a compensation for the noise, interference and other factors that have a negative influence on the useful signal in order to deliver higher capacity and better coverage in the presence of noise and other distortions.

Depending from the distance between the mobile user device and the radio access network, in the areas where the signal level is good, a modulation with a higher data rate and less robust coding is used. On the other hand, in the areas where the signal level is weak, or multi-path reflections exist, a modulation with lower data rate and more robust coding is used in order to minimize the errors. Table 8.2 and Table 8.3 provides the possible modulation coding schemes for 3G, 4G and 5G radio access networks in downlink and uplink direction, respectively [117 – 118].

In the real world, which modulation coding scheme would be applied at what distance is left to decide the network operator itself. In the simulations of this doctoral dissertation is taken the maximum distance between the mobile user device and the base station to be 5000 meters, where at every 500 meters the modulation coding scheme is changed.

Table 8.3 Modulation Coding Scheme (MCS) and Peak Data Rate $R_{\max}$ for 3G, 4G and 5G RAN in Uplink Direction

3G RAN		4G RAN		5G RAN	
MCS	***Peak Data Rate (Mbps))***	***MCS***	***Peak Data Rate (Mbps))***	***MCS***	***Peak Data Rate (Mbps)***
64 QAM / MIMO 4x4	168	64 QAM / MIMO 4x4	1500	256 QAM / MIMO 8x8	30000
64 QAM / MIMO 2x2	50	64 QAM / MIMO 2x2	900	256 QAM / MIMO 4x4	20000
64 QAM / MIMO 1x4	22	64 QAM / MIMO 1x2	800	128 QAM / MIMO 8x8	15000
16 QAM / MIMO 4x4	16	64 QAM	500	128 QAM / MIMO 4x4	10000
16 QAM / 15 codes	11.5	16 QAM / MIMO 4x4	150	128 QAM	8000
16 QAM / 10 codes	5.8	16 QAM / MIMO 2x2	100	64 QAM / MIMO 8x8	5000
16 QAM 5 codes	2	16 QAM/	75	64 QAM / MIMO 4x4	3000
QPSK 5 codes	1.8	QPSK / MIMO 4x4	50	64 QAM /	1500
8PSK	0.384	QPSK / MIMO 2x2	25	16 QAM / MIMO 4x4	1000
4PSK	0.153	QPSK	5	16 QAM	900

Some of the results of the user throughput are given on the Figures below as follows:

1. Downlink direction
    - A Comparison of the User Throughput in 3G Mobile Networks in Cloud and Fog Computing Environment as a Function of the User Devices at a distance from 0 to 500 meters on Figure 8.3;
    - A Comparison of the User Throughput in 3G Mobile Networks in Cloud and Fog Computing Environment as a Function of the Distance for 100 User Devices on Figure 8.4;

- A Comparison of the User Throughput in 4G Mobile Networks in Cloud and Fog Computing Environment as a Function of the User Devices at a distance from 0 to 500 meters on Figure 8.5;
- A Comparison of the User Throughput in 4G Mobile Networks in Cloud and Fog Computing Environment as a Function of the Distance for 100 User Devices on Figure 8.6;
- A Comparison of the User Throughput in 5G Mobile Networks in Cloud and Fog Computing Environment as a Function of the User Devices at a distance from 0 to 500 meters on Figure 8.7;
- A Comparison of the User Throughput in 5G Mobile Networks in Cloud and Fog Computing Environment as a Function of the Distance for 100 User Devices on Figure 8.8;
- A Comparison of the User Throughput in 3G, 4G and 5G Mobile Networks in a Fog Computing Environment as a Function of the User Devices at a distance from 0 to 500 meters on Figure 8.9; and
- A Comparison of the User Throughput in 3G, 4G and 5G Mobile Networks in a Fog Computing Environment as a Function of the Distance for 100 User Devices on Figure 8.10;

2. Uplink direction
    - A Comparison of the User Throughput in 3G Mobile Networks in Cloud and Fog Computing Environment as a Function of the User Devices at a distance from 0 to 500 meters on Figure 8.11;
    - A Comparison of the User Throughput in 3G Mobile Networks in Cloud and Fog Computing Environment as a Function of the Distance for 100 User Devices on Figure 8.12;
    - A Comparison of the User Throughput in 4G Mobile Networks in Cloud and Fog Computing Environment as a Function of the User Devices at a distance from 0 to 500 meters on Figure 8.13;
    - A Comparison of the User Throughput in 4G Mobile Networks in Cloud and Fog Computing Environment as a Function of the Distance for 100 User Devices on Figure 8.14;
    - A Comparison of the User Throughput in 5G Mobile Networks in Cloud and Fog Computing Environment as a Function of the User Devices at a distance from 0 to 500 meters on Figure 8.15;
    - A Comparison of the User Throughput in 5G Mobile Networks in Cloud and Fog Computing Environment as a Function of the Distance for 100 User Devices on Figure 8.16;
    - A Comparison of the User Throughput in 3G, 4G and 5G Mobile Networks in a Fog Computing Environment as a Function of the User Devices at a distance from 0 to 500 meters on Figure 8.17; and
    - A Comparison of the User Throughput in 3G, 4G and 5G Mobile Networks in a Fog Computing Environment as a Function of the Distance for 100 User Devices on Figure 8.18;

From all Figures it can be noticed that user throughput is reduced as the number of the user devices or the distance are increased for any mobile network. Also, 5G mobile network provides much higher user throughput than 3G and 4G mobile networks. Finally, the user throughput in a fog computing environment is greater than the user throughput in a cloud computing environment, independently of the radio access network (3G, 4G, or 5G RAN).

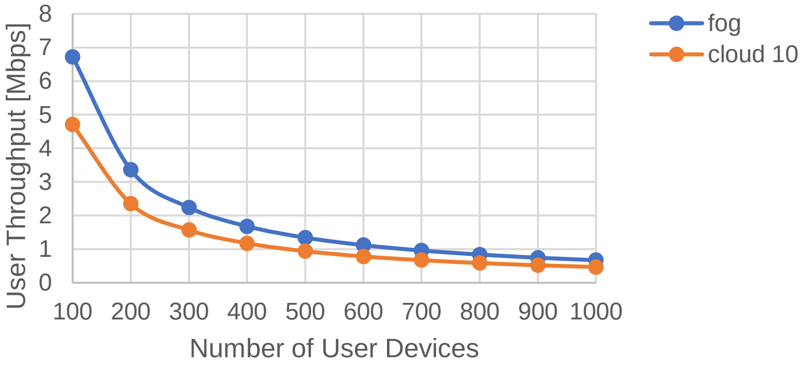


Figure 8.3 A Comparison of the User Throughput in 3G Mobile Networks in Cloud and Fog Computing Environment as a Function of the User Devices at a Distance from 0 to 500 meters in Downlink Direction

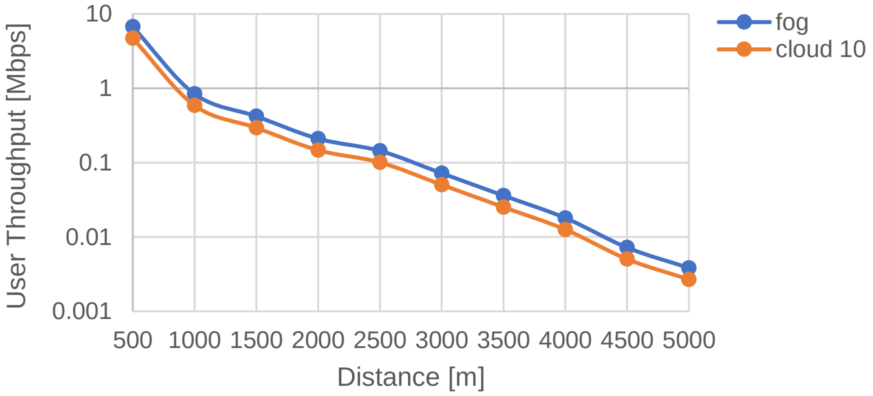


Figure 8.4 A Comparison of the User Throughput in 3G Mobile Networks in a Cloud and Fog Computing Environment as a Function of the Distance for 100 User Devices in Downlink Direction

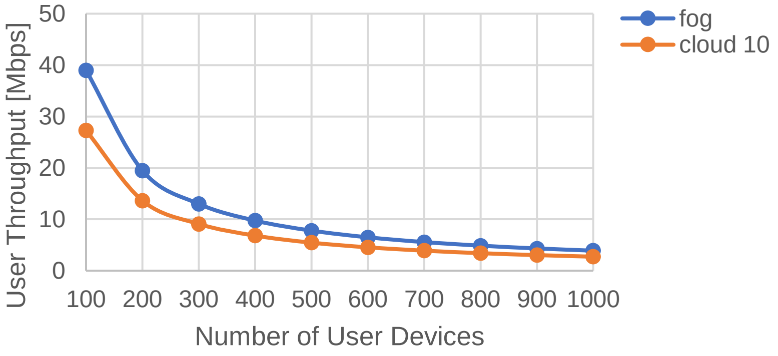


Figure 8.5 A Comparison of the User Throughput in 4G Mobile Networks in a Cloud and Fog Computing Environment as a Function of the User Devices at a Distance from 0 to 500 meters in Downlink Direction

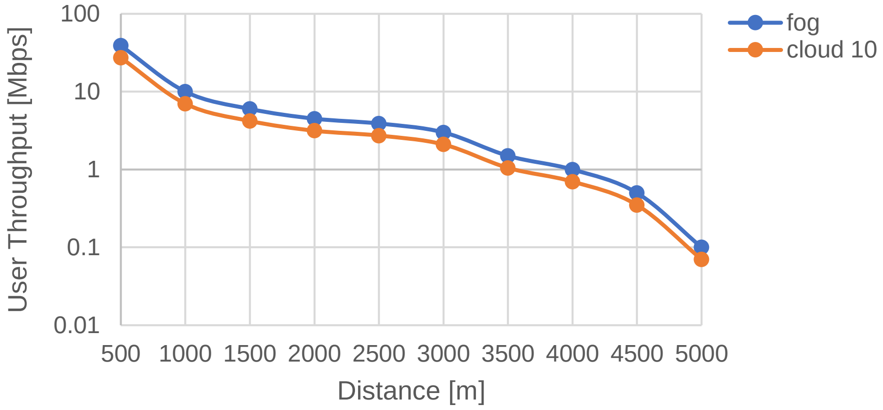


Figure 8.6 A Comparison of the User Throughput in 4G Mobile Networks in a Cloud and Fog Computing Environment as a Function of the Distance for 100 User Devices in Downlink Direction

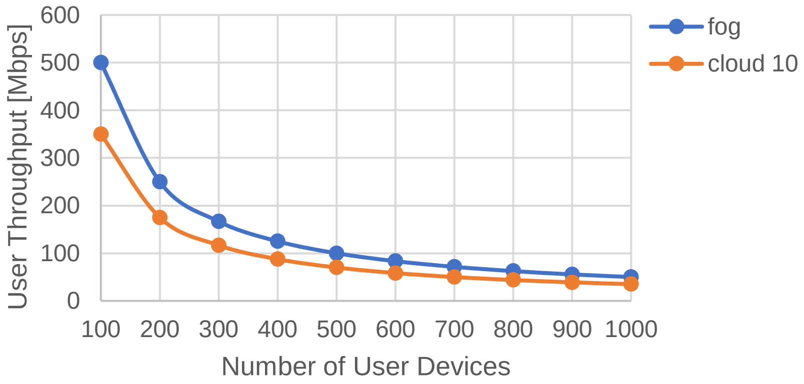


Figure 8.7 A Comparison of the User Throughput in 5G Mobile Networks in a Cloud and Fog Computing Environment as a Function of the User Devices at a Distance from 0 to 500 meters in Downlink Direction

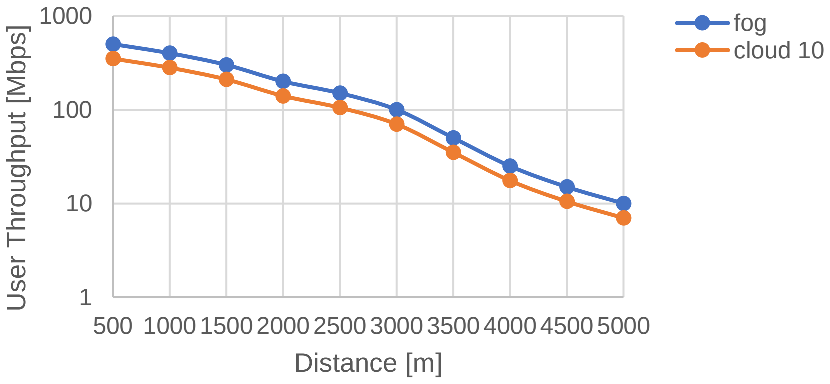


Figure 8.8 A Comparison of the User Throughput in 5G Mobile Networks in a Cloud and Fog Computing Environment as a Function of the Distance for 100 User Devices in Downlink Direction

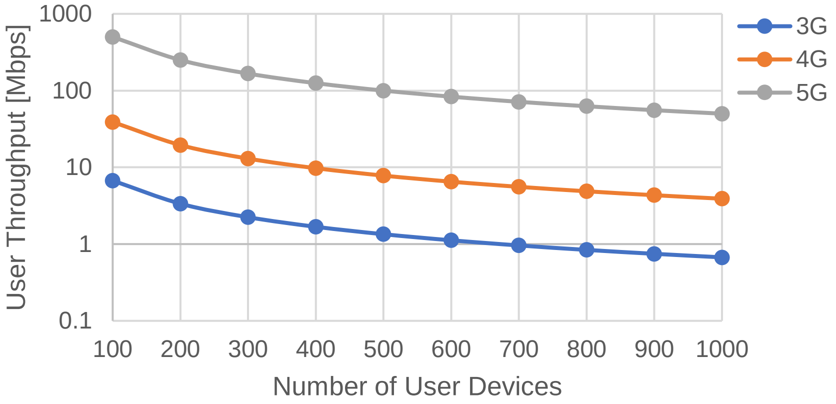


Figure 8.9 A Comparison of the User Throughput in 3G, 4G and 5G Mobile Networks in Fog Computing Environment as a Function of the User Devices at a Distance from 0 to 500 meters in Downlink Direction

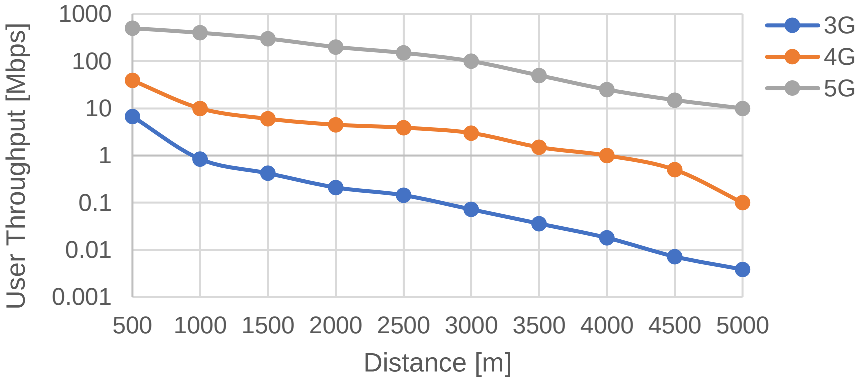


Figure 8.10 A Comparison of the User Throughput in 3G, 4G and 5G Mobile Networks in a Fog Computing Environment as a Function of the Distance for 100 User Devices in Downlink Direction

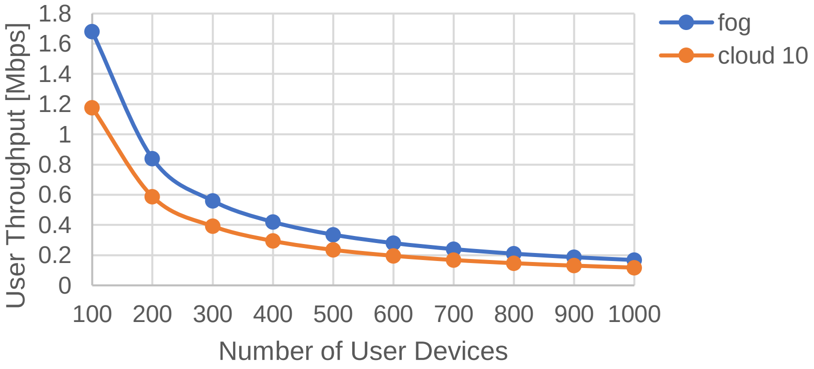


Figure 8.11 A Comparison of the User Throughput in 3G Mobile Networks in a Cloud and Fog Computing Environment as a Function of the User Devices at a Distance from 0 to 500 meters in Uplink Direction

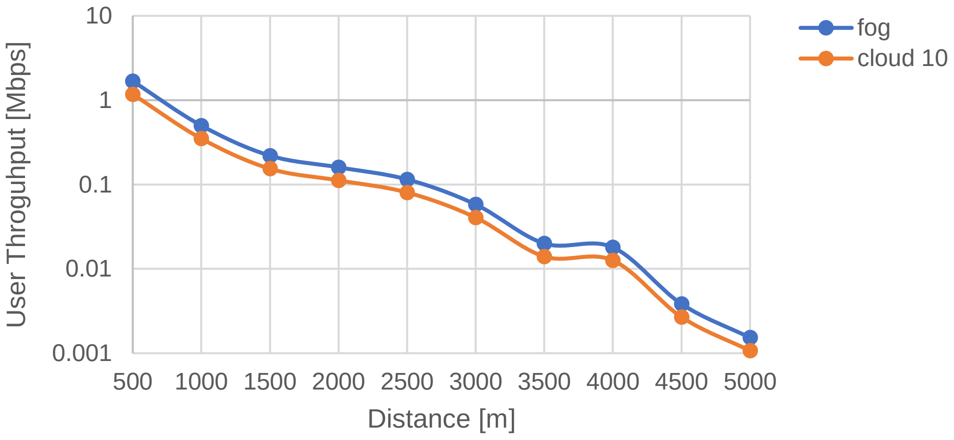


Figure 8.12 A Comparison of the User Throughput in 3G Mobile Networks in a Cloud and Fog Computing Environment as a Function of the Distance for 100 User Devices in Uplink Direction

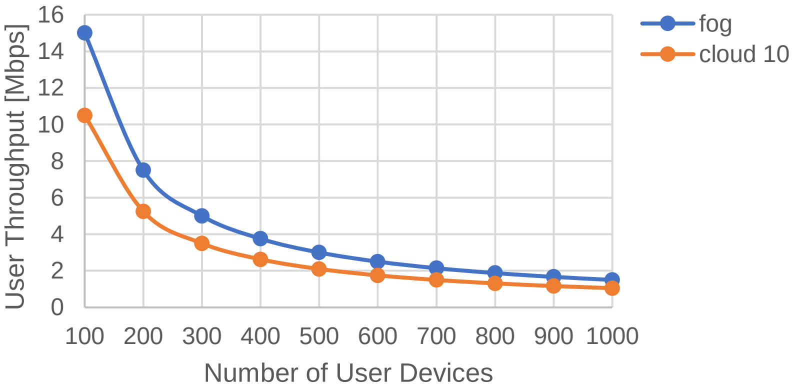


Figure 8.13 A Comparison of the User Throughput in 4G Mobile Networks in a Cloud and Fog Computing Environment as a Function of the User Devices at a Distance from 0 to 500 meters in Uplink Direction

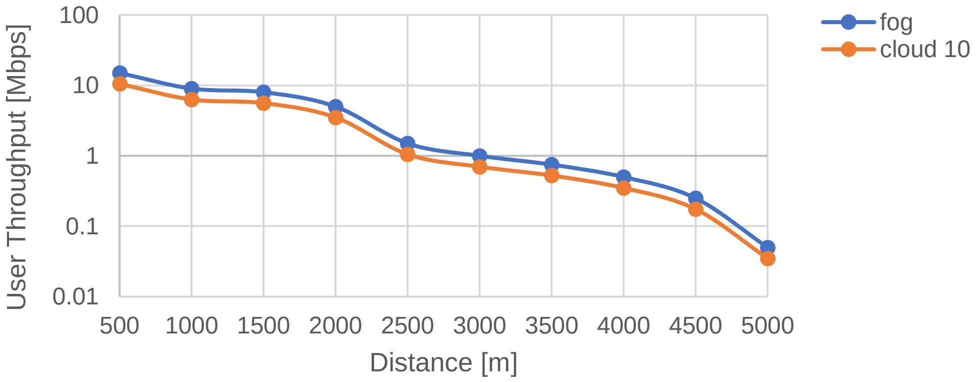


Figure 8.14 A Comparison of the User Throughput in 4G Mobile Networks in a Cloud and Fog Computing Environment as a Function of the Distance for 100 User Devices in Uplink Direction

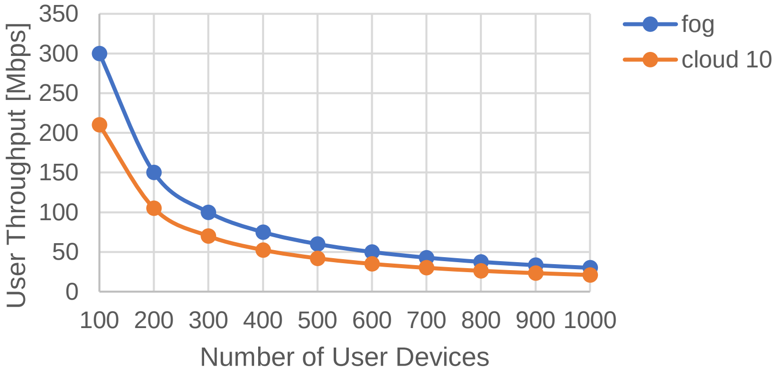


Figure 8.15 A Comparison of the User Throughput in 5G Mobile Networks in a Cloud and Fog Computing Environment as a Function of the User Devices at a Distance from 0 to 500 meters in Uplink Direction

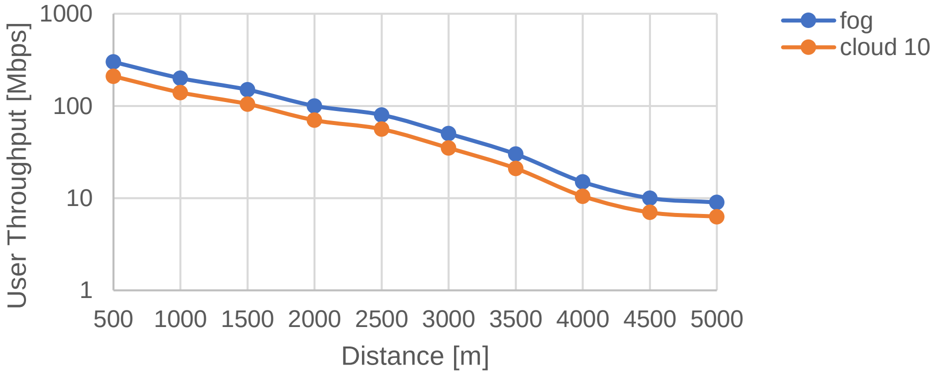


Figure 8.16 A Comparison of the User Throughput in 5G Mobile Networks in a Cloud and Fog Computing Environment as a Function of the Distance for 100 User Devices in Uplink Direction

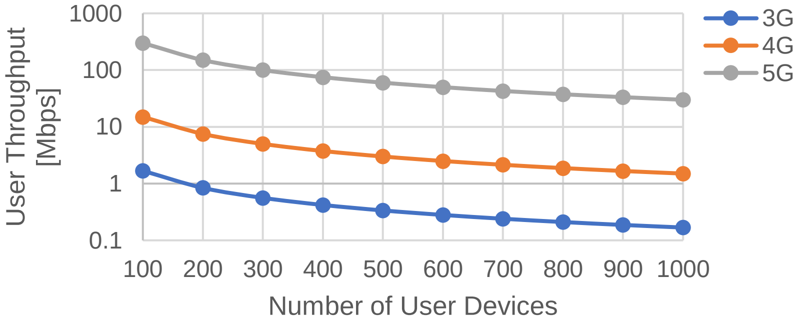


Figure 8.17 A Comparison of the User Throughput in 3G, 4G and 5G Mobile Networks in Fog Computing Environment as a Function of the User Devices at a Distance from 0 to 500 meters in Uplink Direction

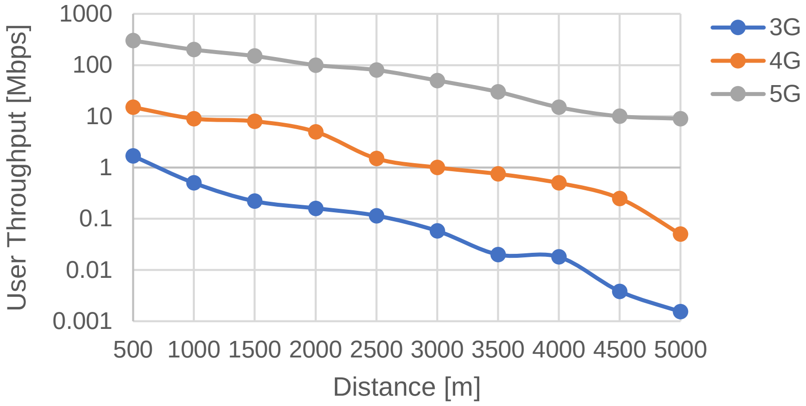


Figure 8.18 A Comparison of the User Throughput in 3G, 4G and 5G Mobile Networks in a Fog Computing Environment as a Function of the Distance for 100 User Devices in Uplink Direction

8.4 Energy Efficiency

The reduction of the power and power consumption by the networks and the devices is of vital importance for the economic and ecological sustainability in the industry. The general principle for minimizing of the power consumption at the network and the device should include all technology generations. This principle is recognized as an ecological goal, and is very important for the reduction of operating expenses in the network management [119]. In addition, the reduction of the power consumption would result to a longer battery life, which would contribute to a greater satisfaction at the mobile device users.

One of the possible methods for the reduction of power consumption in 5G mobile networks is by implementing the fog computing. One of the parameters that can examine the power consumption in the fog computing environment in 5G is through the energy efficiency [117], [120].

The energy efficiency *EE* represents the amount of data that can be transfered through the power consumed per user, usually on a single cell, and is the ratio between the user throughput *R* and the power *P*:

$$EE = \frac{R}{P}\left[\frac{\text{[bit/s/cell]}}{\text{[Joule/s/cell]}}\right] = \frac{R}{P}\left[\frac{\text{bit}}{\text{Joule}}\right]. \tag{4}$$

The reciprocal value of the energy efficiency represents the energy *E* consumed per bit per user:

$$E = \frac{1}{EE} = \frac{P}{R}\left[\frac{\text{[Joule/s/cell]}}{\text{[bit/s/cell]}}\right] = \frac{P}{R}\left[\frac{\text{Joule}}{\text{bit}}\right]. \tag{5}$$

In the relations (4) and (5) *R* is the user throughput which was already discussed in the previous section.

The consumed power *P* can be expressed through the user throughput *R* with the following linear equation [121]:

$$P = \alpha R + \beta \tag{6}$$

where α is the coefficient that gives the power necessary for data transfer (in downlink, or uplink direction), and β is a coefficient that represents the idle power [122]. Table 8.4 gives the values of these coefficients for 3G, 4G and 5G mobile networks.

Table 8.4 Typical Values for the Power Consuming Coefficients

Mobile Network	***α [mW/Mbps] (downlink)***	***α [mW/Mbps] (uplink)***	***β [mW]***
3G	122.12	868.98	817.88
4G	51.97	438.39	1288.04
5G	6.5	65	11475.97

Below are provided some of the results for energy efficiency as follows:

1. Downlink direction
   - A Comparison of the Energy Efficiency in 3G Mobile Networks in Cloud and Fog Computing Environment as a Function of User Devices at a Distance from 0 to 500 meters on Figure 8.18;
   - A Comparison of the Energy Efficiency in 3G Mobile Networks in Cloud and Fog Computing Environment as a Function of the Distance for 100 User Devices on Figure 8.19;
   - A Comparison of the Energy Efficiency in 4G Mobile Networks in Cloud and Fog Computing Environment as a Function of User Devices at a Distance from 0 to 500 meters on Figure 8.20;
   - A Comparison of the Energy Efficiency in 4G Mobile Networks in Cloud and Fog Computing Environment as a Function of the Distance for 100 User Devices on Figure 8.21;
   - A Comparison of the Energy Efficiency in 5G Mobile Networks in Cloud and Fog Computing Environment as a Function of User Devices at a Distance from 0 to 500 meters on Figure 8.22;
   - A Comparison of the Energy Efficiency in 5G Mobile Networks in Cloud and Fog Computing Environment as a Function of the Distance for 100 User Devices on Figure 8.23;
   - A Comparison of the Energy Efficiency in 3G, 4G, and 5G Mobile Networks in a Fog Computing Environment as a Function of User Devices at a Distance from 0 to 500 meters on Figure 8.24; and
   - A Comparison of the Energy Efficiency in 3G, 4G, and 5G Mobile Networks in a Fog Computing Environment as a Function of the Distance for 100 User Devices on Figure 8.25;

2. Uplink direction
   - A Comparison of the Energy Efficiency in 3G Mobile Networks in Cloud and Fog Computing Environment as a Function of User Devices at a Distance from 0 to 500 meters on Figure 8.26;
   - A Comparison of the Energy Efficiency in 3G Mobile Networks in Cloud and Fog Computing Environment as a Function of the Distance for 100 User Devices on Figure 8.27;
   - A Comparison of the Energy Efficiency in 4G Mobile Networks in Cloud and Fog Computing Environment as a Function of User Devices at a Distance from 0 to 500 meters on Figure 8.28;

- A Comparison of the Energy Efficiency in 4G Mobile Networks in Cloud and Fog Computing Environment as a Function of the Distance for 100 User Devices on Figure 8.29;
- A Comparison of the Energy Efficiency in 5G Mobile Networks in Cloud and Fog Computing Environment as a Function of User Devices at a Distance from 0 to 500 meters on Figure 8.30;
- A Comparison of the Energy Efficiency in 5G Mobile Networks in Cloud and Fog Computing Environment as a Function of the Distance for 100 User Devices on Figure 8.31;
- A Comparison of the Energy Efficiency in 3G, 4G, and 5G Mobile Networks in a Fog Computing Environment as a Function of User Devices at a Distance from 0 to 500 meters on Figure 8.32; and
- A Comparison of the Energy Efficiency in 3G, 4G, and 5G Mobile Networks in a Fog Computing Environment as a Function of the Distance for 100 User Devices on Figure 8.33;

From all Figures it can be noticed that energy efficiency is reduced as the number of the user devices or the distance are increased for any mobile network. 5G mobile network provides much higher energy efficiency than 3G and 4G mobile networks. Also, the energy efficiency in 3G, 4G or 5G mobile networks has better performances in fog computing environment, rather than in cloud computing environment.

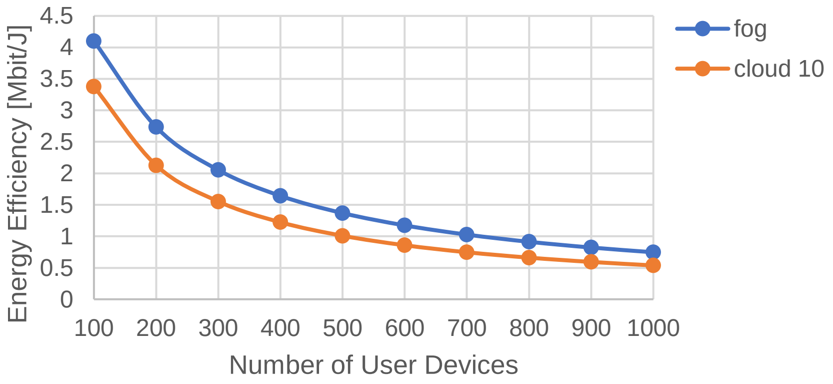


Figure 8.19 A Comparison of the Energy Efficiency in 3G Mobile Networks in Cloud and Fog Computing Environment as a Function of the User Devices at a Distance from 0 to 500 meters in Downlink Direction

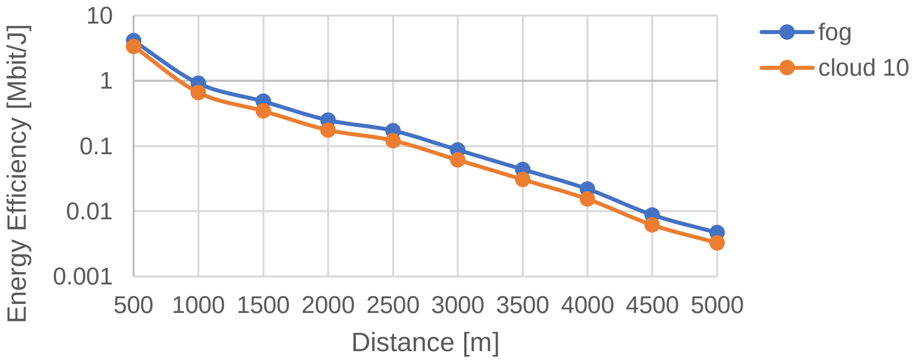


Figure 8.20 A Comparison of the Energy Efficiency in 3G Mobile Networks in Cloud and Fog Computing Environment as a Function of the Distance for 100 User Devices in Downlink Direction

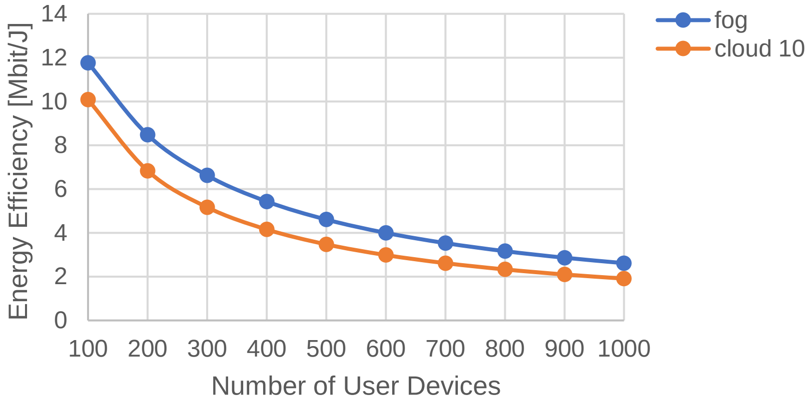


Figure 8.21 A Comparison of the Energy Efficiency in 4G Mobile Networks in Cloud and Fog Computing Environment as a Function of the User Devices at a Distance from 0 to 500 meters in Downlink Direction

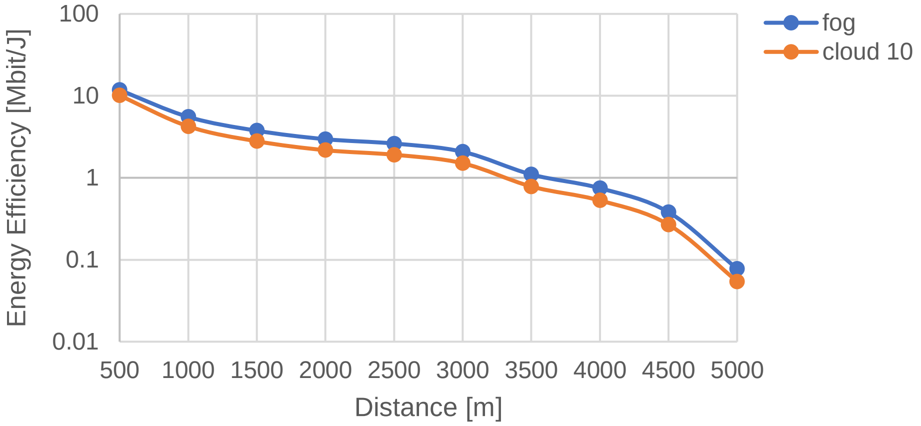


Figure 8.22 A Comparison of the Energy Efficiency in 4G Mobile Networks in Cloud and Fog Computing Environment as a Function of the Distance for 100 User Devices in Downlink Direction

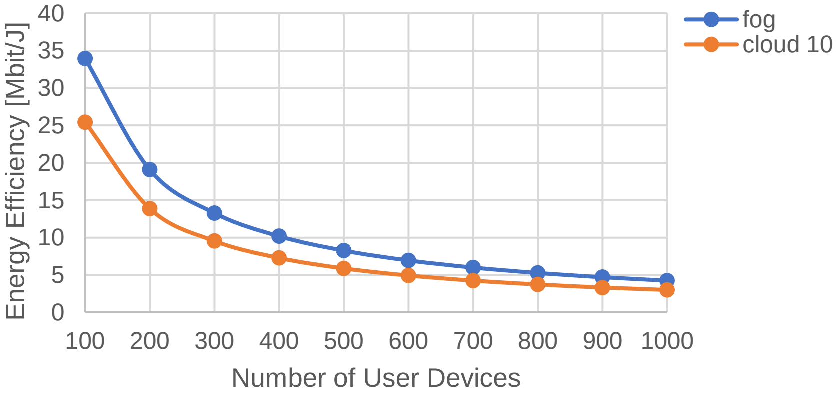


Figure 8.23 A Comparison of the Energy Efficiency in 5G Mobile Networks in Cloud and Fog Computing Environment as a Function of the User Devices at a Distance from 0 to 500 meters in Downlink Direction

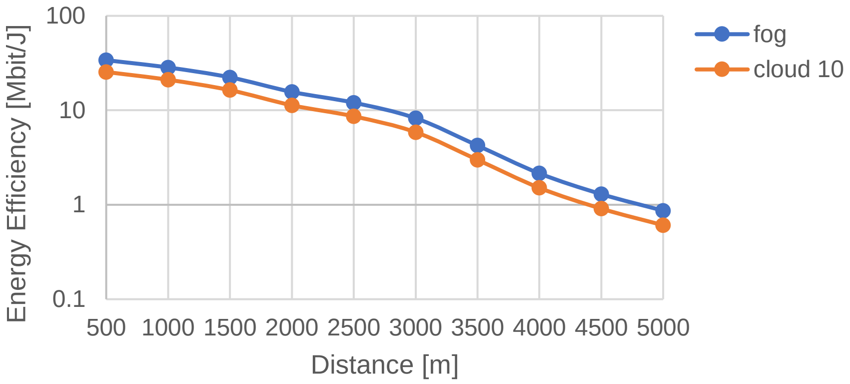


Figure 8.24 A Comparison of the Energy Efficiency in 5G Mobile Networks in Cloud and Fog Computing Environment as a Function of the Distance for 100 User Devices in Downlink Direction

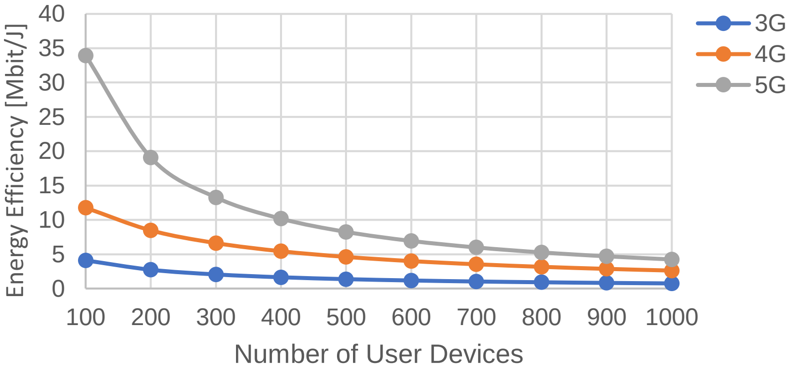


Figure 8.25 A Comparison of the Energy Efficiency in 3G, 4G and 5G Mobile Networks in a Fog Computing Environment as a Function of the User Devices at a Distance from 0 to 500 meters in Downlink Direction

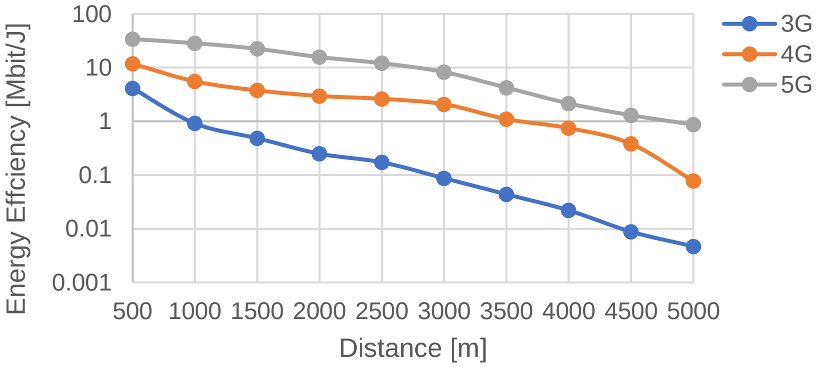


Figure 8.26 A Comparison of the Energy Efficiency in 3G, 4G and 5G Mobile Networks in a Fog Computing Environment as a Function of the Distance for 100 User Devices in Downlink Direction

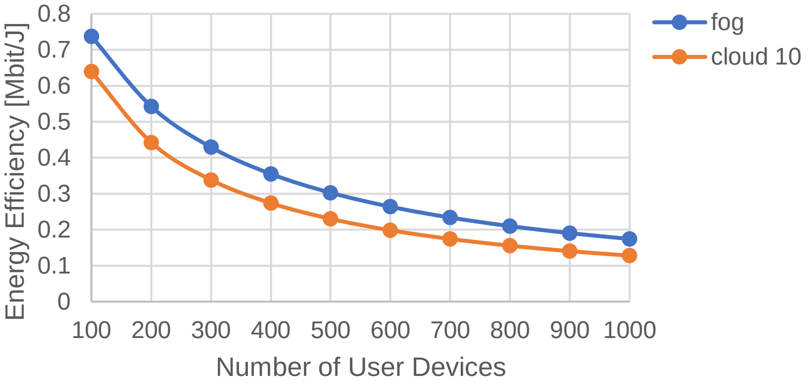


Figure 8.27 A Comparison of the Energy Efficiency in 3G Mobile Networks in Cloud and Fog Computing Environment as a Function of the User Devices at a Distance from 0 to 500 meters in Uplink Direction

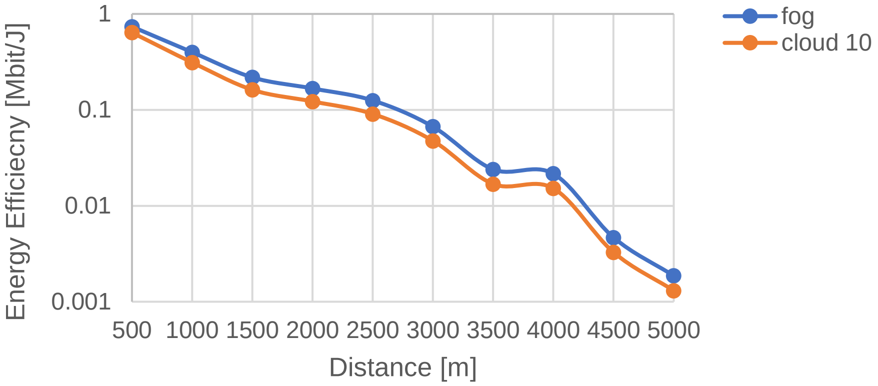


Figure 8.28 A Comparison of the Energy Efficiency in 3G Mobile Networks in Cloud and Fog Computing Environment as a Function of the Distance for 100 User Devices in Uplink Direction

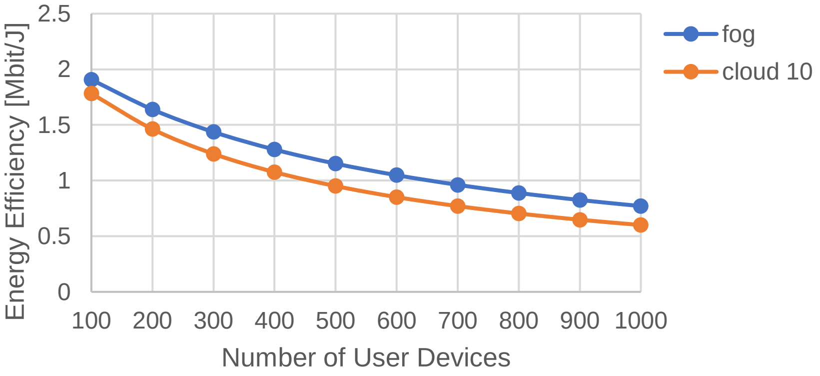


Figure 8.29 A Comparison of the Energy Efficiency in 4G Mobile Networks in Cloud and Fog Computing Environment as a Function of the User Devices at a Distance from 0 to 500 meters in Uplink Direction

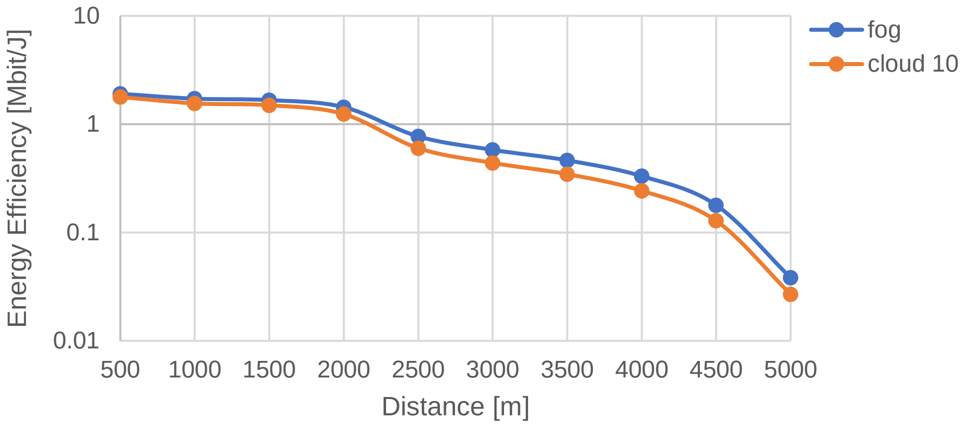


Figure 8.30 A Comparison of the Energy Efficiency in 4G Mobile Networks in Cloud and Fog Computing Environment as a Function of the Distance for 100 User Devices in Uplink Direction

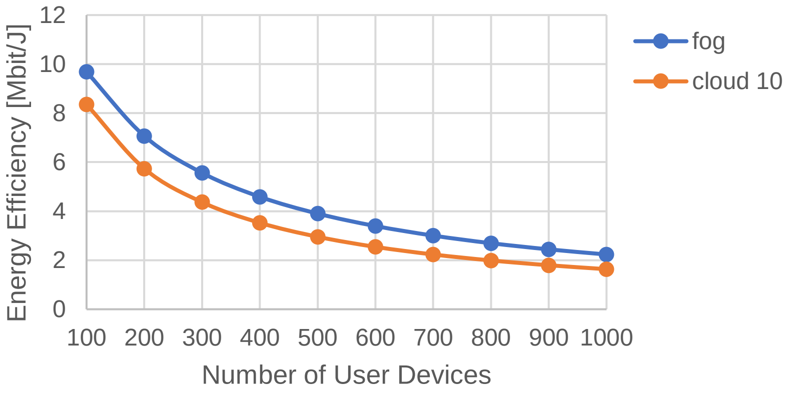


Figure 8.31 A Comparison of the Energy Efficiency in 5G Mobile Networks in Cloud and Fog Computing Environment as a Function of the User Devices at a Distance from 0 to 500 meters in Uplink Direction

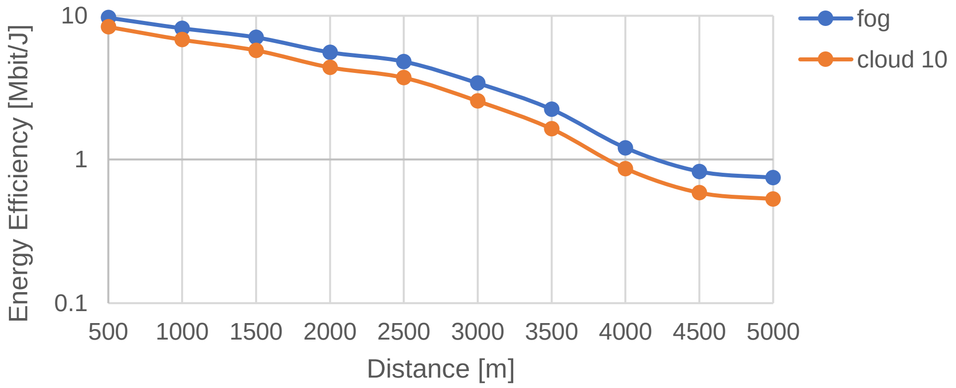


Figure 8.32 A Comparison of the Energy Efficiency in 5G Mobile Networks in Cloud and Fog Computing Environment as a Function of the Distance for 100 User Devices in Uplink Direction

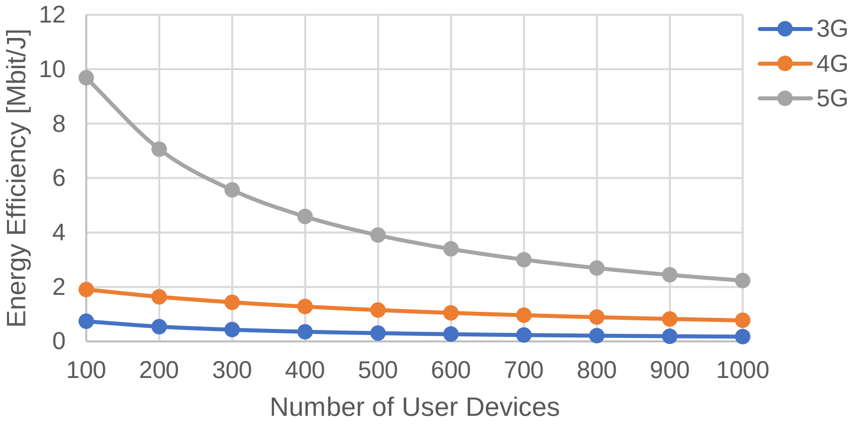


Figure 8.33 A Comparison of the Energy Efficiency in 3G, 4G and 5G Mobile Networks in a Fog Computing Environment as a Function of the User Devices at a Distance from 0 to 500 meters in Uplink Direction

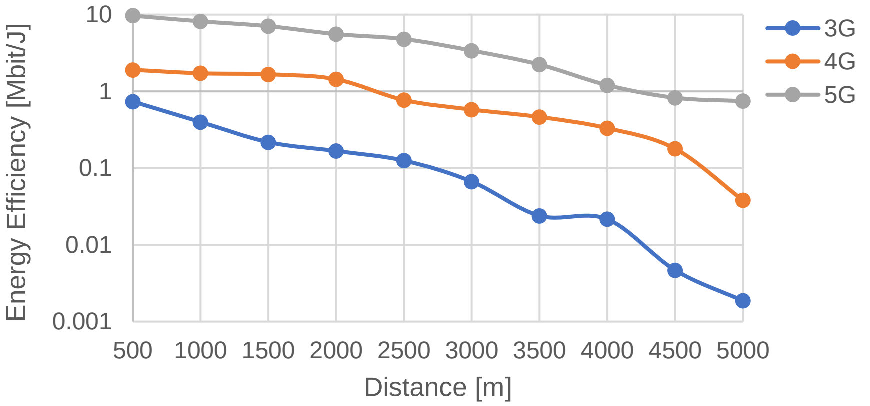


Figure 8.34 A Comparison of the Energy Efficiency in 3G, 4G and 5G Mobile Networks in a Fog Computing Environment as a Function of the Distance for 100 User Devices in Uplink Direction

8.5 Discussion of the Results

This chapter presented some of the results through which it was evaluated the quality of cloud and fog orchestrated services in 5G mobile networks. For the evaluation purposes the following parameters were used: RTT latency (E2E delay), user throughput in downlink and uplink direction and energy efficiency of the user device in downlink and uplink direction. These parameters were initially compared separately in 3G, 4G and 5G mobile networks, and then they were compared in all three networks together, as a function of the number of user devices and the distance between the user device and the base station.

It was concluded that if the fog computing environment is implemented independently in any mobile network gives better results than the cloud computing environment. Also, 5G network provides much better performances than 3G and 4G network in both cloud and fog computing environment.

These results clearly demonstrate the benefits of implementing the fog computing service orchestration mechanisms in 5G mobile networks, and provide a proof for the main hypothesis and the two special thesis given at the introduction part of this doctoral dissertation. The selection of an optimal network that would provide the best user satisfaction depends primarily from the user service requirements, such as high user throughput, low latency, better energy efficiency, etc.

In particular, the big data analytics that requires real time processing and very often has stringent time requirement can only be carried out in the fog. This is essential for critical usage cases of IoT devices and Tactile Internet where 1 ms end-to-end latency is required in the network in order to provide virtual-reality-type interfaces between humans and machines (human-machine interaction and machine-machine interaction) [109].

9 Algorithm for Optimal Selection of 5G Radio Access Network

In 5G mobile network, every user device will use cloud and data services through different radio access networks and different clouds. Therefore, it is necessary an algorithm for optimal selection of the best radio access network. The selection of such radio access network primarily depends from the data service requirements which are being transmitted through 5G network, such as throughput bandwidth (downlink and uplink direction), RTT latency (E2E delay), the energy efficiency of the radio access network, etc. [110], [117].

9.1 Requirements for Different Types of Services

Figure 9.1 illustrates the requirments for different types of services in terms of latency, and throughput bandwidth. It can be notices that some of services can be delivered only with 5G network, while others can be delivered with the existing network technologies.

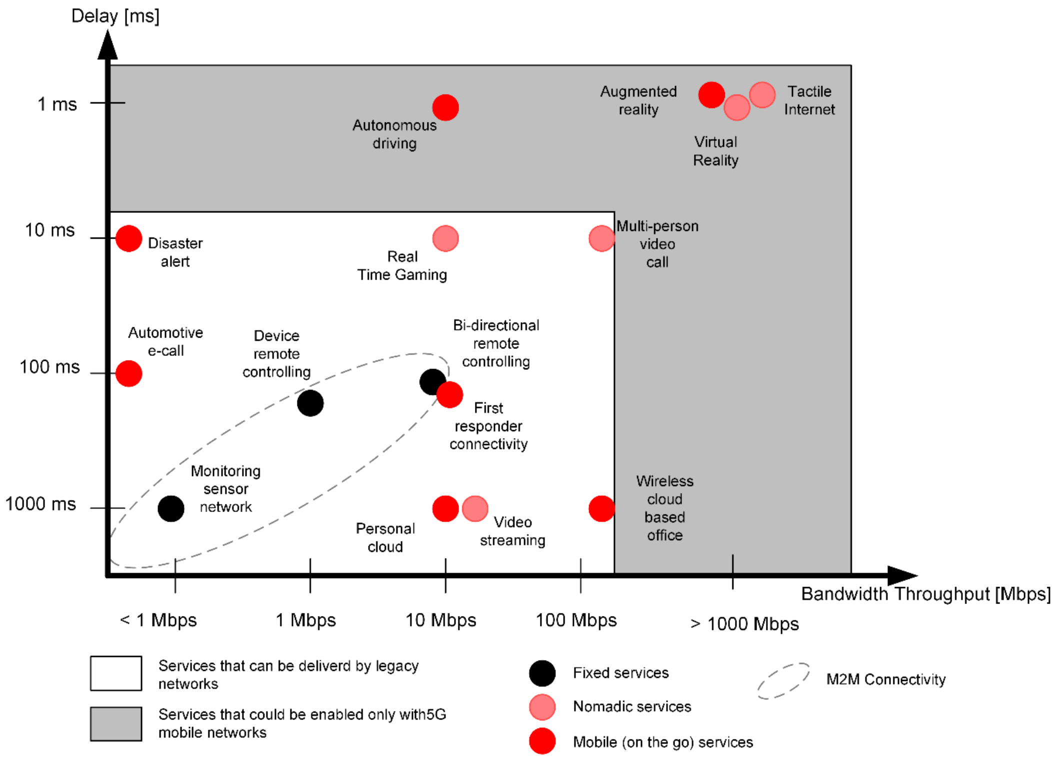


Figure 9.1 Requirements for the Latency (Delay) and the Bandwidth Throughput for Different Types of Services [113]

As can be noticed the following services can be delivered only with 5G: virtual reality, augmented reality, tactile Internet and autonomous driving/connected vehicles. Although the wireless cloud based office can be delivered with the existing 4G mobile networks, in the future it is expected this service to have requirements for higher bit rates and lower latency. The M2M communications is already enabled with the existing technologies. However, in the future the number of the devices that use such type of communication is expected to grow up from almost 250 million in 2014 to almost 2 billion in 2020 [113], which would result with demands for increased data rate and lower latency.

Apart from the latency and the throughput bandwidth, another important parameter for the optimal selection of radio access network is the energy efficiency, which has an important role in the reduction of operational costs at the network operator, and an increased battery life of the smart user devices.

9.2 Possible Algorithms for Optimal Selection of Radio Access Network

Because there are different types of services that have different requirements for the throughput bandwidth, latency, energy efficiency, etc. in 5G network it is necessary to implement an algorithm for optimal selection of radio access network that would deliver cloud computing services in an efficient manner.

In this doctoral dissertation 5 algorithms are proposed. In the analysis it is made a comparison between the results obtained when the algorithm is applied, and the results obtained when the algorithm is not applied. The proposed algorithms are:

1. **Algorithm for an optimal selection of radio access network according to criterion for minimal RTT latency – Algorithm 1**

   Find the minimal latency subject to:
   - *Given number of user devices N*
   - *Given distance d between the user device and the 5G base station*
   - *User throghput* $R \geq R_{\min}$ *(in downlink and uplink direction)*
   - *Energy Efficiency* $EE \geq EE_{\min}$ *(in downlink and uplink direction)*

   Notice:
   A) If *R* (downlink and uplink), or *EE* (downlink and uplink) cannot be found that fulfill the above conditions, then for *R* (downlink and uplink) and *EE* (downlink and uplink) are chosen the ones that are close to $R_{\min}$ (downlink and uplink) and $EE_{\min}$ (downlink and uplink), respectively.
   B) If more than one 5G base stations that fulfill the criterion are found, then it should be selected the candidate with the lowest latency.

2. **Algorithm for an optimal selection of radio access network according to criterion for maximum user throughput *R* (downlink direction) – Algorithm 2**

   Find the maximum user throughput R (downlink direction) subject to:
   - *Given number of user devices N*
   - *Given distance d between the user device and the 5G base station*
   - *Latency (delay)* $RTT \leq RTT_{\max}$
   - *User throghput* $R \geq R_{\min}$ *(in uplink direction)*
   - *Energy Efficiency* $EE \geq EE_{\min}$ *(in downlink and uplink direction)*

   Notice:
   A) If *RTT*, *R* (uplink) and *EE* (downlink and uplink) cannot be found that fulfill the above conditions, then for *RTT, R* (uplink) and *EE* (downlink and uplink) are chosen the ones that are close to $RTT_{\max}$, $R_{\min}$ (uplink) and $EE_{\min}$ (downlink and uplink), respectively.
   B) If more than one 5G base stations that fulfill the criterion are found, then it should be selected the candidate with the highest user throughput (downlink direction).

3. Algorithm for an optimal selection of radio access network according to criterion for maximum user throughput *R* (uplink direction) – Algorithm 3

Find the maximum user throughput R (uplink direction) subject to:

- *Given number of user devices N*
- *Given distance d between the user device and the 5G base station*
- *Latency (delay)* $RTT \le RTT_{\max}$
- *User throghput* $R \ge R_{\min}$ *(in downlink direction)*
- *Energy Efficiency* $EE \ge EE_{\min}$ *(in downlink and uplink direction)*

Notice:
A) If *RTT*, *R* (downlink) and *EE* (downlink and uplink) cannot be found that fulfill the above conditions, then for *RTT, R* (downlink) and *EE* (downlink and uplink) are chosen the ones that are close to $RTT_{\max}$, $R_{\min}$ downlink) and $EE_{\min}$ (downlink and uplink), respectively.
B) If more than one 5G base stations that fulfill the criterion are found, then it should be selected the candidate with the highest user throughput (uplink direction).

4. Algorithm for an optimal selection of radio access network according to criterion for maximum energy efficiency *EE* (downlink direction) – Algorithm 4

Find the maximum energy efficiency EE (in downlink direction) subject to:

- *Given number of user devices N*
- *Given distance d between the user device and the 5G base station*
- *Latency (delay)* $RTT \le RTT_{\max}$
- *User throghput* $R \ge R_{\min}$ *(in downlink and uplink direction)*
- *Energy Efficiency* $EE \ge EE_{\min}$ *(in uplink direction)*

Notice:
A) If *RTT*, *R* (downlink and uplink) and *EE* (uplink) cannot be found that fulfill the above conditions, then for *RTT, R* (downlink and uplink) and *EE* (uplink) are chosen the ones that are close to $RTT_{\max}$, $R_{\min}$ (downlink and uplink) and $EE_{\min}$ (uplink), respectively.
B) If more than one 5G base stations that fulfill the criterion are found, then it should be selected the candidate with the highest energy efficiency (downlink direction).

5. Algorithm for an optimal selection of radio access network according to criterion for maximum energy efficiency *EE* (uplink direction) – Algorithm 5

Find the maximum energy efficiency EE (in uplink direction) subject to:

- *Given number of user devices N*
- *Given distance d between the user device and the 5G base station*
- *Latency (delay)* $RTT \le RTT_{\max}$
- *User throghput* $R \ge R_{\min}$ *(in downlink and uplink direction)*
- *Energy Efficiency* $EE \ge EE_{\min}$ *(in downlink direction)*

Notice:
A) If *RTT*, *R* (downlink and uplink) and *EE* (downlink) cannot be found that fulfill the above conditions, then for *RTT*, *R* (downlink and uplink) and *EE* (downlink) are chosen the ones that are close to $RTT_{\max}$, $R_{\min}$ (downlink and uplink) and $EE_{\min}$ (downlink), respectively.
B) If more than one 5G base stations that fulfill the criterion are found, then it should be selected the candidate with the highest energy efficiency (uplink direction).

In order to evaluate of all of these algorithms the number of user devices *N* is taken to vary 100 to 1000, with an incrementing step of 100, and the distance *d* to vary from 500 to 5000 meters, with an incrementing step of 500.

The parameters $RTT_{\max}$, $R_{\min}$, $EE_{\min}$ are decision thresholds for the selecting the RTT latency, user throughput (downlink and uplink direction) and energy efficiency (downlink and uplink direction). As values for these threshold parameters are taken the calculated theoretical values for the RTT latency, user throughput (downlink and uplink direction) and the energy efficiency (downlink and uplink direction), depending from the number of user devices and the distance between the user and the base station of the mobile network, which were obtained through the analysis in Chapter 8.

9.3 Scenario for Algorithms' Evaluation

In order to perform an evaluation for the algorithms in the previous section the following scenario is used. There is a single region that is covered by several 5G RANs, that serve many user devices. One user device is selected for evaluation, which has a different distance to the base stations of 5G RANs, that are connected to several cloud computing centers.

Let the user device uses a certain cloud computing service that can be offered by all 5G RANs. The user device in its own database obtains measurement reports for the latency, user throughput (downlink and uplink) and the energy efficiency (downlink and uplink) from every 5G radio access network separately.

Depending from the distance between the user device and the base station of 5G RAN, i.e. the applied modulation coding scheme and the radio conditions, every 5G base station would offer to the mobile user device certain RTT latency, user throughput (downlink and uplink) and energy efficiency (downlink and uplink). For the evaluation purposes of the proposed algorithms random values are generated for latency, user throughput (downlink and uplink), and on the basis of the user throughput randomly is generated the power consumed by the mobile user device and then the energy efficiency (downlink and uplink) is being calculated.

Depending from the service type being requested by the 5G smart user device, the 5G smart user device selects one of the five algorithms previously described, and then performs an optimal selection of 5G radio access network on the basis of the selected algorithm.

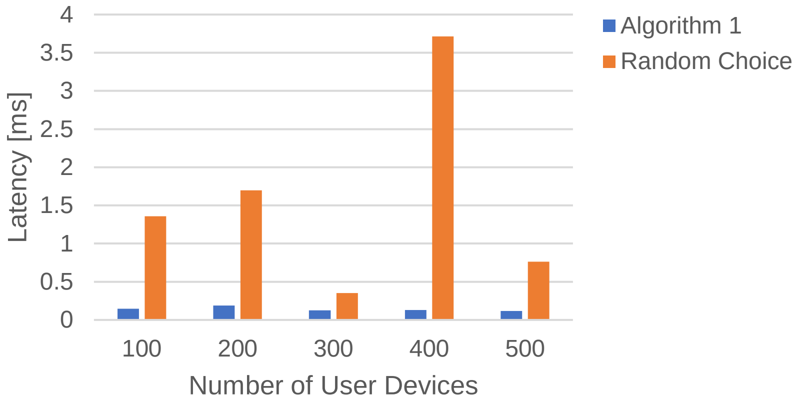
This simulation scenario can be easily constructed in MATLAB. The obtained results represent an average value of 10 simulations.

9.4 Analysis and Evaluation of the Obtained Results according to Algorithm 1

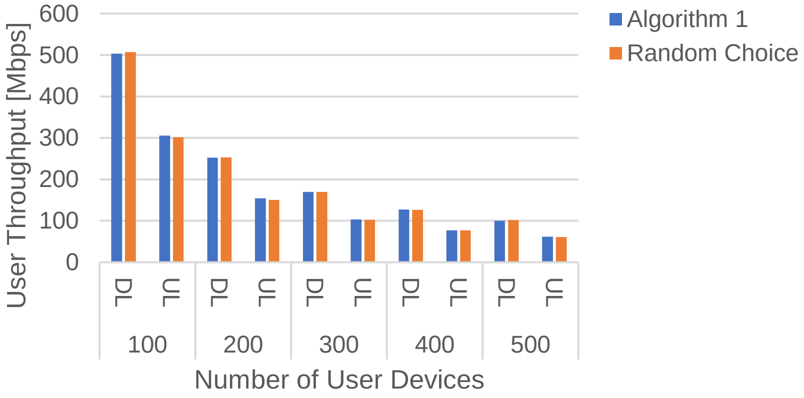
The results obtained from Algorithm 1 are given on the Figures below as follows:

- A comparison of the obtained results for latency, user throughput, and energy efficiency in 5G mobile networks in a cloud and fog computing environment as a function of the number of user devices at a given distance on Figure 9.2; and
- A comparison of the obtained results for latency, user throughput, and energy efficiency in 5G mobile networks in a cloud and fog computing environment as a function of the distance for a given number of user devices on Figure 9.3.

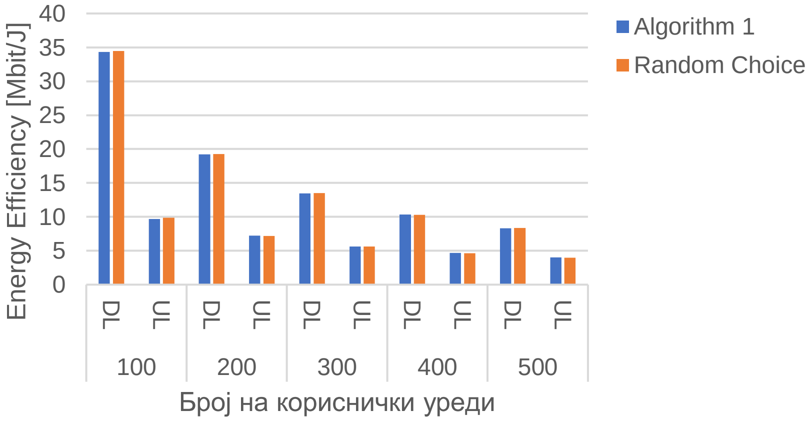
It can be noticed that lower latency is obtained when algorithm 1 is applied, rather than when it is not applied. In addition, in some cases better, or similar values are obtained for the user throughput (downlink and uplink) and the energy efficiency (downlink and uplink).



a) Latency

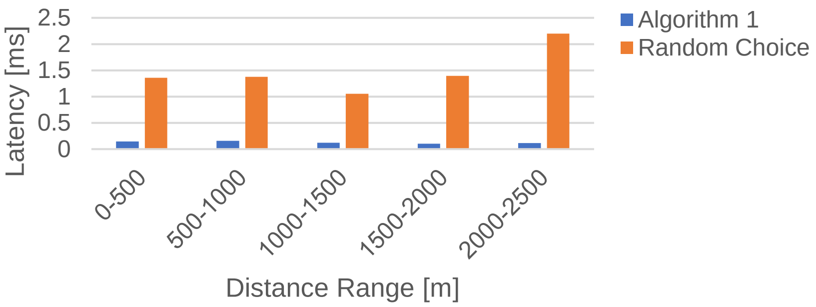


b) User Throughput

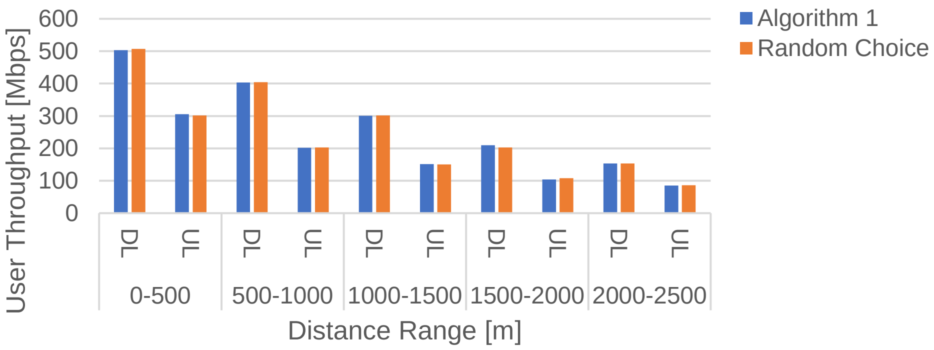


c) Energy Efficiency

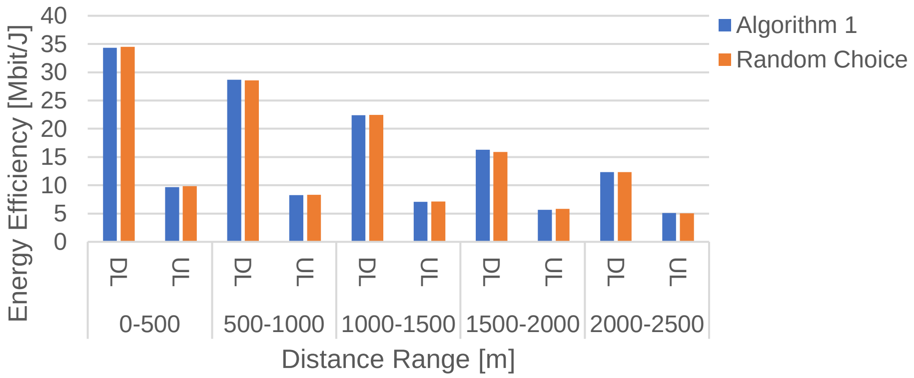
Figure 9.2 A Comparison of Latency, User Throughput and Energy Efficiency in 5G Mobile Networks in Cloud and Fog Computing Environment as a Function of the Number of User Devices at a Distance from 0 to 500 meters, with and without application of Algorithm 1



a) Latency



b) User Throughput



c) Energy Efficiency

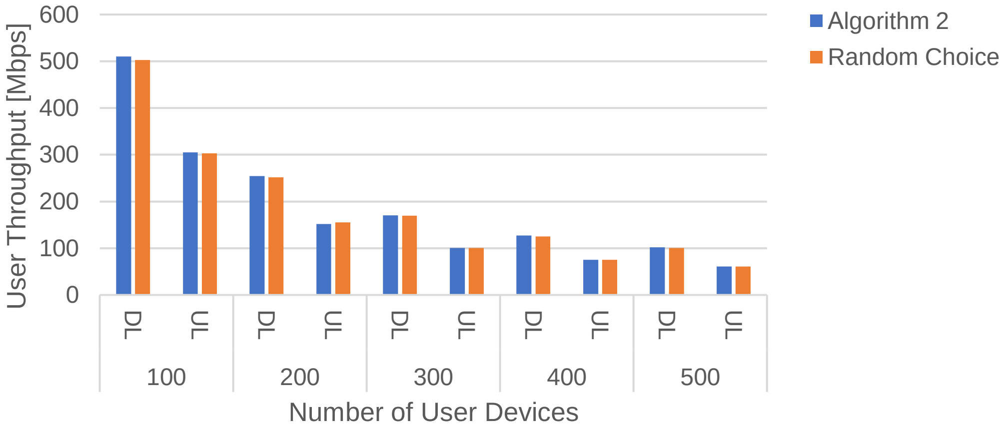
Figure 9.3 A Comparison of Latency, User Throughput and Energy Efficiency in 5G Mobile Networks in Cloud and Fog Computing Environment as a Function of the Distance for 100 User Devices, with and without application of Algorithm 1

9.5 Analysis and Evaluation of the Obtained Results according to Algorithm 2

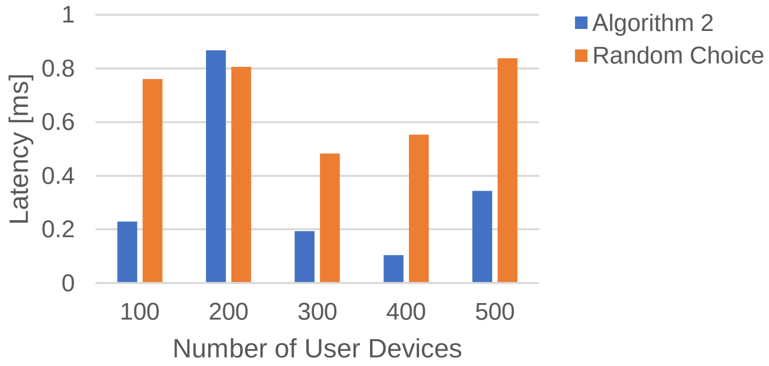
The results obtained from Algorithm 2 are given on the Figures below as follows:

- A comparison of the obtained results for latency, user throughput, and energy efficiency in 5G mobile networks in a cloud and fog computing environment as a function of the number of user devices at a given distance on Figure 9.4; and
- A comparison of the obtained results for latency, user throughput, and energy efficiency in 5G mobile networks in a cloud and fog computing environment as a function of the distance for a given number of user devices on Figure 9.5.

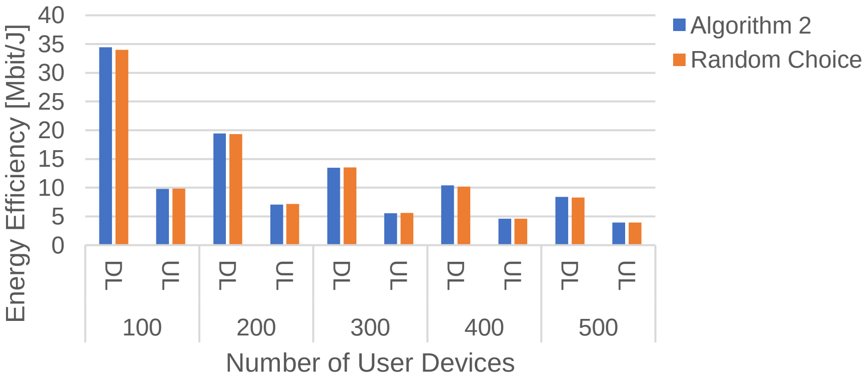
It can be noticed that higher user throughput in downlink direction is obtained when algorithm 2 is applied, rather than when it is not applied. In addition, in some cases better, or similar values are obtained for the latency, user throughput (uplink) and the energy efficiency (downlink and uplink).



a) User Throughput

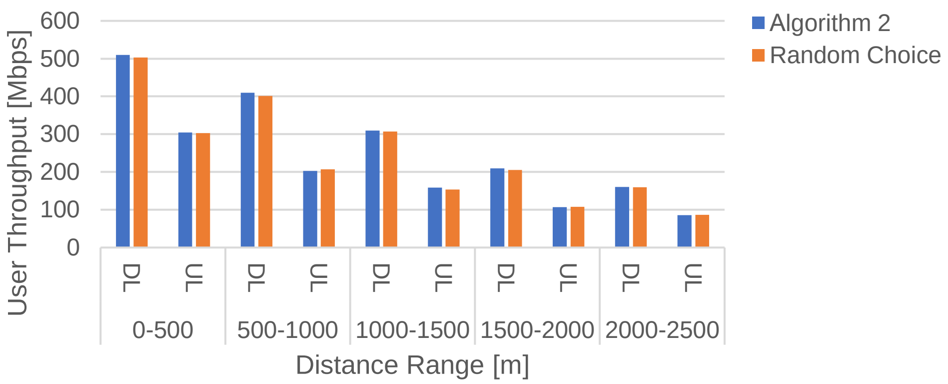


b) Latency

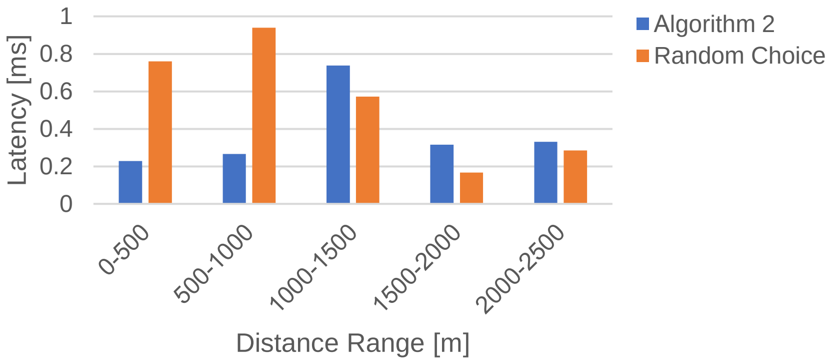


c) Energy Efficiency

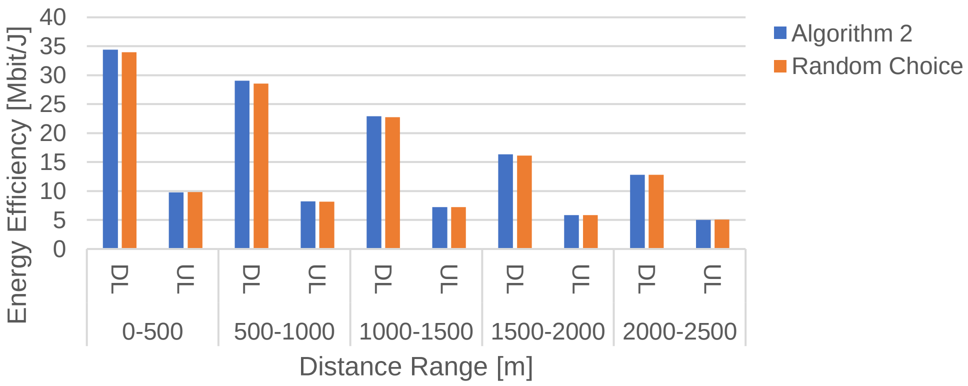
Figure 9.4 A Comparison of Latency, User Throughput and Energy Efficiency in 5G Mobile Networks in Cloud and Fog Computing Environment as a Function of the Number of User Devices at a Distance from 0 to 500 meters, with and without application of Algorithm 2



a) User Throghput



b) Latency



c) Energy Efficiency

Figure 9.5 A Comparison of Latency, User Throughput and Energy Efficiency in 5G Mobile Networks in Cloud and Fog Computing Environment as a Function of the Distance for 100 User Devices, with and without application of Algorithm 2

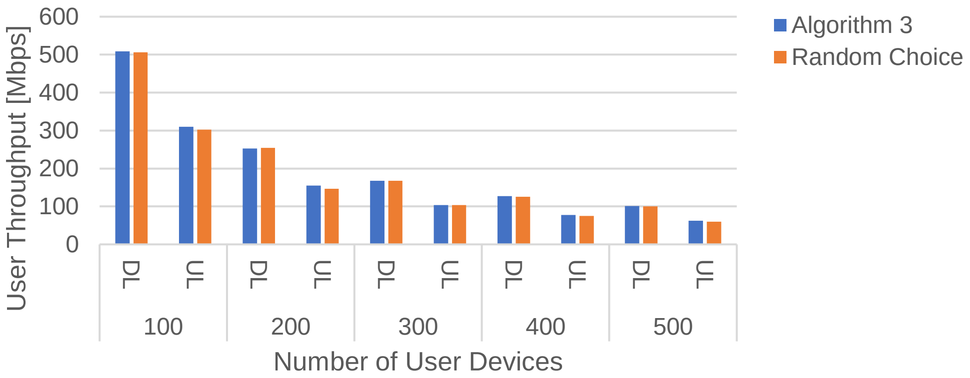
9.6 Analysis and Evaluation of the Obtained Results according to Algorithm 3

The results obtained from Algorithm 3 are given on the Figures below as follows:

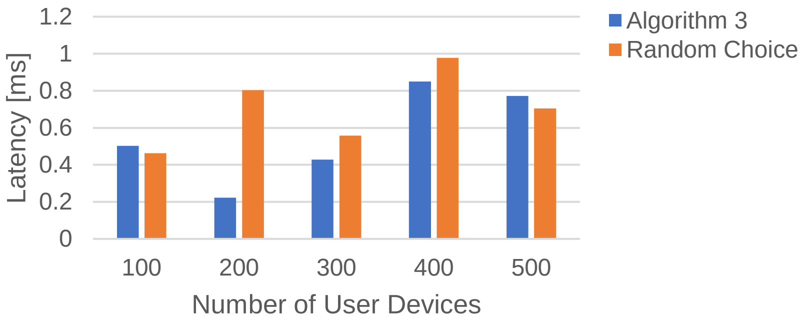
- A comparison of the obtained results for latency, user throughput, and energy efficiency in 5G mobile networks in a cloud and fog computing environment as a function of the number of user devices at a given distance on Figure 9.6; and

- A comparison of the obtained results for latency, user throughput, and energy efficiency in 5G mobile networks in a cloud and fog computing environment as a function of the distance for a given number of user devices on Figure 9.7.

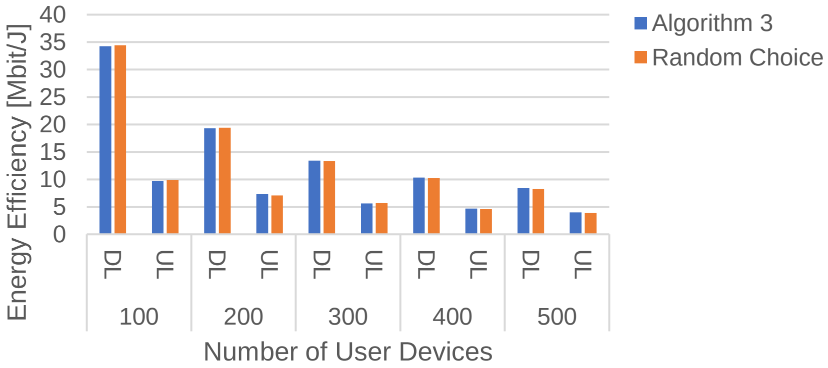
It can be noticed that higher user throughput in uplink direction is obtained when algorithm 3 is applied, rather than when it is not applied. In addition, in some cases better, or similar values are obtained for the latency, user throughput (downlink) and the energy efficiency (downlink and uplink).



a) User Throughput

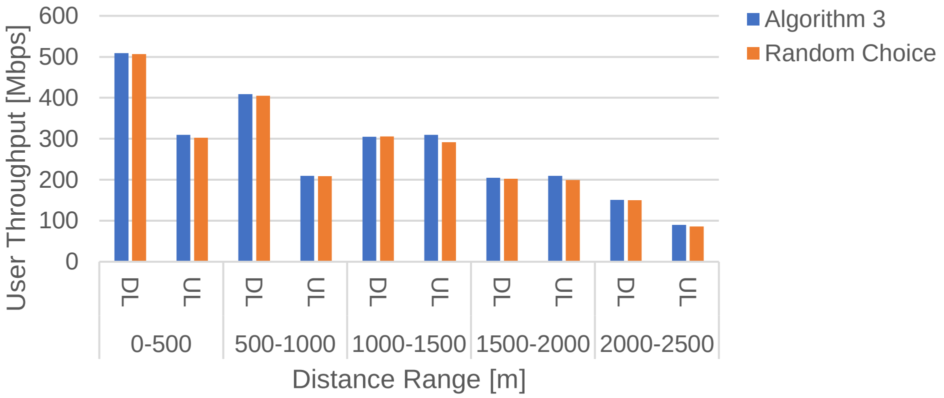


b) Latency

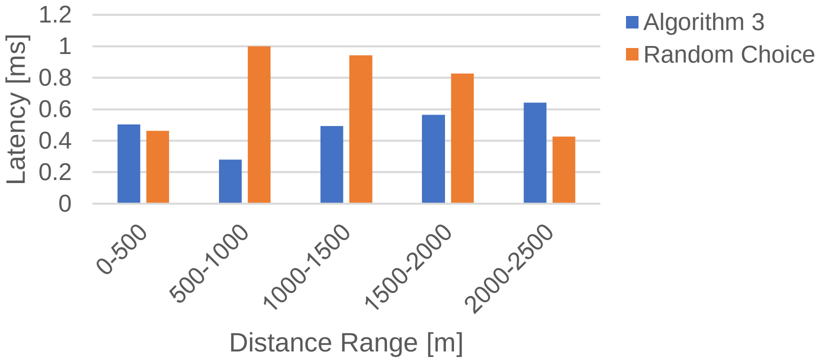


c) Energy Efficiency

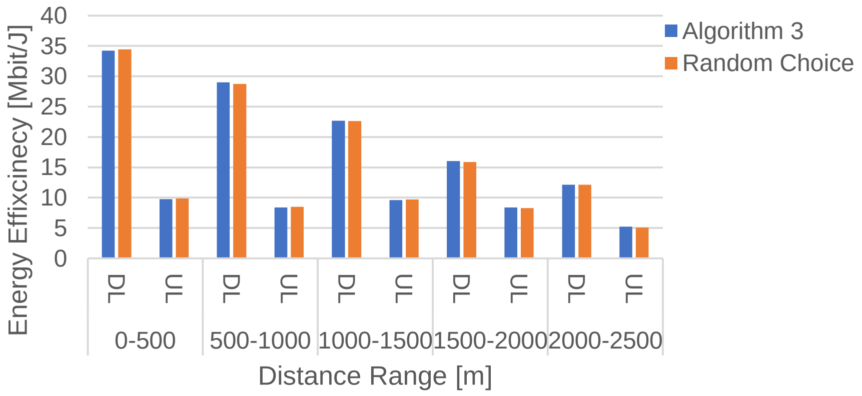
Figure 9.6 A Comparison of Latency, User Throughput and Energy Efficiency in 5G Mobile Networks in Cloud and Fog Computing Environment as a Function of the Number of User Devices at a Distance from 0 to 500 meters, with and without application of Algorithm 3



a) User Throughput



b) Latency



c) Energy Efficiency

Figure 9.7 A Comparison of Latency, User Throughput and Energy Efficiency in 5G Mobile Networks in Cloud and Fog Computing Environment as a Function of the Distance for 100 User Devices, with and without application of Algorithm 3

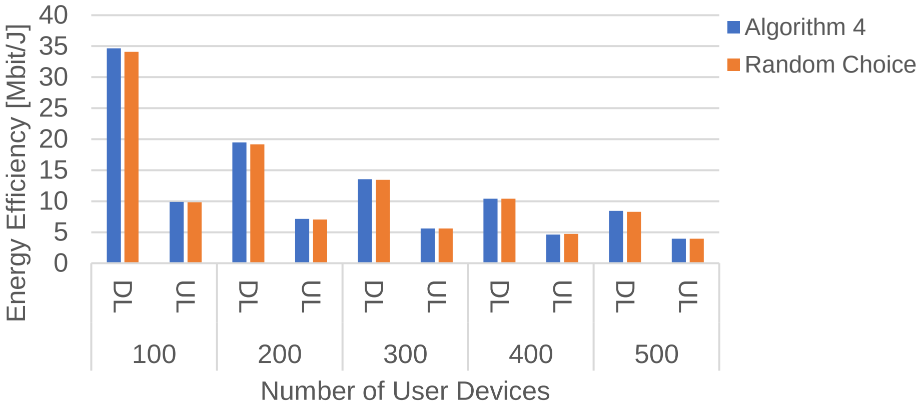
9.7 Analysis and Evaluation of the Obtained Results according to Algorithm 4

The results obtained from Algorithm 4 are given on the Figures below as follows:

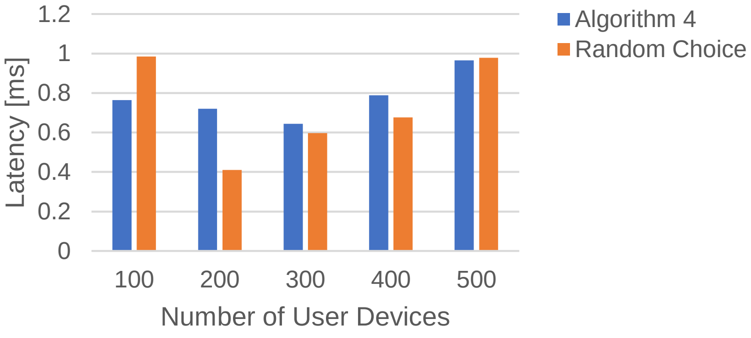
- A comparison of the obtained results for latency, user throughput, and energy efficiency in 5G mobile networks in a cloud and fog computing environment as a function of the number of user devices at a given distance on Figure 9.8; and

- A comparison of the obtained results for latency, user throughput, and energy efficiency in 5G mobile networks in a cloud and fog computing environment as a function of the distance for a given number of user devices on Figure 9.9.

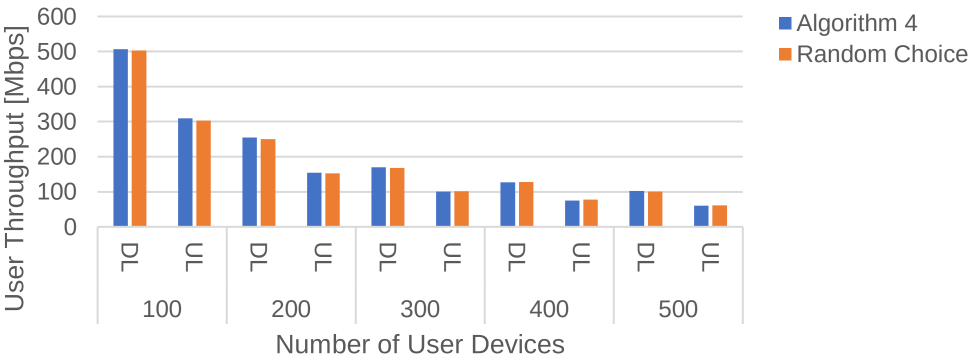
It can be noticed that better energy efficiency in downlink direction is obtained when algorithm 4 is applied, rather than when it is not applied. In addition, in some cases better, or similar values are obtained for the latency, user throughput (downlink and uplink) and the energy efficiency (uplink).



a) Energy Efficiency

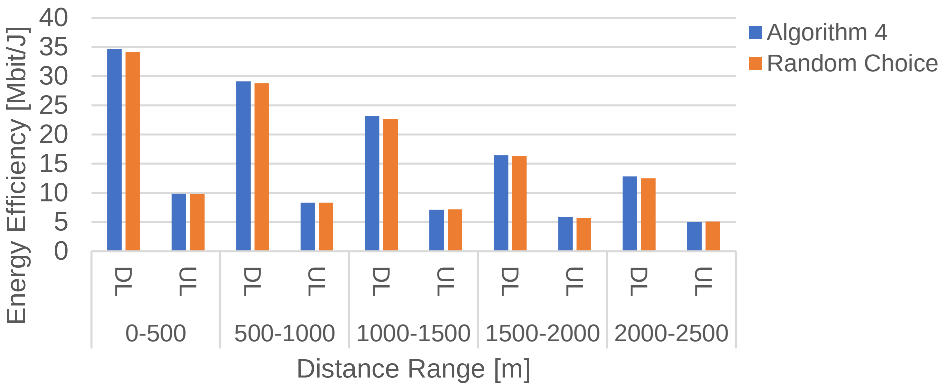


b) Latency

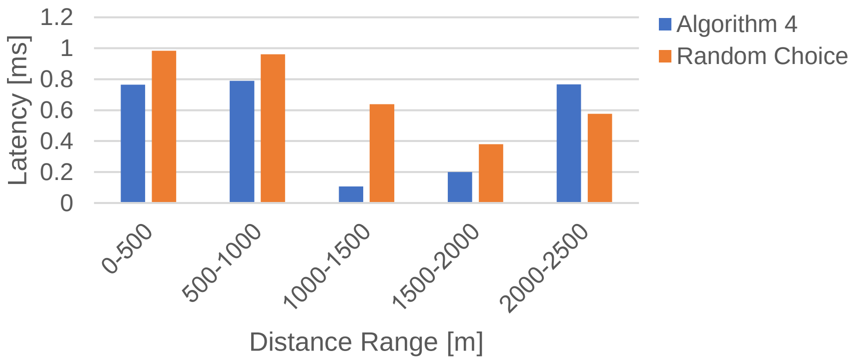


c) User Throughput

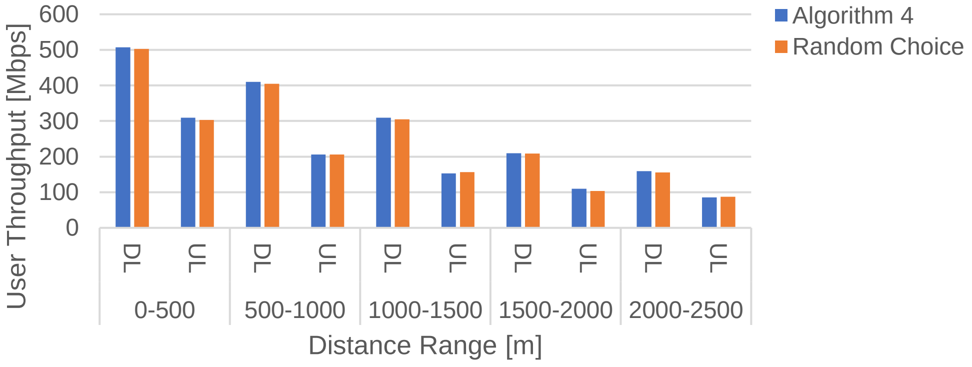
Figure 9.8 A Comparison of Latency, User Throughput and Energy Efficiency in 5G Mobile Networks in Cloud and Fog Computing Environment as a Function of the Number of User Devices at a Distance from 0 to 500 meters, with and without application of Algorithm 4



a) Energy Efficiency



b) Latency



C) User Throughput

Figure 9.9 A Comparison of Latency, User Throughput and Energy Efficiency in 5G Mobile Networks in Cloud and Fog Computing Environment as a Function of the Distance for 100 User Devices, with and without application of Algorithm 4

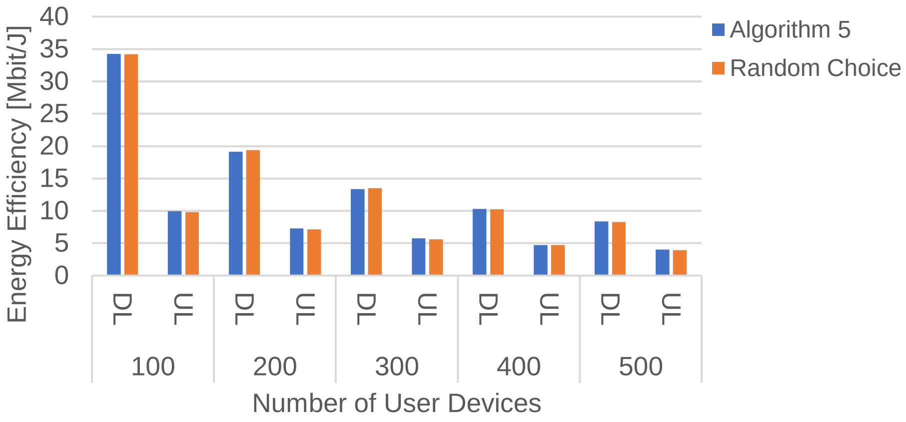
9.8 Analysis and Evaluation of the Obtained Results according to Algorithm 5

The results obtained from Algorithm 4 are given on the Figures below as follows:

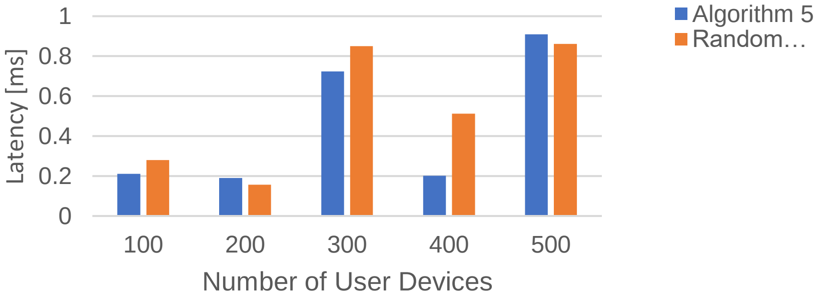
- A comparison of the obtained results for latency, user throughput, and energy efficiency in 5G mobile networks in a cloud and fog computing environment as a function of the number of user devices at a given distance on Figure 9.10; and

- A comparison of the obtained results for latency, user throughput, and energy efficiency in 5G mobile networks in a cloud and fog computing environment as a function of the distance for a given number of user devices on Figure 9.11.

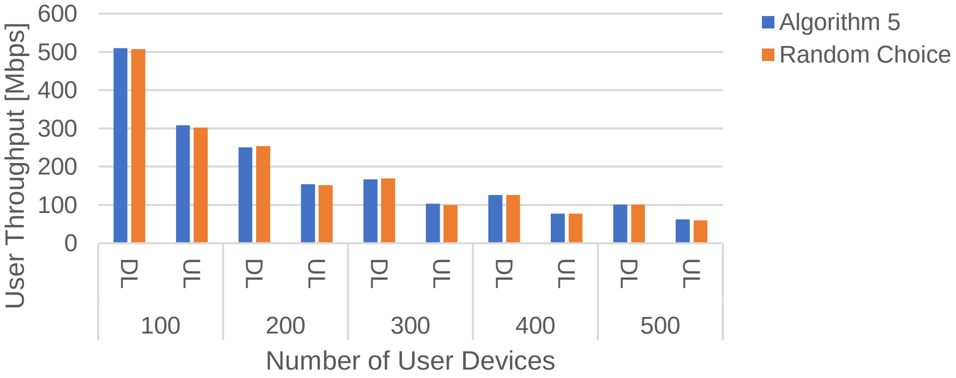
It can be noticed that better energy efficiency in uplink direction is obtained when algorithm 4 is applied, rather than when it is not applied. In addition, in some cases better, or similar values are obtained for the latency, user throughput (downlink and uplink) and the energy efficiency (downlink).



a) Energy Efficiency

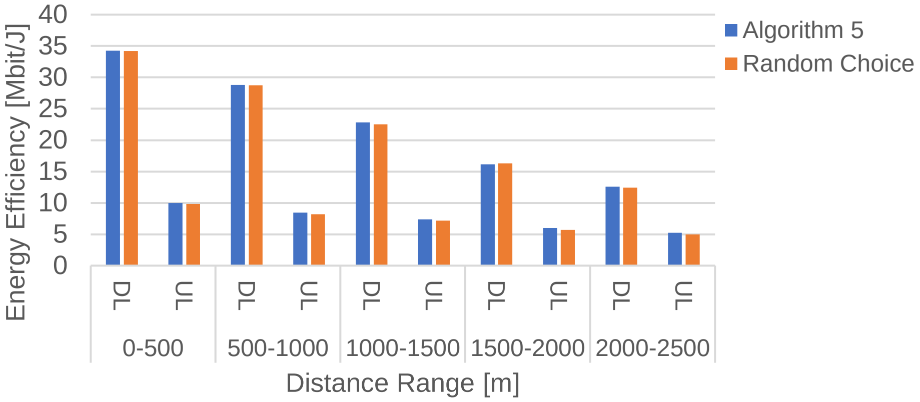


b) Latency

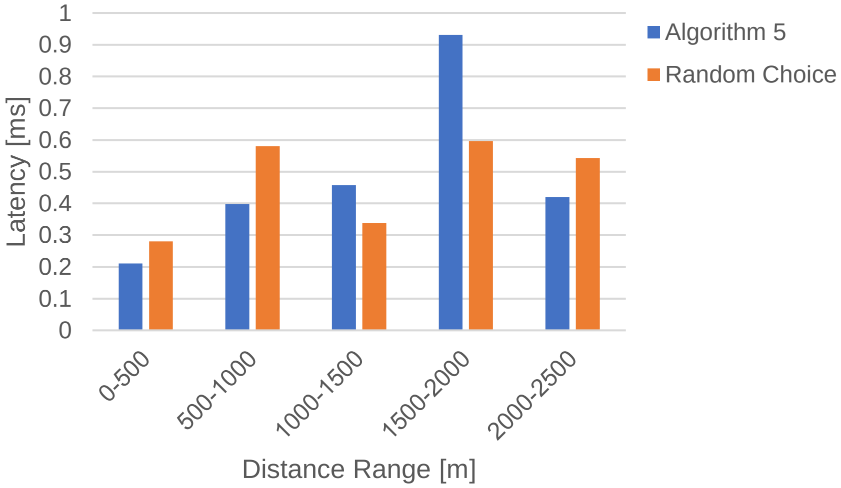


c) User Throughput

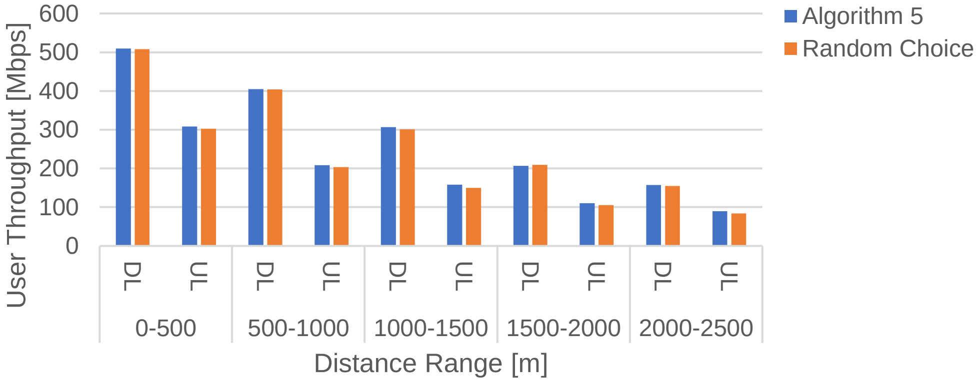
Figure 9.10 A Comparison of Latency, User Throughput and Energy Efficiency in 5G Mobile Networks in Cloud and Fog Computing Environment as a Function of the Number of User Devices at a Distance from 0 to 500 meters, with and without application of Algorithm 5



a) Energy Efficiency



b) Latency



c) User Throughput

Figure 9.11 Comparison of Latency, User Throughput and Energy Efficiency in 5G Mobile Networks in Cloud and Fog Computing Environment as a Function of the Distance for 100 User Devices, with and without application of Algorithm 4

9.9 Discussion of the Results

The optimal selection of a radio access network primarily depends from the service requirements of data being transferred such as user throughput (downlink and uplink), latency (end-to-end delay) and energy efficiency (downlink and uplink).

This chapter presented 5 algorithms for an optimal selection of 5G RAN, 5G base station. During the evaluation it was made a comparison of the obtained results when each algorithm is applied and the obtained results when each algorithm is not applied. Here:

- With algorithm 1 is achieved lowest latency;
- With algorithm 2 is achieved the highest downlink throughput;
- With algorithm 3 is achieved the highest uplink throughput;
- With algorithm 4 is achieved the highest energy efficiency in downlink direction; and
- With algorithm 5 is achieved the highest energy efficiency in uplink direction.

Which algorithm would be applied, depends primarily from the user preferences and the user needs, as well as the Service Level Agreement (SLA) between the user and the service provider.

Part Three

10 Conclusion

5G mobile networks from the user perspective, should support [52 - 53], [123]:

- 1000 times greater volume of mobile data services per area;
- 10 to 100 times more connected devices; and
- 10 to 100 times more higher typical data rates per user.

In 2017 already are announced the 5G specifications [124]:

- Peak data rate of 20 Gbps in downlink direction;
- Peak data rate of 10 Gbps in uplink direction;
- Average latency of 1ms (even lower), and the maximum latency should not exceed above 4 ms;
- Support of 1 million devices per square kilometer;
- Every frequency carrier should have at least 100 MHz available bandwidth, with a possibility to be extended to 1GHz wherever it is possible;
- 5G base stations should provide support to the user that moves with a speed from 0 km/h to 500 km/h.
- Peak spectral efficiency of 30 bits per second per hertz in downlink direction;
- Peak spectral efficiency of 15 bits per second per hertz in uplink direction;
- User data rate of 100 Mbps in downlink direction; and
- User data rate of 50 Mbps in uplink direction.

In this final chapter, are pointed out the main conclusions and the contributions of the doctoral dissertation that relate to the key identifying problems according to the research questions that were addressed at the Introduction part of this PhD dissertation.

10.1 Concluding Remarks

The main goal of this doctoral dissertation was to perform an evaluation of the quality of cloud and fog orchestrated services in 5G mobile networks, as well as to propose an algorithm for an optimal selection of 5G base station in order to satisfy the demands of 5G smart user devices, that have increased demands, which results in rapid increase of the traffic volume.

With this doctoral dissertation was proved and analyzed that it is possible to achieve and provide the necessary QoS of present and future 5G mobile and wireless networks and services.

First of all was considered the cloud computing environment in 5G mobile networks, because the best results can be achieved if the analysis is started from the advantages and disadvantages of the knowledge that already exists, and which must be deepened and advanced, or at least to use some inspiring ideas from it, in order to discover something new and better. Then the new fog computing paradigm for 5G was considered.

The basic goal of the research in this doctoral dissertation was to check the quality of cloud and fog orchestrated services in 5G mobile networks, in the direction of improving the QoS parameters, such as latency, throughput and energy efficiency. Therefore, it was initially proposed a model and architecture for hybrid orchestration of the services in a fog computing environment in 5G mobile networks, and then it was performed an evaluation of QoS parameters. In addition, 5 possible algorithms for an optimal selection of 5G radio access network on the basis of the user service requirements of these QoS parameters were proposed.

The basic task of the purpose was to answer on the following questions:

1. How to implement the protocol stack layer and the advanced intelligence of orchestrating mechanisms in a fog computing environment in the future mobile and wireless networks?

Answer: By using the **hybrid environment service orchestrator**, that consists of several levels: regional, domain and federated service orchestrators, that enable distribution of the cloud through all entities in 5G network (from the core, to the radio access network, and up to the 5G smart user device).

2. How to get a better quality of service in the future 5G mobile and wireless networks (for any service type)?
   Answer: With a joint implementation of the cloud and the fog in 5G mobile networks. Through the evaluation of QoS parameters it was confirmed that better results are achieved in terms of latency, throughput and energy efficiency by implementing the fog computing environment in 5G networks.
3. How to build the new mobile and wireless IP networks in order data processing to be executed in a fog computing environment?
   Answer: The answer is again with using the **hybrid environment service orchestrator,** as well as the implementation of five proposed algorithms for optimal selection of 5G base station, with which is achieved lower latency, higher data rates, and better energy efficiency.
4. Where the fog computing orchestrating mechanisms should be positioned in the 5G network architecture in order to have a better quality of orchestrated services?
   Answer: By implementing the **hybrid environment service orchestrator** in 5G, the cloud computing orchestrating mechanisms are actually distributed across the whole 5G network architecture starting from the 5G core, then the radio access network and up to the 5G smart user device.

Further in the introduction part was given the following **main hypothesis**:

The introduction of Fog Computing Service Orchestration Mechanisms contributes to a higher satisfactory level of all QoS parameters and obtains higher estimations of QoE parameters of any multimedia services.

And the following two special thesis:

1) *The introduction of Fog Computing Service Orchestration Mechanisms in 4G homogeneous mobile and wireless IP networks contributes to a higher satisfactory level of all QoS parameters for a given multimedia service.*
2) *The introduction of Fog Computing Service Orchestration Mechanisms in 5G heterogeneous mobile and wireless IP networks contributes to a higher satisfactory level of QoS parameters any given multimedia services.*

The main hypothesis and the special thesis where checked through the evaluation of QoS parameters (latency, throughput, and energy efficiency) in Chapter 8, with which actually it was confirmed that the fog computing environment has better performances than the cloud computing environment independently of the generation of mobile network.

Through the performed evaluation of QoS parameters in Chapter 8 was concluded if the fog computing environment is implemented independently in any mobile network gives better results than the cloud computing in terms of latency, throughput and energy efficiency. In addition, 5G in the cloud and fog computing environment has much better performances than 3G and 4G networks. The analysis demonstrate that compared to the cloud, the fog computing environment provides higher level of satisfied users for any multimedia service, highest level, or rate of multimedia service access, broadband data rates, optimal usage of all resources from the available radio access networks, minimum cost per service and intelligent balancing of traffic load trough the various radio access technologies.

Finally, in Chapter 9 were demonstrated 5 possible algorithms for optimal selection of 5G base station. The selection of such radio access network primarily depends from the service requirements of the data being transferred such as latency (end-to-end delay), throughput (downlink and uplink) and energy efficiency (downlink and uplink), etc. Which algorithm would be applied depends from the user needs and user preferences, as well as the Service Level Agreement (SLA), that exists between the user and the Service provider.

As a conclusion the cloud in 5G network would be diffused among the client devices often with mobility too, i.e. the cloud would become fog [56]. More and more virtual network functionality would be executed in a fog computing environment, which would provide ***mobiquitous*** service to the users. This would enable new services paradigms such as Anything as a Service (AaaS) where devices, terminals, machines, and also smart things and robots would become innovative tools that would produce and use applications, services and data.

10.2 Contributions of the Doctoral Dissertation and Application of the Research Results

This doctoral dissertation gives the following contributions:

1. It is proposed a novel 5G mobile network architecture in cloud and fog computing environment with advanced cellular RAT that are added advanced Wireless Local Area Networks – WLAN or WiFi and active distant nodes on the backhaul connection from the radio network up to the fundamental network;
2. It is offered one advanced mechanism for the quality of the cloud orchestrated services at the existing and future 5G heterogeneous mobile and wireless networks, with whom it is achieved best cost perfomances, it is selected the best network, or networks for a given multimedia service in order to provide an excellent quality of service at a maximum level. In that way it is resulted with a higher QoS/QoE level at all users, i.e. better satisfaction of all users for all the services that are being used by them;
3. Furthermore, the contribution is in providing user-oriented network where the user does not need to know through which network is performed the orchestration and distribution of the services with their connected mobile device. These orchestration mechanisms are opened and independent from all possible existing and future mobile and wireless networks. Therefore, if such service orchestrating mechanism is implemented in 5G networks, it would have a possibility to use any radio access technology, because it is defined independently from the technologies on the network IP layer. Finally, this mechanism is scalable and flexible, because it can be easily implemented to any new radio interface and new radio access technology.
4. Also with these service orchestration mechanisms, one smart 5G mobile device that would combine and use aggregated data traffic from all available radio access networks for all used multimedia services, easily can achieve the planned 5G aggregated data rates (from the order of 50 Gbps for low mobility, 10 Gbps for high mobility everywhere including the edge of the network). In this way this doctoral dissertation would contribute for the high level of QoS/QoE support of all services, high data rates, improved mobility, and omnipresence of the future 5G mobile networks, delivering the communication on the whole planet, on one truly higher level and higher quality.
5. It is achieved better energy efficiency, which results with longer battery life of the mobile devices, which contributes to higher satisfaction of mobile device users. Also, the improved energy efficiency has a positive impact on the reduction of the operating costs for the network management at the network operators.

The obtained results in this doctoral dissertation for the quality of cloud orchestrated services in 5G mobile network, have a broad scientific application, and may have an influence these cloud and fog computing orchestrating mechanisms to enter in the final specifications in the standard of 5G mobile networks, or at least to participate in the creation of the logics, the paradigm and the architecture of 5G mobile networks.

References

[1] F. Boccardi; et al., “Five Disruptive Technology Directions for 5G,” *IEEE Communications Magazine*, vol. 52, no. 2, pp. 74-80, 2014.

[2] N. Bhushan; et al., “Network Densification: The Dominant Theme for Wireless Evolution into 5G,” *IEEE Communications Magazine,* vol. 52, no. 2, pp. 82-89, 2014.

[3] B. Bangerter, S. Talwar, R. Arefi and K. Stewart, “Networks and Devices for the 5G Era,” *IEEE Communications Magazine*, vol. 52, no. 2, pp. 90-96, 2014.

[4] C. Wang; et al., “Cellular Architecture and Key Technologies for 5G Wireless Communication Networks,” *IEEE Communications Magazine,* vol. 52, no. 2, pp. 122-130, 2014.

[5] T. Janevski, “5G Mobile Phone Concept,” in *6th IEEE Consumer Communications and Networking Conference (CCNC)*, Las Vegas, USA, 2009.

[6] W. W. Lu, “An Open Baseband Processing Architecture for Future Mobile Terminals Design,” *IEEE Wireless Communications,* vol. 15, no. 2, pp. 110-119, 2008.

[7] A. Tudzarov and T. Janevski, “Design for 5G Mobile Network Architecture,” *International Journal of Communication Networks and Information Security (IJCNIS),* vol. 3, no. 2, pp. 112-123, 2011.

[8] M. Josef Noll and M. Chowdhury, “5G – Service Continuity in Heterogeneous Environments,” *Wireless Personal Communications*, vol. 57, no. 3, pp. 413-429, 2011.

[9] T. Janevski, *NGN Architectures, Protocols, and Services,* John Wiley and Sons, 2014.

[10] T. Janevski, *Internet Technologies for Fixed and Mobile Networks*, Artech House, 2015.

[11] A. Tudzarov and T. Janevski, “Functional Architecture for 5G Mobile Networks,” *International Journal of Advanced Science and Technology (IJAST),* vol. 32, pp. 65-78, 2011.

[12] V. Tikhvinskiy and G. Bochechka, “Prospects and QoS Requirements in 5G Networks,” *Journal of Telecommunications and Information Technologies,* vol. 1, no. 1, pp. 23-26, 2015.

[13] S. Zhang, S. Zhang, X. Chen and X. Huo, “Cloud Computing, Research and Development Trend,” in *Proceedings of the 2nd IEEE International Conference on Future Networks*, Sanya, Hainan, 2010.

[14] ITU-T, “Cloud Computing Framework and High Level Requirements,” in *ITU-T Recommendation Y.3501,* 2013.

[15] ITU-T, “Cloud Computing Infrastructure Requirements,” ITU-T Recommendation Y.3510, 2013.

[16] NIST, “NIST Cloud Computing Standards Roadmap,” NIST Special Publication 500-291, Version 2, 2013.

[17] T. H. Dihn, C. Lee, D. Niyato and P. Wang, “A Survey of Mobile Cloud Computing: Architecture, Applications, and Approaches,” *Wireless Communications and Mobile Computing, Wiley,* vol. 13, no. 18, p. 1587–1611, 2011.

[18] S. S. Qureshi, T. Ahmad and K. R. Shuja-ul-islam, “Mobile Cloud Computing as Future for Mobile Applications – Implementation Methods and Challenging Issues,” in *Proceedings of IEEE Conference on Cloud Computing and Intelligence Systems (CCIS),* Beijing, China, 2011.

[19] D. Huang; et al, “Mobile cloud computing,” *IEEE COMSOC Multimedia Communications Technical Committee (MMTC) E-Letter,* vol. 6, no. 10, p. 27–31, 2011.

[20] S. Kitanov and T. Janevski, “State of the Art: Mobile Cloud Computing,” in *Proceedings of the Sixth IEEE International Conference on Computational Intelligence, Communication Systems and Networks 2014 (CICSYN 2014),* Tetovo, Macedonia, 2014.

[21] M. Armbrust; et al, “A view of cloud computing,” *Communications of the ACM,* vol. 53, no. 4, p. 50–58, 2010.

[22] N. Liu, X. Li and Q. Wang, “A resource & capability virtualization method for cloud manufacturing systems,” in *IEEE International Conferference on Systems, Man, and Cybernetics,* Anchorage, USA, 2011.

[23] B. Sotomayor, R. S. Montero, I. M. Llorente and I. Foster, “Virtual infrastructure management in private and hybrid clouds,” *IEEE Internet Computing,* vol. 13, no. 5, p. 14–22, 2009.

[24] F. Bonomi, R. Milito, J. Zhu and S. Addepalli, “Fog Computing and its Role in the Internet of Things,” in *Proceedings of the First Edition of the ACM SIGCOMM Workshop on Mobile Cloud Computing (MCC 2012),* Helsinki, Finland, 2012.

[25] L. M. Vaquero and L. Rodero-Merino, “Finding your Way in the Fog: Towards a Comprehensive Definition of Fog Computing,” *ACM SIGCOMM Computer Communication Review Newsletter,* vol. 44, no. 5, pp. 27-32, 2014.

[26] H. T. Luan, L. Gao, Z. Li and L. X. Y. Sun, “Fog Computing: Focusing on Mobile Users at the Edge,” *arXiv preprint arXiv:1502.01815[cs.NI],* 2015.

[27] I. Stojmenovic and S. Wen, “The Fog Computing Paradigm: Scenarios and Security Issues,” in *Proceedings of the Federated Conference on Computer Science and Information Systems (FedCSIS), ACSIS,* Warsaw, Poland, 2014.

[28] B. Brech, J. Jamison, L. Shao and G. Wightwick, “The Interconnecting of Everything,” IBM Corp., 2013.

[29] V. L. Kalyani and D. Sharma, “IoT: Machine to Machine (M2M), Device to Device (D2D) Internet of Everything (IoE) and Human to Human (H2H): Future of Communication,” *Journal of Management Engineering and Information Technology (JMEIT),* vol. 2, no. 6, pp. 17-23, 2015.

[30] M. Chiang, “Fog Networking: An Overview on Research Opportunities,” White Paper, 2015.

[31] M. Yannuzzi; R. Milito; R. Serral-Gracia; et al., “Key ingredients in an IoT recipe: Fog Computing, Cloud computing, and more Fog Computing,” in *IEEE 19th International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD),* Athens, Greece, 2014.

[32] Cisco, “Visual Networking Index: Global Mobile Data Traffic Forecast Update 2014 - 2019,” Cisco White Paper, 2015.

[33] Cisco, “Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2016–2021,” Cisco White Paper, 2017.

[34] E. Hedin, “Social Networks and Mobile Devices, ” Stockholm, Sweden: Master of Science Thesis at KTH Information and Communication Technology, 2008.

[35] Ericsson, “Ericsson Mobility Report on the Pulse of the Networked Society,” Ericsson, Stockholm, Sweden, 2016.

[36] M. M. ud in Mir and S. Kumar, “Evolution of Mobile Wireless Technology from 0G to 5G,” *International Journal of Computer Science and Information Technologies (IJCSIT),* vol. 6, no. 3, pp. 2545-2551, 2015.

[37] T. Shuminoski, "QoS Mechanisms and Vertical Multi-Homing for 5G Mobile and Wireless Networks", Skopje, Republic of Macedonia: Doctoral Dissertation, Ss Cyril and Methodius University, Faculty of Electrical Engineering and Information Technologies, 2016.

[38] M. Rahnema, "Overview of the GSM System and Protocol Architecture," *IEEE Communications Magazine* , vol. 31, no. 4, pp. 92-100, 1993.

[39] J. Cai and D. J. Goodman, "General Radio Packet Service in GSM," *IEEE Communications Magazine* , vol. 35, no. 10, pp. 122-131, 1997.

[40] K. Zangi, A. Furuskar and M. Hook, "EDGE: Enhanced Data Rates for Global Evolution of GSM and IS-136," in *Multi Dimensional Mobile Communications (MDMC 1998)*, 1998.

[41] H. Axelsson, P. Björkén, P. d. Bruin, S. Eriksson and H. Persson, "GSM/EDGE continued evolution," Ericsson Review No. 1, 2006.

[42] K. Richardson, "UMTS overview," *Electronics & Communication Engineering Journal*, vol. 12, no. 3, pp. 93-100, 2000.

[43] H. Kaaranen, et al. eds. , UMTS Networks: Architecture, Mobility and Services. 2nd ed., England: John Wiley & Sons Ltd., 2005.

[44] Risavy Research Report, "EDGE, HSPA, and LTE: Broadband Evolution," 3G Americas, 2008.

[45] S. Kitanov and T. Janevski, "Modeling and Analyzing LTE Networks with EstiNet Network Simulator and Emulator," in *XVIII International Scientific Conference on Information, Communication and Energy Systems and Technologies (ICEST)*, Ohrid, Republic of Macedonia, 2013.

[46] S. Kitanov and T. Janevski, "Performance Evaluation of Scheduling Strategies for LTE Networks in Downlink Direction," in *XI International Conference of Society for Electronics, Telecommunications, Automation and Informatics ETAI 2013*, Ohrid, Republic of Macedonia, 2013.

[47] S. Sesia, I. Toufik and M. Baker, *LTE – The UMTS Long Term Evolution From Theory to Practice 2nd Edition including Release 10 for LTE Advanced, 2nd ed.*, John Wiley & Sons Ltd., 201.

[48] S. G. Glisic, *Advanced Wireless Networks 4G Technologies*, 1st ed., University of Oulu, Finland: John Wiley & Sons Ltd., 2006.

[49] M. Cudak, "IEEE 802.16m System Requirements," IEEE 802.16 Broadband Wireless Access Working Group , 2008.

[50] Nokia, "LTE-Advanced Pro Pushing LTE capabilities towards 5G," Nokia Networks white paper , 2015.

[51] Nokia, "LTE-M – Optimizing LTE for the Internet of Things," Nokia White Paper , 2015.

[52] Datang Telecom Technology and Industry Group, "Evolution, Convergence and Innovation," Datang Mobile Wireless Innovation Center 5G White Paper, 2013.

[53] SK Telecom, "SK Telecom's View on 5G Vision, Architecture, Technology and Service and Spectrum," SK Telecom Network Technology Research and Development Center 5G White Paper, 2014.

[54] S. Kitanov and T. Janevski, "Fog Computing Mechanisms in 5G Mobile Networks," in *Networks of the Future: Architectures, Technologies and Implementations*, Florida, USA, CRC Press Taylor and Francis Group, 2017.

[55] GSA - Global mobile Suppliers Association, “The Road to 5G: Drivers, Applications, Requirements and Technical Development,” A GSA (Global mobile Suppliers Association) Executive Report from Ericsson, Huawei and Qualcomm, 2015.

[56] S. Kitanov, E. Monteiro and T. Janevski, “5G and the Fog – Survey of Related Technologies and Research Directions,” Proceedings of the 18th Mediterranean IEEE Electrotechnical Conference MELECON 2016, Limassol, Cyprus, 2016.

[57] A. Tudzarov, “New Methods for Quality Control and Improvement of the IP based Services in Mobile and Wireless Networks,” Skopje, Republic of Macedonia: Doctoral Dissertation, Ss Cyril and Methodius University, Faculty of Electrical Engineering and Information Technologies, 2011.

[58] S. Barbarossa, S. Sardelliti and P. D. Lorenzo, “Communicating While Computing – Distributed Mobile Cloud Computing Over 5G Heterogeneous Networks,” *IEEE Signal Processing Magazine International Journal of Advanced Science and Technology (IJAST),* vol. 1, no. 6, pp. 45 - 55, 2014.

[59] D. Guinard, V. Trifa, F. Mattern and E. Wilde, “From the Internet of Things to the Web of Things: Resource-oriented Architecture and Best Practices,” in *Architecting the Internet of Things*, Berlin Heidelberg, Springer Verlag, 2011, pp. 97 - 129.

[60] ABI, Internet of Things vs. Internet of Everything What’s the Difference, ABI Research White Paper, 2014.

[61] S. Mitchell, N. Villa, M. Stewart-Weeks and A. Lange, “The Internet of Everything for Cities Connecting People, Process, Data, and Things to Improve the ‘Livability’ of Cities and Communities,” CISCO, white paper., 2013.

[62] D. Evans, “The Internet of Things: How the Next Evolution of the Internet Is Changing Everything,” Cisco White Paper, 2011.

[63] T. Harris, “Cloud Computing - An Overview”.

[64] Dialogic Corporation, “Introduction to Cloud Computing,” Dialogic Corporation White Paper, 2010.

[65] A. R. Khan, M.Othman, S. A. Madani and S. U. Khan, “Survey of Mobile Cloud Computing Application Models,” *IEEE,* vol. 16, no. 1, pp. 393-413, 2014.

[66] S. Fowler, “Survey on Mobile Cloud Computing Challenges Ahead,” *IEEE CommSoft E-Letters,* vol. 2, no. 1, pp. 13-17, 2013.

[67] S. Goyal, “Public vs Private vs Hybrid vs Community – Cloud Computing: A Critial Review,” *International Journal of Computer Network and Information Security(IJCNIS),* vol. 6, no. 3, pp. 20-29, 2014.

[68] ITU Centres of Excellence for Europe, “Module 4: Cloud Computing,” in *NGN, FUTURE NETWORKS AND CLOUD COMPUTING,* A Course provided by ITU Centres of Excellence for Europe, 2014.

[69] H. M. Sqalli, M. Al-saeedi, F. Binbeshr and M. Siddiqui, “UCloud: Simulated Hybrid Cloud for a University Environment,” in *IEEE 1st International Conf. on Cloud Networking (CLOUDNET),* Paris, 2012.

[70] O. Consortium, “OpenFog Reference Architecture for Fog Computing,” OpenFog Consortium Architecture Working Group, 2017.

[71] L. Gao, T. H. Luan, W. Z. B. Liu and S. Yu, “Fog Computing and Its Applications in 5G,” in *5G Mobile Communications,* Springer International Publishing Switzerland, 2017, pp. 571-593.

[72] S. Kitanov and T. Janevski, “Fog Networking for 5G and IoT,” in *5G Mobile: From Research and Innovations to Deployment Aspects, Architectures, Technologies and Implementations*, Nova Publishers, 2017.

[73] Z. Zhang, J. Zhang and L. Ying, “Multimedia streaming in cooperative mobile social networks,” Preprint..

[74] F. M. F. Wong, S. Ha, C. Joe-Wong, Z. Liu and M. Chiang, “Mind Your Own Bandwidth: Adaptive Traffic Management on Network Edge,” in *Proceedings of IEEE IWQoS*, 2015.

[75] Y. Du, E. Aryafar, J. Camp and M. Chiang, “iBeam: Intelligent client-side multi-user beamforming in wireless networks,” in *IEEE INFOCOM*, 2014.

[76] E. Aryafar, A. Keshavarz-Haddard, M. Wang and M. Chiang, “RAT selection games in HetNets,” in *Proceedings of IEEE INFOCOM* , 2013.

[77] J. Chong, C. Joe-Wong, S. Ha and M. Chiang, “CYRUS: Toward client-defined Cloud storage,” in *Proceedings of EuroSys.*, 2015.

[78] A. Chakraborty, V. Navda, V. N. Padmanabhan and R. Ramjee, ““Coordinating cellular background transfers using LoadSense,” in *ACM Mobicom*, 2013.

[79] L. Canzian and M. v. d. Schaar, “Real time stream mining: Online knowledge extraction using classifier networks,” *IEEE Networks Special Issue on Networking for Big Data*, 2014.

[80] S. J. Stolfo, M. B. Salem and A. D. Keromytis, “Fog Computing: Mitigating Insider Data Theft Attacks in the Cloud,” in *Proceeding of IEEE Symposium on Security and Privacy Workshops (SPW)*, San Francisco, California, USA, 2012.

[81] G. Brown, “Converging Telecom and IT in the LTE RAN,” white paper at Heavy Reading on behalf od Samsung, 2013.

[82] S. Kitanov and T. Janevski, “Mobile Cloud Computing in 5G Networks,” in *the XI International Conference of Society for Electronics, Telecommunications, Automation and Informatics ETAI 2015*, Ohrid, Macedonia, 2015.

[83] M. Peng, S. Yan, K. Zhang and C. Wang, “Fog Computing based Radio Access Networks: Issues and Challenges,” *IEEE Network*, p. arXiv:1506.04233v1 [cs.IT], 2015.

[84] M. Peng, Y. Li, Z. Zhao and C. Wang, “System architecture and key technologies for 5G heterogeneous cloud radio access networks,” *IEEE Network*, vol. 29, no. 2, pp. 6-14, 2015.

[85] M. Peng, Y. Li, J. Jiang, J. Li and C. Wang, “Heterogeneous cloud radio access networks: A new perspective for enhancing spectral and energy efficiencies,” *IEEE Wireless Communications*, vol. 21, no. 6, pp. 126-135, 2014.

[86] J. Marinho and E. Monteiro, “Cognitive radio: Survey on Communication Protocols, Spectrum Decision Issues, and Future Research Directions,” *Springer Wireless Networks*, vol. 18 , no. 2, p. 147–164. , 2012.

[87] P. Bahl, R. Y. Han, L. E. Li and M. Satyanarayanan, “Advancing the State of Mobile Cloud Computing,” in *Third ACM Workshop on Mobile Cloud Computing and Services*, 2012.

[88] E. Ahmed, A. Gani, M. Sookhak, S. H. A. Hamid and F. Xia, “Application Optimization in Mobile Cloud Computing: Motivation, Taxonomies, and Open Challenges,” *Journal of Network and Computer Applications*, vol. 2, pp. 52-68, 2015.

[89] D. B. Hoang and L. Chen, “Mobile cloud for assistive healthcare (mocash),” in *Services Computing Conference (APSCC),* Asia-Pacific, 2010.

[90] S. Adibi, “Biomedical Sensing Analyzer (BSA) for Mobile-Health (m-Health) – LTE,” *IEEE Journal of Biomedical and Health Informatics,* vol. 18, no. 1, pp. 345-351, 2014.

[91] G. Sun and J. Shen, “Facilitating Social Collaboration in Mobile Cloud-based learning: a Teamwork as a Service (TaaS) Approach,” *IEEE Transactions on Learning Technologies* , vol. 7, no. 3, p. 207–220, 2014.

[92] J. Wu, C. Yuen, N.-M. Cheung, J. Chen and C. Chen, “Enabling Adaptive High-frame-rate Video Streaming in Mobile Cloud Gaming Applications,” *IEEE Transactions on Circuits and Systems for Video Technology,* vol. PP, no. 99, pp. 1-1, 2015.

[93] K. Sabarish and R. Shaji, “A Scalable Cloud Enabled Mobile Governance Framework,” in *Global Humanitarian Technology Conference - South Asia Satellite (GHTC-SAS),* 2014.

[94] E. Ahmed, A. Gani, M. K. Khan, R. Buyya and S. U. Khan, “Seamless Application Execution in Mobile Cloud Computing: Motivation, Taxonomy and and Open Challenges,” *Journal of Network and Computer Applications,* vol. 52, p. 54–172, 2015.

[95] Erricsson, “Ericsson Mobility Report, 2013; http://www.ericsson.com/res/docs/2013/ericsson-mobility- report-june-2013.pdf.,” Ericsson, 2013.

[96] L. Atzori, A. Iera and G. Morabito, “The internet of things: A survey,” *Computer networks,* vol. 54, no. 15, p. 2787–2805, 2010.

[97] M. Satyanarayanan, P. Bahl, R. Caceres and N. Davies, “The Case for VM-based Cloudlets in Mobile Computing,” *IEEE Pervasive Computing,* vol. 8, no. 4, pp. 14-23, 2009.

[98] A. Ahmed and E. Ahmed, “A Survey on Mobile Edge Computing,” in *IEEE International Conference on Itelligent System and Control (ISCO 2016),* 2016.

[99] M. Beck, M. Werner, S. Feld and S. Schimper, “Mobile edge computing: A taxonomy,” in *Sixth International Conference on Advances in Future Internet,* 2014.

[100] Y. Yu, “Mobile Edge Computing Towards 5G: Vision, Recent Progress, and Open Challenges,” *China Communications,* vol. Supplement No. 2, pp. 89-99, 2016.

[101] R. Buyya; J. Broberg; A. Goscinski A; et al., “Cloud Computing Principles and Paradigms,” in *Cloud Computing 2011.*

[102] B. Pfaff; J. Pettit; T. Koponen; et al., “The design and implementation of open vSwitch,” *Networked Systems Design and Implementation,* 2015.

[103] G. Liu, X. Hou and J. Jin, “3-D-MIMO With Massive Antennas Paves the Way to 5G Enhanced Mobile Broadband: From System Design to Field Trials,” *IEEE Journal on Selected Areas in Communications,* vol. 35, no. 6, pp. 1222-1233, 2017.

[104] I. Jovović and M. P. I. Forenbacher, “Massive Machine-Type Communications: An Overview and Perspectives Towards 5G,” in *The 3rd International Virtual Research Conference In Technical Disciplines,* 2015.

[105] O. N. C. Yilmaz, Y.-P. E. Wang and N. A. Johansson, “Analysis of ultra-reliable and low-latency 5G communication for a factory automation use case,” in *IEEE International Conference on Communication Workshop (ICCW),* London, UK, 2015.

[106] Nebbiolo Technologies, “Fog vs Edge Computing v1.1,” Nebbiolo Technologies Inc. Report, Milpitas, California, USA, 2015.

[107] S. Singh, Y. Chiu and Y. T. J. Yang, “Mobile Edge Fog Computing in 5G Era: Architecture and Implementation,” in *2016 International Computer Symposium*, Chiayi, Taiwan, 2016.

[108] S. Kitanov and T. Janevski, “Hybrid Environment Service Orchestration for Fog Computing,” in *51st International Scientific Conference on Information, Communication and Energy Systems and Technologies (ICEST 2016),* Ohrid, Macedonia, 2016.

[109] ITU, “The Tactile Internet,” ITU Technology Watch Report, 2014.

[110] S. Kitanov and T. Janevski, “Fog Computing Service Orchestration Mechanisms for 5G Networks.,” *Journal of Internet Technology (JIT),*, no. ISSN 1607-9264, Taiwan, 2018. (2016 IMPACT FACTOR: 1.930), 2018.

[111] S. Kitanov and T. Janevski, “Quality Evaluation of Cloud and Fog Orchestrated Services in 5G Network,” *European Journal of Information Science and Technology (EJIST),* vol. 1, no. 1, pp. 62-72, 2016.

[112] S. Kitanov and T. Janevski, “Fog Computing as a Support for 5G Network,” *Journal of Emerging Research and Solutions in ICT (ERSICT) published by Faculty of Information and Communication Technologies in Bitola, Macedonia,* vol. 1, no. 2, pp. 47-60, 2016.

[113] D. Warren and C. Dewar, “Understanding 5G: Perspective on Future Technological Advancements in Mobile,” GSMA Intelligence Analysis Report, 2014.

[114] S. Kitanov and T. Janevski, “State of the Art : Fog Computing for 5G Networks,” in *XXIV International Telecommunication Forum TELFOR 2016,* Belgrade, Serbia, 2016.

[115] S. Kitanov and T. Janevski, “Fog Computing for 5G Networks,” *Accepted paper as a poster presentation on 8th ICT -Innovations International Conference,* Ohrid, Macedonia, 2016.

[116] A. Roessler, M. M. J. Schlienz and M. Kottkamp, “LTE-Advanced (3GPP Rel.12) Technology Introduction,” Rhode and Scwharz White Paper, 1MA252_2E, 2014.

[117] S. Kitanov and T. Janevski, “5G Networks in the Fog Computing Environment,” in *XII International Conference of Society for Electronics, Telecommunications, Automation and Informatics ETAI 2016*, Struga, Macedonia, 2016.

[118] S. Kitanov and T. Janevski, “Energy Efficiency of Fog Computing and Networking Services in 5G Networks,” in *XVII IEEE International Conference on Smart Technologies EUROCON 2017*, Ohrid, Macedonia, 2017.

[119] W. V. Heddeghem, S. Lambert, B. Lannoo, D. Colle, M. Pickavet and P. Demeester, “Trends in worldwide ICT electricity consumption from 2007 to 2012,” *Journal of Computer Communications,* vol. 50, pp. 64-76, 2014.

[120] S. Kitanov and T. Janevski, “Energy Efficiency of 5G Mobile Networks in Hybrid Fog and Cloud Computing Environment,” in *II International Conference Recent Trends and Applications in Computer Science and Information Technology RTA-CSIT*, Tirana, Albania, 2016.

[121] J. Huang, F. Qian, A. Gerber, Z. M. Mao, S. S and O. Spatscheck, “A Close Examination of Performance and Power Characteristics of 4G LTE Networks,” in *Mobile Systems, Applications and Services (MobiSys 2012) Conference,* Low Wood Bay, Lake District, UK, 2012.

[122] C. Ragona, F. Granelli and C. Fiandrino, “Energy-Efficient Computation Offloading for for Wearable Devices and Smartphones in Mobile Cloud Computing,” in *IEEE Global Communications Conference (GLOBECOM),* 2015.

[123] V. Nikolikj, “Dynamic QoS and Operationally Efficient Support in the Conditions of Growing Traffic in LTE and Future Heterogeneous Mobile Networks,” Skopje, Republic of Macedonia: Doctoral Dissertation, Ss Cyril and Methodius University, Faculty of Electrical Engineering and Information Technologies, 2015.

[124] ITU, “DRAFT NEW REPORT ITU-R M.[IMT-2020.TECH PERF REQ] - Minimum requirements related to technical performance for IMT-2020 radio interface(s),” ITU-R SG05 Contribution 40, 2017.