MEC towards 5G: A Survey of Concepts, use Cases, Location Tradeoffs

Belghol Hibat Allah and Idrissi Abdellah
IPSS Team, Computer Science Department, Faculty of Sciences
Mohammed V University in Rabat
hibatbelghol@gmail.com; idrissi@fsr.ac.ma

ABSTRACT

In recent years, there has been a new trend to push everything to a centralized cloud, triggered by virtualization and pushed by the need to reduce costs and increase suppleness. In the process, mobile operators and industry players forgot how prominent the location of the functionality is to performance, optimal use of network resources and user experience. As they progressively grasp the influence of location in wireless networks and specifically in virtualized networks they start to look for efficient ways to deploy network taking in consideration those metrics. On the one hand Distributed Cloud RAN (DC-RAN) which consists in deploying Base Band Units (BBUs) in distributed way instead of pooling the units at centralized data center and Multiple-access Edge Computing (MEC), Previously known as Mobile Edge Computing which consists in mixing the IT & Telco domains and bringing their capabilities within the close proximity of mobile subscribers to better serve them are gaining acceptance in mobile networks. On the other hand, Software Defined Networking (SDN) and Network Function Virtualization (NFV), two promising concepts are expected to take mobile networks to a high level of agility. While SDN is based on the separation of control and data planes, NFV separates software from hardware enabling flexible network deployment and dynamic operation. Virtualization, originally used as the support for shifting to the centralized cloud, is even more basic in permitting hybrid models because it offers service providers the chance to choose a location, hardware, and software separately to optimize end-to-end network performance and Quality of Experience (QoE). In this perspective, NFV, SDN, DC-RAN, and MEC in distinct but in complementary ways man up the necessity to put processing and storage where it's suitable to preserve a healthful harmony between what lingers centralized and what have to be distributed to the edge based on parameters such as applications, traffic type, and network conditions. In this paper, we will present MEC use cases that have gained an attraction to date and we will shed the light on the importance of the edge location and criteria to take in consideration while deploying MEC within network along with the MEC location tradeoffs.

Keywords: MEC, NFV, SDN, DC-RAN, 5G, EDGE.

1   Introduction

The transition from LTE to 5G will be the deepest revolution in the wireless industry since the evolution from analog to digital [1]. There are many reasons for this, primarily because new uses for wireless technology are creating demand for new approaches to connectivity, bandwidth, and network

architecture; 5G is expected to fulfill a very wide range of requirements such as higher bandwidth, ultra-dense network, very low latency and huge capacity. This evolution will bring about several new ways of designing networks so that the promise of always-on, high-bandwidth, low latency and massive networks can become reality. According to the last Cisco forecasts, the overall mobile data traffic especially the video one in its multiple forms; e.g, downloaded, streamed, uploaded, interactive is expected to know a phenomenal growth by 2021, a seven-fold increase over 2016. This tremendous rise of traffic is driven mainly by five major trends: evolution toward Smarter Mobile Devices, defining Cell Network Advances - the evolution of 4G and the coming introduction of 5G, mass mobile Internet of Everything (IoE) adoption, versatility of Wi-Fi and expansion of its use and coverage and identification of new mobile applications and requirements [2]. However, due to resource constraints (processing power, battery lifetime, and storage capabilities), mobile devices are unable to perform effectively and cater for this myriad of bandwidth-intensive applications with stringent requirements, which calls for Cloud Computing (CC) involvement. Thus, the traditional CC centralized model (processing, storage...) does not align with the 5G requirements and may lead to some limitations such as high latency, limited bandwidth, and scalability problems [3].

IT and Telecom Networking domains are converging bringing with them new opportunities and capabilities that can be deployed into the network. A key revolution has been the capacity to run IT servers at network edge nearby the mobile subscriber applying cloud computing concept, we define this as MEC. This is one of the most relevant initiatives proposed during the past few years to address the above challenges. MEC speeds up content, services, and applications, augmenting reactivity from the edge. Laying by this the ground for e-health, self-driving cars, augmented reality, intelligent houses and cities, enhanced mobile broadband and many other applications. The mobile subscriber's experience can be enhanced through effective network and service operations, based on an understanding of the radio and network conditions [4]. In today's networks, the dominant approach is to boost performance, given the accessible financial resources. This generally means enhancing network capacity using a pressure approach by deploying the recent technological features that increase throughput. But far less effort is put into optimizing the use of the capacity available. There are many ways to optimize network resource utilization and, under severe pressure to improve performance without increasing costs, mobile operators have started to take steps toward achieving this optimization goal. At the RAN level, two approaches have to be considered: DC-RAN deploys BBUs in distributed manner nearby the Remote Radio Unit (RRU) instead of pooling the units at centralized data center while MEC is in general geared to balance the processing, storage and control in the network by moving processing, storage and control to the edge of mobile network, but the edge location and its prominence remains blur which leads us to explore the different possible locations and assess their impact on the global architecture. For the core network, SDN-NFV integration could bring huge benefits enabling a flexible, agile and scalable network. The remainder of this paper is structured as follows: An overview of MEC is given in Section II. Section III, lists MEC use cases gaining attraction then the importance of the edge location, as well as location tradeoffs are examined in depth. Finally, Section IV concludes the paper.

2 Overview of Multiple-Access Edge Computing (MEC)

Among the different actions taken to realize edge computing and, more specifically, mobile edge computing, European Telecommunications Standards Institute (ETSI) MEC occupies a central role because it provides a structure to move processing, storage, and control to the edge. MEC was conceived to

respond to mobile operators need to shift the processing and storage of some services and applications to the brink, and to improve mobile network performance and resource usage in real time. MEC standardization work began in 2014 at ETSI, with a seminal white paper authored by the founders of the MEC ISG Huawei, IBM, Intel, Nokia, NTT DOCOMO, and Vodafone [5]. This white paper discusses the business, consumer and technical value/benefits offered by this technology, the enablers, the requirements and challenges for MEC as well as the objectives of the MEC initiative. Additionally, it presents the high-level architectural blueprint of MEC as well as a number of use cases where MEC seems to have a prominent impact. Content caching, application computation offloading and connected vehicles are examples of uses cases among other existing. The list of active participants has since expanded to cover more vendors and operators. The focal point was first the mobile networks, but it has extended to fixed networks, to mirror the close amalgamation of mobile and fixed networks, which often host the same services and applications. The MEC acronym not any more assigned to Mobile Edge Computing and rather refers to Multiple-access Edge Computing. Albeit the change was catalyzed by ETSIs internal process requirements, it was a welcome change as it expands the reach and potential of MEC in today's rapidly evolving wireless networks, which consist of a various set of access technologies and spectrum regulatory structures. The extension to non-cellular technologies means that Wi-Fi is now included within MECs scope. This is a welcome addition that mirrors the fact Wi-Fi in the majority of markets stands for the bulk of traffic to mobile devices. The ETSI MEC standardization work creates the foundation for edge computing deployments in mobile and, progressively, fixed networks. During the first term (2015-2017), ETSI ISG finished the basic work, delivered the elemental specifications, and emboldened the creation of the ecosystem. More work is needed during the second terms (2017-2018) not only to extend beyond mobile networks but also to bolster the connections with other edge computing initiatives while evading the fragmentation of exertions. Beyond standardization and industry collaborative initiatives, there is a need to investigate various business and deployment models and review the role that stakeholders will have in the development chain [6]. MEC and other promising edge computing initiatives such as Fog Computing, Cloudlets address the need to place processing and storage at the network edge in close proximity to the final user. In this context, a high-level overview was proposed in [7] to show that those three approaches include partially overlapping concepts and are also complementary. In many ways, edge computing concept is nothing new. There has been edge computing solutions all the time to respond to a specific needs and optimization issues in mobile networks. What differs today is that network virtualization gives a framework to extend edge computing capacities, adding adaptability, trustworthiness, and elasticity. This will permit to operators to garner profits in terms of improved QoE and resource usage. While MEC and C-RAN are not straight linked to each other, the C-RAN function, in theory, could be deployed over any cloud service. However, to make that possible, the front-haul transport between the RRUs at the cell sites and the virtualized BBUs in the centralized data centers would require high throughput and low latency connections. Achieving this kind of transport requires that the centralized BBUs be placed within close proximity to the RRUs. With this perspective, it's conceivable that C-RAN could be implemented at the edge as a way of augmenting the benefits of MEC [8]. And that was the main goal of the work done in [9]. The authors presented in [9] an integration of mobile edge services in scenarios where an aggregation of small cells is able to share certain processing and storage resources in a local perspective. Therefore, rather than centralizing RAN and MEC functions into Big Data center, the proposed solution creates a distributed light data center by coordinating the micro servers associated with different small cells in a cluster. MEC and Fog computing, the two are closely related. Both describe

standards-based architectures where computing, storage, and networking resources are made available in a cloud fashion on host servers located at the network edge. The main difference is in the context in which they are referenced. Fog computing is often mentioned in the context of Internet of Things (IoT), where host servers are typically routers, access points or even computing devices co-located with sensors and actuators. MEC is often mentioned in the context of mobile networks, where host servers are integrated with the mobile network infrastructure, such as base stations or aggregation sites. It is considered as a sub-category of Fog computing. IoT, which is highly associated with the mobile network, will benefit from extending the Software Defined Mobile Edge Computing (SD-MEC) concept to other areas (VANET, WSN, etc.) to enable wide deployment for IoT. In [10], the authors have tried to shed light on these new trends and drawn their vision of the Internet of the future. Based on a virtualized platform, MEC is recognized by the European 5G Infrastructure Public Private Partnership (5G PPP) research body as one of the key emerging technologies for 5G networks together with NFV and SDN [11].

To the best of our knowledge, the integration of MEC, DC-RAN, NFV, SDN and Big Data Analytics in one overall network architecture putting a special focus on the prominence of the choice of the edge location has not been addressed in the existing works. A Cloud-based Wireless Network with virtualized, reconfigurable and smart Wireless Network was proposed by the authors in [12]. However, the architecture didn't include MEC and DC-RAN which are two key factors to boost the network responsiveness and decrease latency to cater for the wide variety of 5G requirements. In essence, this paper will discuss the most relevant MEC uses cases and analyze the importance of the MEC server location and the key parameters to take in consideration in order to ensure the balance of the whole network.

3 MEC Applications Scenarios and Edge Location Tradeoffs

3.1   Applications Scenarios

In this section, we will start by presenting MEC use cases that have attracted attention to date and which are divided into three main categories: Operator and third party services, user-oriented services and network performance and QoE improvements. Then we will discuss the importance of the edge location along with the tradeoffs to be considered.

3.1.1   Intelligent Video Acceleration:

Fig.1 shows a radio analytics application providing the video server with an indication on the throughput estimated to be available at the radio downlink interface. This information can be used to assist TCP congestion control decisions and also to ensure that the application-level coding matches the estimated capacity at the radio downlink and then enables improved video quality and throughput. It uses data about network conditions especially RAN conditions and RAN load to generate advice for content and application providers on how to manage traffic exchanged with the subscriber. When the network has sufficient capacity, the providers can share content at the highest quality available. When the network is capacity constrained or congested, the content and TCP transmission can be adapted to provide subscribers the best experience possible given the real-time availability of network resources [13].
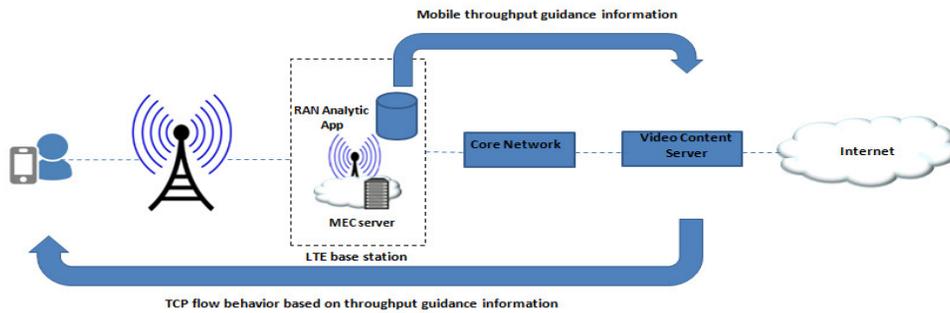
Figure 1. Intelligent Video Acceleration

3.1.2   Local Content Caching:

Fig.2 depicts how subscribers at a given location and time tend to watch a remarkably consistent and narrow set of content so not only traffic is growing, it is often concentrated in specific locations and times that are the backdrop for perfect use case to justify moving processing and storage to the edge. MEC cannot lower the latency in the RAN, but it can reduce the end-to-end latency by storing locally the most popular content consumed in a specific geographical area. Once requested, the content is provided from the local cache. In this case no need to transfer the content always through the core network [13].
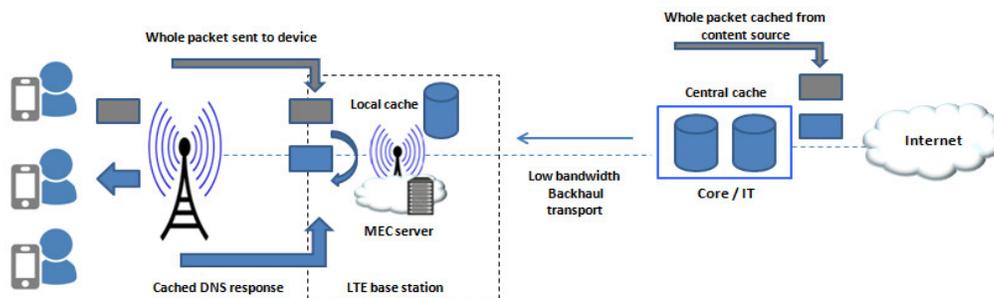

Figure 2. Local Content Caching

3.1.3   Augmented Reality:

Fig.3 refers to Augmented Reality. This MEC application analyses the output from a device's camera and the precise location. Objects viewed on the device camera are overlaid with local augmented reality content. It enables the unique experience of a visitor to a museum or other (indoors or outdoors) points of interest and ensures low latency and high rate of data processing [14].
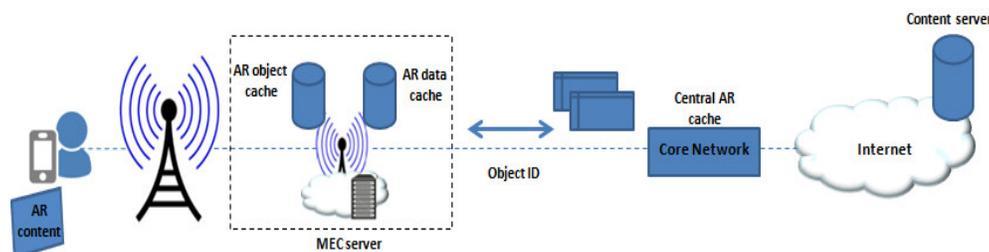

Figure 3. Augmented Reality

Application computation offloading: In Fig.4 the mobile edge host executes compute-intensive functionalities with high performance instead of the mobile devices, improving the user experience and the consumer can use low complexity devices by off-loading compute capacity to the mobile edge host. Therefore, new compute-intensive services like auto-translation and recommendations synthesized from more than one mobile edge application can be delivered in near real time to low complexity devices. In this way, the operator can provide value added services by utilizing application data available at the mobile edge host [13].
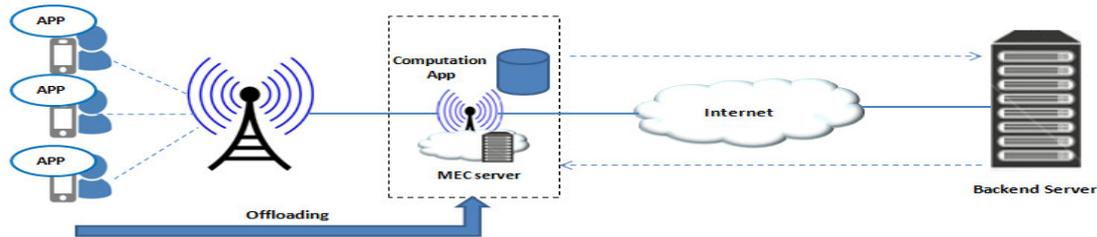

Figure 4. Application computation offloading

3.1.4   Video Analytics:

Fig.5 presents the distributed live video streams analytics application where Events are triggered automatically (e.g. movement, missing objects, crowd) enabling fast detection and action triggering; which is useful and applicable to public safety, smart cities [13].
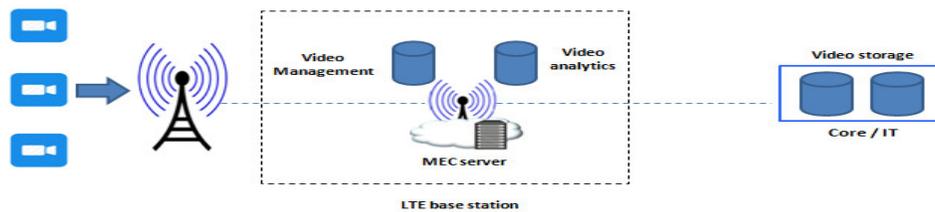

Figure 5. Video Analytics

3.1.5   Location Based Services:

Fig.6 shows how a real-time location of an active device is tracked and provided in a passive way independent of GPS information Mitigating indoor limitation of GPS and helps to locate specific users and understand how the crowd is distributed. It is applicable to Smart City, Geo-Fencing, Retail, and advertising [13].
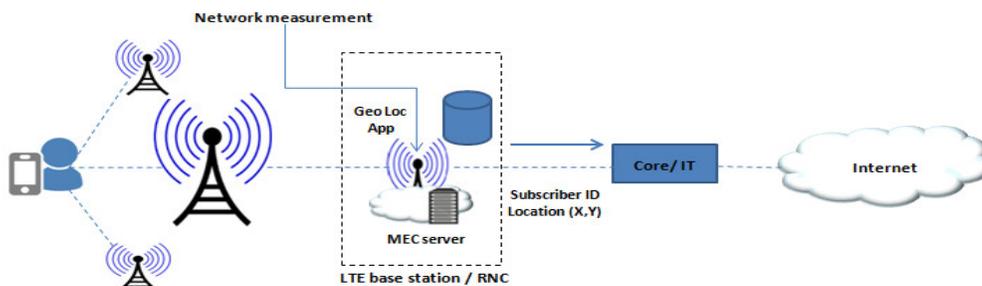

Figure 6. Location Based Services

3.1.6   Connected Vehicles:

Existing cloud services are extended into the highly distributed mobile base station environment, leveraging the existing LTE connectivity. The MEC application presented in fig.7 operates as a roadside unit for vehicle-to-infrastructure (V2I) and road hazards can be recognized and warnings can be sent to nearby cars with extremely low latency which enables a nearby car to receive data in a matter of milliseconds, and the driver to react instantly [14].
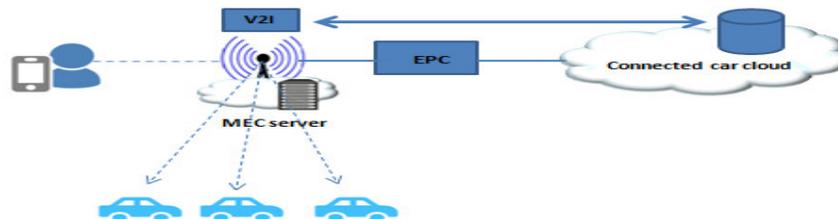

Figure 7. Connected Vehicles

3.2   The prominence of MEC server location

We have talked so far about MEC concept and how shifting processing and storage to the edge of mobile network could bring huge benefits in term of QOE and network performance. However, is not clear at all where the edge could be or, more meticulously, where possible edge locations are, and which one or ones should be selected to deploy MEC server. MEC requires more than just deploying a server at a given edge location. This   is crucial to ensure that selected edge functionality location brings both performance improvement and cost saving to the service provider. To fully harness the benefits of moving some processing and storage capabilities to the edge an evaluation of possible scenarios should be done. If the edge location is too close to the end user, MEC may become overly expensive and complex. If the edge location is too close to the centralized core, the benefits dissipate, leading to more complexity but no gain in term of performance. There are several locations where deployment of MEC functionalities could be a clever step. Inspecting the tradeoffs that these locations offer is essential for successful deployment. It is important to choose the edges location within the network painstakingly to prevent unnecessary investment. For example, Location-based applications may be housed in an aggregation point that covers the venue. An enterprise deployment may be hosted in a location that reaches all the enterprise's buildings or a part of them. For applications requiring video caching, there is more flexibility in choosing the MEC location. Since service providers may pick the location based on subscriber usage trends and see where the caching contribution could be maximized. For the abovementioned applications, stationary edge location through time may be preferable. But for a network with fluctuating loads, dynamic edge shifting according to real-time network conditions is possible and desirable in a virtualized environment. To achieve this, huge orchestration capabilities are needed in the core network. On one hand, MEC links the RAN with the UE. On the other hand, links the applications with content. It gives also the opportunity to manage traffic in real time based on factors such as application, content requirements, and network conditions to optimize available network resources use [15].

3.3   Edge Location tradeoffs

Moving processing and storage to the far edge would improve considerably the QoE and performance. However, there are many constraining parameters such as cost, complexity, space and latency on which

the choice of MEC location is based. To select the best location the choice should rest on a compromise between the criteria cited above.

Table I, shows how the MEC location impact the network key parameters as it gets closer to the UE and the opposite as it moves toward the centralized cloud:

Table 1. Impact of MEC location on network performance

Network constraining parameters	MEC location	Closer to UE	Closer to centralized cloud
Latency		Low	High
Storage Capacity		Limited	High
Processing power		Limited	High
End to end complexity		High	Low
Resources optimization		Low	High
Covered area Size		Small	Big

As the MEC location gets closer to the UE, the latency gets lower and this improves performance and QOE especially for time sensitive applications. Hardware has less storage capacity, limiting the amount of content that can be stored at the edge. Processing power becomes limited, and hence it may not be suitable to run some applications from the edge. End-to-end network complexity increases since network operators have to deploy, integrate and manage more hardware, at a higher number of locations. If applications could be smoothly run from a centralized location resources allocation at the edge may be pointless. Another key factor is the covered area. As the edge host moves away from the centralized cloud where mobility management and RAN optimization are processed, the covered zone gets smaller. Edge traffic optimization is adapted to this area. So, geographically distributed MEC servers can optimize their respective covered areas, but in separated way. But if only one MEC server covers the whole area, optimization can be coordinated across the entire area. This coordination makes optimization more efficient, but at the same time, it increases the distance to the RAN which may collapse the optimization capabilities. Access to edge applications and services is preserved as subscribers move from one area to another within the same network. If the edge host covers multiple networks, application access can be preserved across networks. As the subscriber moves to an area that has not the same MEC capabilities or doesn't support MEC at all, the subscriber may experience degradation in the quality of the access to the edge application or lose connectivity as they move out of the covered area. In many cases, this last aspect is desired especially for actors that want their services to be available only within a specific area. When the selected MEC location moves toward the centralized cloud the inverse holds giving high latency, more storage capacity, big processing power, less complexity, more efficient use of resources [15]. Table II, presents factors that play a major role in the selection of MEC location:

Table 2. Key factors defining the choice of MEC location

Factors	How MEC address the network needs
RAN resources	The MEC capacities must be adequate to the RAN capabilities to serve the covered area.
Backhaul resources	MEC may address capacity and/or latency limitations in the backhaul and prevent the backhaul from becoming a performance bottleneck.
Applications	Latency, processing and storage requirements that affect MEC location vary across applications, so the best MEC location should vary by application.
Subscriber & client expectations	The expectation for service performance and QoE may be different for enterprise employees, mall visitors.
Operator preferences	Operators may want to position themselves in the market in a specific way e.g., provide a higher-quality service for a specific enterprise client or venue owner.

4 Conclusion

In this paper, we have seen how MEC allows content, services, and applications to be accelerated by increasing responsiveness from the edge. The impact on latency, QoE and performance from moving processing and storage to the edge is easy to grasp, even though it is not trivial to quantify over a network because it depends on multiple environmental factors that are variable. We discussed also the MEC use cases that have attracted attention and their impact on network performance then we shed the light on the importance of the edge location along with the tradeoffs to be considered while deploying MEC servers across the network.

This paper is a solid basis that paves the ground to our future work that will propose the integration of MEC, DC-RAN, NFV, SDN and Big Data Analytics in an overall network architecture and discuss the benefits brought back by this combination. Furthermore, as an extension of this work, we also think that we can reuse the constraint programming concepts as proposed in [16 - 23] to bring better approaches to all these systems.

REFERENCES

[1]   [1]INTERDIGITAL, "Putting Together the Pieces of the 5G Puzzle," http://www.interdigital.com/ : White paper, October 2016, p.2.

[2]   Cisco VNI Mobile, "Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 20162021," http://www.cisco.com: White paper, February 2017, pp.1-28.

[3]   G. A. Lewis, "Mobile computing at the edge (keynote)," in Proceedings of the 1st International Conference on Mobile Software Engineering and Systems. ACM, 2014, pp. 69–70.

[4]   N. Sprecher, "MEC Activity Report 2015," https://portal.etsi.org/.

[5]   M. Patel, J. Joubert, J. Ramos, N. Sprecher, S. Abeta, A. Neal "Mobile-Edge Computing – Introductory Technical White Paper," https://portal.etsi.org/:White paper, September 2014.

[6]   D. Sabella, N. Sprecher, M. Patel, V. Young, Y. Chao Hu "Mobile Edge Computing a key technology towards 5G," http://www.etsi.org/: Whitepaper, September 2015.

[7]   E. Borcoci, "Fog Computing, Mobile Edge Computing, Cloudlets - which one?." https://www.iaria.org/: SoftNet 2016 Conference, August 2016.

[8]   INTERDIGITAL, "What the MEC? An Architecture for 5G," http:// www.interdigital.com/ : White paper, December 2016.

[9]   J. O. Fajardo, F. Liberal, I. Giannoulakis, E. Kafetzakis, V. Pii, I. Trajkovska, T. M. Bohnert, L. Goratti, R. Riggio, J. G. Lloreda et al.,"Introducing mobile edge computing capabilities through distributed 5g cloud enabled small cells," Mobile networks and applications, vol. 21, no. 4, pp. 564–574, 2016.

[10]   O. Salman, I. Elhajj, A. Kayssi, and A. Chehab, "Edge computing enabling the internet of things," in Internet of Things (WF-IoT), 2015 IEEE 2nd World Forum on. IEEE, 2015, pp. 603–608.

[11]   5Gppp, "5G Vision," https://5g-ppp.eu/: February 2015.

[12]   M. Chen, Y. Zhang, L. Hu, T. Taleb, and Z. Sheng, "Cloud-based wireless network: Virtualized, reconfigurable, smart wireless network to enable 5g technologies," Mobile Networks and Applications, vol. 20, no. 6, pp. 704–712, 2015.

[13]   G. Klas, "ETSI MEC Use Cases and Requirements," http://www.etsi.org/: October 2016.

[14]   N. Sprecher, "Mobile Edge Computing An enabler for enhanced Car2X communication," http://www.etsi.org/: June 2016.

[15]   S. Fili, M. Paolini, "Power at the edge processing and storage move from the central core to the network edge," : White paper, February 2017.

[16]   Abourezq, M., Idrissi, A. Introduction of an outranking method in the Cloud computing research and Selection System based on the Skyline. Proceedings - International Conference on Research Challenges in Information Science. 2014.

[17]   Idrissi, A. Some methods to treat capacity allocation problems. Journal of Theoretical and Applied Information Technology. 2012.

[18]   Idrissi, A., Abourezq, M. Skyline in cloud computing. Journal of Theoretical and Applied Information Technology. 2014.

[19]   Abourezq, M., Idrissi, A. A cloud services research and selection system. Proceedings of International Conference on Multimedia Computing and Systems. 2014.

[20]   Yakine, F., Idrissi, A. Energy-aware topology control and QoS routing in ad-hoc networks. Procedia Computer Science. 2015.

[21]   Idrissi, A., Rehioui, H., Laghrissi, A., Retal, S. An improvement of DENCLUE algorithm for the data clustering. 5th International Conference on Information and Communication Technology and Accessibility, ICTA 2015. 2015.

[22]   Rehioui, H., Idrissi, A., Abourezq, M., Zegrari, F. DENCLUE-IM: A New Approach for Big Data Clustering. Procedia Computer Science. 2016.

[23]   Idrissi, A., Yakine, F. Multicast routing with quality of service constraints in the ad hoc wireless networks. Journal of Computer Science. 2014.