Bringing computation closer towards user network: Is edge computing the solution?

Article *in* IEEE Communications Magazine · November 2017

DOI: 10.1109/MCOM.2017.1700120

Arif Ahmed
Université de Rennes 1

Junaid Shuja
COMSATS University Islamabad

Bringing computation closer towards user network: Is edge computing the solution?

Article *in* IEEE Communications Magazine · September 2017

Bringing computation closer towards user network: Is edge computing the solution?

Ejaz Ahmed, *Member, IEEE,* Arif Ahmed, Ibrar Yaqoob, Junaid Shuja, Abdullah Gani, *Senior Member, IEEE,* Muhammad Imran *Member, IEEE*, Muhammad Shoaib

Abstract—The virtually unlimited available resources and wide range of services provided by the cloud have resulted in the emergence of new cloud-based applications, such as smart grids, smart building control, and virtual reality. These developments however, have also been accompanied by a problem for delay-sensitive applications that have stringent delay requirements. The current cloud computing paradigm cannot realize the requirements of mobility support, location awareness, and low latency. Hence, to address the problem, an edge computing paradigm that aims to extend the cloud resources and services and enable them to be nearer the edge of an enterprise's network has been introduced. In this article, we highlight the significance of edge computing by providing real-life scenarios that have strict constraints requirements on application response time. From previous literature, we devise a taxonomy to classify the current research efforts in the domain of edge computing. We also discuss the key requirements that enable edge computing. Finally, current challenges in realizing the vision of edge computing are discussed.

Index Terms—Edge computing, cloudlet, fog computing, cloud computing, distributed computing.

1 Introduction

Rapid advancements in computing technologies have enabled a wide range of applications usually categorized as future Internet applications. Examples of these applications are road traffic and smart surveillance. However, majority of these emerging applications are compute-intensive in nature and impose stringent requirements on delay. Their compute-intensive nature makes such applications difficult to run on resource-constrained mobile devices. The limited capabilities of the mobile devices are augmented by leveraging the resources of cloud servers. The cloud provides virtually unlimited resources and a wide range of services to enable such compute-intensive applications on the resource-constrained mobile devices. The delay-sensitive compute-intensive applications still suffer from large Wide Area Network (WAN) latency that transforms euphoria into a problem because of the strict delay requirements. The current cloud computing paradigm augments the capabilities of resource-constrained devices, but it cannot fulfill the requirements of location awareness, mobility support, and low latency [1].

In this context, researchers introduced a vision of edge computing to enable the applications on billion of smart connected devices to run directly at the network edge. Similar to the cloud, edge computing also assists the user by providing compute, data, storage,

- *E. Ahmed (Corresponding Author), I. Yaqoob, and A. Gani are with the Department of Computer System & Technology, University of Malaya, 50603 Kuala Lumpur, Malaysia. (E-mail: {ejazahmed, abdullahgani}@ieee.org, {ibraryaqoob@siswa.um.edu.my)*
- *A. Ahmed works for the Department of Computer Science & Engineering, National Institute of Technology, Silchar, Silchar, India. (Email: arifch2009@gmail.com)*
- *J. Shuja is working with COMSATS Institute of Information Technology, Abbottabad, Pakistan. (E-mail: junaidshuja@ciit.net.pk)*
- *M. Imran (Corresponding Author) and M. Shoaib are working with the College of Computer and Information Sciences, King Saud University, Saudi Arabia. (Email: dr.m.imran@ieee.org, muhshoaib@ksu.edu.sa)*

TABLE 1: Differences Between Cloud Computing and Edge Computing

Parameters	Cloud Computing	Edge Computing
Service Location	Within the Internet	In edge network
Distance (Number of Hops)	Multiple hops	Single hop
Latency	High	Low
Jitter	High	Very low
Location awareness	No	Yes
Geo-distribution	Centralized	Distributed
Mobility Support	Limited	Supported
Data enroute attacks	High probability	Very low probability
Target user	General Internet users	Mobile users
Service Scope	Global	Limited
Hardware	Scalable capabilities	Limited capabilities

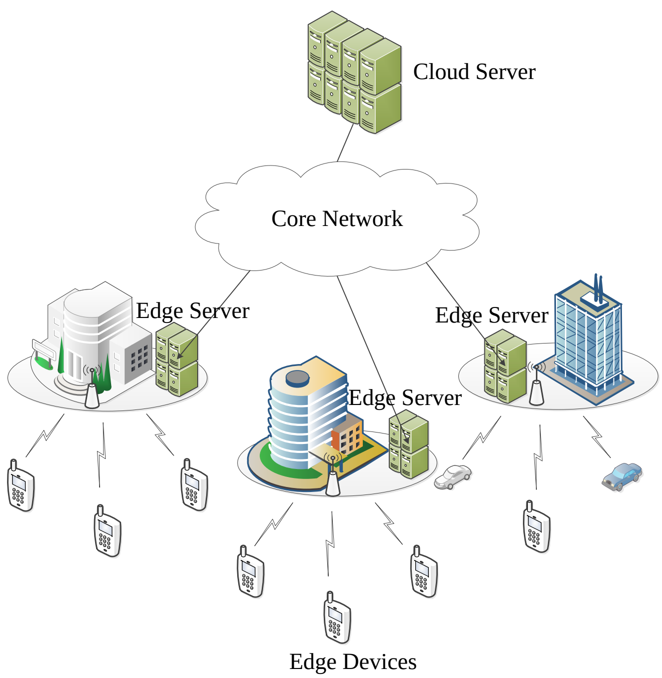

Fig. 1: Edge Computing Architecture

and application services [2]. The distinguishing characteristics of edge computing include its dense geographical distribution, support for mobility, and proximity to the end users. Edge computing aims to provide location awareness, maintain low latency, support heterogeneity, and ameliorate Quality of Service (QoS) for real-time applications, such as transportation, industrial automation, networks of actuators and sensors, and real-time big data analytics. Services are deployed at edge devices, such as access points or set-top boxes. Table 1 highlights the differences between cloud computing and edge computing.

Current edge computing architectures are modeled as three-level hierarchies as illustrated in Figure 1. In these hierarchies, the smart things can connect to edge servers, these devices can interconnect with each other, and each edge device is also connected to the cloud.

The contributions of the article are as follows: (a) we present possible use cases for edge computing; (b) the study creates a classification of existing literature by devising a taxonomy; (c) identification of the key requirements to enable edge computing; and (d) highlighting the current challenges in realizing the vision of edge computing. The article also enables edge computing application engineers and service providers to leverage on the relevant features that can minimize communication and computation latencies while providing edge services to users. The identified requirements can serve as a guide for framework designers in incorporating specific features to execute efficiently the application in the edge computing paradigm. Similarly, these identified challenges highlight future research directions.

These contributions are provided in Sections 2–5, and concluding remarks are provided in Section 6.

2 Possible Use Cases

This section presents the application scenarios in the domain of edge computing. Edge computing can resolve several issues in various scenarios, such as real-time image processing, gaming, smart grid, smart traffic lights and connected vehicles, smart building control, and smart health environment.

2.1 Real-time Image Processing

Peter, a foreign visitor in Japan, wants to know the details on the food listed in a hotel menu written in Japanese. Thus, he takes a snapshot of the menu to process it using a character recognition application. The execution of the application is compute-intensive. Therefore, he is unable to run the application on his mobile phone. In this context, he can execute his application on the edge server available inside the hotel network. He registers with the edge server and migrates the application to the server. The results are then sent back to him upon completion of the task.

2.2 Gaming

Peter's son, who is fond of computer games, wants to play a high definition 3D game from his mobile phone on a train while traveling from Seoul to Pohang-si, South Korea. He finds an available edge server on the train that provides the services to passengers of the train to run their applications. The game engine is migrated to the edge server where the actual business logic of the game will run. The game interface is the only component that runs on the mobile device. The execution of the business logic on the edge server extends the battery life of the mobile device and enables game execution on the resource-constrained mobile device.

2.3 Smart Grid

Energy load balancing applications running on the smart meters and microgrids enable automatic switching to alternative energies, such as wind and solar energy after considering the lowest price, energy demands, and availability. The data generated by grid devices and sensors are processed by the edge collector at the edge network. Edge devices process the delay-sensitive data in the edge network and send the rest of the data to the cloud server. In a smart grid environment, the lowest tier stores temporary data whereas the highest tier stores the semi-permanent data.

2.4 Smart Transportation

Edge computing can also contribute to improving the functionalities and services provided by smart transportation. Smart lights that can serve as edge devices can take sensing information of the flashing lights of an ambulance from a video camera and consequently open lanes for the ambulance. Similarly, smart street lights take information from sensors to detect the presence of bikers and pedestrians and then turn the lights on or off depending on when movement is detected or traffic has passed.

2.5 Real-time Big Data Analytics

Big data processing has been a hot research area for computer science researchers. Real-time big data analytics in cloud computing is a challenging process because of the enormous volume of data and large WAN latency. Edge computing furnishes on-demand resources for processing huge amounts of real-time big data. Processing big data in the edge network reduces the traffic in the network and the workload on the cloud server [3]. Edge computing can complement the services provided by cloud computing. For instance, in a large-scale environment monitoring system, regional and local data can be collected and mined at edge servers, thereby enabling timely responses to emergency cases. Compute-intensive tasks, such as detailed analysis can be performed in the cloud server. Such edge computing-based big data analytics can be useful for the Internet of Things (IoT) and smart cities where sensor devices continuously generate enormous amounts of data [4] [5].

3 TAXONOMY

A taxonomy on the edge computing paradigm is shown in Figure 2. This taxonomy is classified broadly into the following attributes: a) access technologies, b) computing devices, c) computing paradigms, d) objectives, e) enabling technologies, f) computational hierarchy, and g) applications. The rest of the section discusses each attribute of the taxonomy by providing a definition of each term.

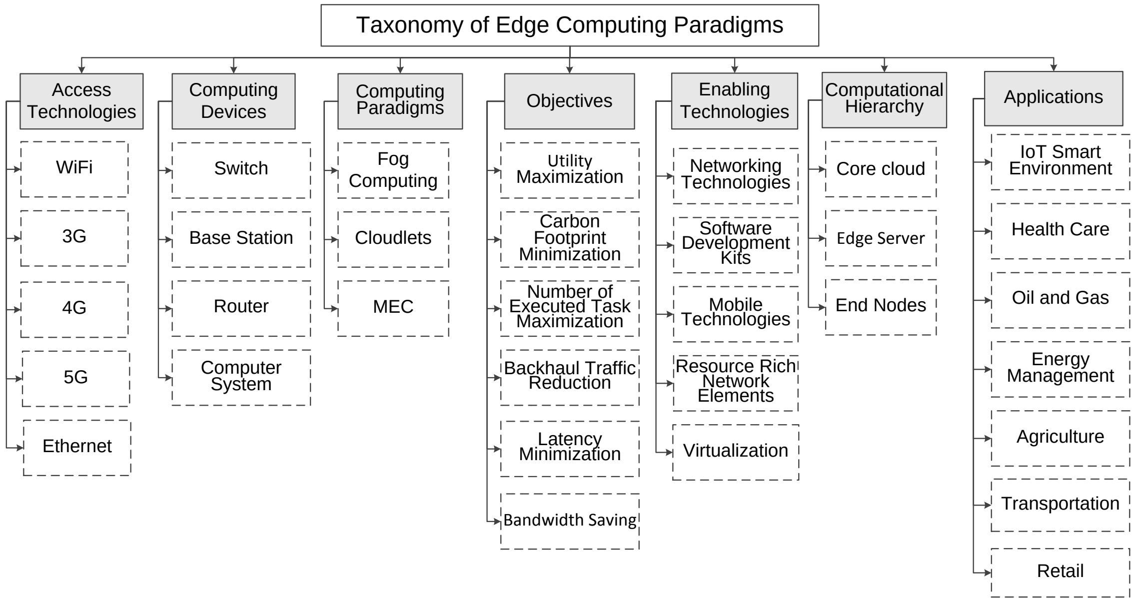

Fig. 2: Edge Computing Taxonomy

3.1 Access Technologies

The end user can access the edge services using both wireless and wired access technologies. Wired communication, i.e. Ethernet, is used to connect within an office network and in the server room to provide a link between short distances. Telecommunication network operators with the assistance of third party application developers can deploy new services for consumers and enterprise business services rapidly at the base station. Mobile users can subscribe to 3G, 4G, and 5G networks to access the applications [6]. Deploying computational offloading services at the WiFi access point to handle large compute-intensive application targeting few users is also possible. IoT devices, such as small embedded devices and smart bulbs, use WiFi networks to perform machine-to-machine (M2M) communication.

3.2 Computing Devices

Unlike traditional cloud data centers, edge computing devices are deployed closer to the end user in a distributed geographical location; however, such devices have certain weaknesses, such as computational, storage and network bandwidth. The service applications can be deployed in the telecoms network (base station), network devices (such as router and switches), and computer systems.

3.3 Computing Paradigms

The sole purpose of edge computing is to bring cloud computational resources and services within close proximity to the user. Three emerging paradigms that aim to achieve this goal are fog computing [7], cloudlet [8], and mobile edge computing (MEC) [9], [10]. Fog computing and MEC represent the commercialization of the idea of cloudlet. These paradigms provide an intermediate layer between the end device and the central cloud to reduce the network latency. In MEC, the virtualized server provides all the services and deploys them in the telecommunication network. In fog computing, the main computational services are available at network routers and switches.

3.4 Objectives

Edge computing brings the centralized computing platform of the core network to the edge network. The objectives of bringing the computational part from the core network to the edge network are as follows:

a. Utility maximization: The Quality of Experience (QoE) of users can be improved significantly. The end users are in the proximity of the service provider, and the location and context of the user can be read easily from the low-level network information, i.e. dynamic network environment. This information can be utilized by cloud providers to provide dynamic services.

b. Carbon footprint minimization: The distributed data center also consumes considerable energy and emits a huge amount of carbon to the surrounding atmosphere. Energy consumption and the amount of carbon emission is one of the challenges of the current data center. The joint resource allocation can minimize carbon footprint of the distributed nodes of the edge servers [11]. Tiny IoT devices require a considerably small amount of energy for computing and energy can be harvested from other renewable energy sources.

c. Number of executed task maximization: Efficient utilization of hardware resources in edge computing is one of the prime objectives. Distributed computing in the edge environment helps to allocate the task jointly to the edge nodes. The distributed task allocation not only reduces the idle state of the devices but also maximizes the average number of executed tasks on the device.

d. Backhaul traffic reduction: Edge computing can provide advanced caching services at the base stations to reduce backhaul traffic. Such edge computing architecture can enable the design of a cost-effective backhaul by transmitting voice and data from the radio cell site to an edge switch.

e. Latency minimization: Network latency is critical for many applications, such as augmented reality and m-game, if the application is required to connect to the centralized cloud server in the core network. The network latency of the core network is larger than that of the edge network [12]. The QoE of users improve if the computational part is deployed in the edge network to reduce the latency of delay-sensitive applications.

f. Bandwidth saving: Several applications, such as video streaming, consume a considerable amount of bandwidth. The edge network can apply mechanisms, such as web caching, to minimize bandwidth consumption of such applications.

3.5 Enabling Technologies

The realization of edge computing can be attributed to the recent development in the telecommunication and distributed computing. Key enablers of edge computing are presented as follows:

a. Networking technologies: The development of telecommunication networks has resulted in the increase in bandwidth capacity of wireless links. Mobile subscribers use wireless communication, such as 3G or 4G telecommunication network to access edge servers. These networking technologies provide highly reliable data communication between edge server and subscribers.

b. Software Development Kit (SDK): Many software provide SDK that consists of an application programming interface (API) that assists in the seamless integration of new services into the existing software package. It enhances the usability of the software package and fosters the development of new edge applications. Most SDKs are open source packages and are available to the developer.

c. Mobile Technologies: The development of the mobile technologies has become notable with the invention of smart phones and portable devices, such as a tablet. The number of mobile subscribers has been increasing exponentially in the recent years. Because majority of the users use mobile devices to access cloud services, mobile technologies are key enablers of the edge computing.

d. Resource Rich Network Elements: The recent development of computer hardware has enabled telecommunication networks to deploy a resource-rich network infrastructure.

e. Virtualization: Virtualization creates virtual or logical infrastructure above the same physical hardware. Virtualization can be both in hardware and software. Virtual resources are highly efficient in terms of usability of the physical resources and provide flexible and effective resource management. Edge computing services can be deployed in virtual machines and containers to increase efficiency in resource utilization.

3.6 Computational Hierarchy

Computational hierarchy shows that the execution in edge computing can be performed at different levels, such as core cloud, edge server, and end nodes. Although the goal of edge computing is to execute the compute-intensive delay-sensitive part of an application in the edge network, some applications in the edge server communicate with the core cloud for synchronization of data with the global application. It is also notable that the hierarchy represents the computing capacity of edge computing elements and their characteristics. The lowest element comprises end nodes that have less computational capability and subscribe mostly for the edge services. The intermediate node hosts the edge computing services closer to the end user. The edge server sometimes accesses the core cloud located at the far distance from the end nodes.

3.7 Applications

Application attributes define the range of edge computing applications hosted by edge servers. In IoT environment, the edge server hosts services for the intelligent data processing of the collected IoT data [4]. The IoT devices are controlled and managed directly by the edge server. In transportation, edge server can control the road traffic condition intelligently by acquiring data from the roadside camera and sensor devices. Edge server can host online vehicle parking for smart parking of the city dwellers. In healthcare, a person's health data can be collected from wearable devices, and the edge server can monitor the data for any emergency health assistance. Agriculture services, such as monitoring the environmental parameters (i.e., temperature, soil, and weather), can be deployed at the edge server. In retail, manufacturing and food processing can be automated with the help of edge computing [13].

4 Requirements for Edge Computing

Edge computing migrates the utility services hosted in the centralized cloud data centers to decentralized edge network devices. The edge network that relayed data packets to and from users can also act as a computing entity. In this manner, cloud services can be provided by the edge network with minimal latency. However, the edge computing paradigm must ensure several requirements, such as scalability to workload, to penetrate the IT utility market. Cloud computing services have captured most of the IT utility market. Edge computing must add incentives to the utility model of cloud computing. The listed requirements also apply to both cases of the standalone functioning of edge computing as a mini-cloud and interplay with the cloud resources. We list the requirements of cloud computing below.

4.1 Reliability

How will edge computing provide reliable services in instances of edge server failure and high user mobility?

Edge devices and architectures are often dynamic and evolving. Moreover, most devices connected to the edge have high mobility. Therefore, a failover plan must be presented for user services in case of dynamic changes in the foggy environment. Reliable edge computing ensures failover mechanisms in scenarios, such as the following: (a) failure of individual node/server/end user, (b) lack of network coverage because of the limitations of the access network and mobility of end devices, and (c) failure of edge network and service platform. In cloud computing, the reliability of services is ensured by duplication/replication of servers and data center nodes. However, due to the decentralized nature of edge networks, new nodes cannot be added immediately to the infrastructure. Protocols similar to the packet path finding need to be implemented to find services elsewhere in a failover plan. Moreover, client session data stored on edge devices need a consistent backup in case of device/network failure.

In general, three techniques can be applied to improve the reliability of edge computing: (a) check pointing, i.e., periodically saving the state of end devices and user services, (b) replication of edge servers and services in multiple geographical locations, and (c) rescheduling of failed tasks.

4.2 Scalability

How will edge computing scale to the number of mobile end users, diverse applications, and low bandwidth access networks?

Edge computing is scalable if it can penetrate into diverse networks, dynamic sensor, and end user devices, and provide a steady performance in case of a rapid increase in the number of end users and applications. Edge computing can be scaled by (a) adding a new point of services (geographic expansions), (b) adding new service nodes to existing points of service, and (c) utilizing cloud interplay. While cloud interplay can lead to far greater scalability for both entities, the standalone functioning of edge devices requires pre-planning because of the limited resources of edge devices. Edge should ensure redundancy of resources in case it acts as an independent mini-cloud. Intelligent management of the scalability options in different scenarios is critical to the functioning of edge networks.

4.3 Security

Which technologies can enable the sandboxing of user data on edge networks and secure edge devices from malicious users?

Security in edge computing has two aspects: (a) isolation of data paths and memory as multiple users share edge devices and (b) security of edge devices from malicious users because their addresses are advertised to end devices for service discovery. Application and session data from sensors and IoT devices are often cached on edge facilities. Moreover, multiple users contend for the limited resources of edge servers. While cloud computing provides isolation in the form of virtualized devices, edge computing needs to define sandboxes for user applications to ensure data isolation and monitor resource usage. A candidate for the fulfillment of security requirement in the edge is network function virtualization (NFV). NFV can be applied to network nodes to ensure security and isolation of user domains. However, NFV is still a new technology and not all network vendors adhere to NFV standards. Moreover, as network devices advertise their available resources, they become more vulnerable to rogue attacks.

4.4 Resource management

How will resources be federated by an edge network in the presence of multiple players and decentralized geo-dispersed resources?

Resource management comprises several activities, including resource allocation, reallocation, load balancing, and monitoring. The requirement of resource management is magnified in edge computing because of the decentralized and geo-dispersed nature of resources along with interplay of cloud computing. The cloud interplay also adds complexity to the resource usage price model. Moreover, network resources have to be monitored because they largely dictate the execution of applications over edge and cloud infrastructure. Resource management has to be performed within an edge network server and between edge networks, thereby making the task more complicated. There- fore, multi-level resource management techniques need to be adopted in edge computing. These resource management techniques can be applied independently and locally at the edge network level or in coordination with multiple edge networks and remote cloud providers. In this regard, resource discovery and synchronization requirements also arise because of the high mobility of user devices. NFV and software-defined networking (SDN) are potential technologies that can enable ease of management of edge resources.

4.5 Interoperability

Given the magnified heterogeneity of edge devices, protocols, and techniques, how will edge players interact and interoperate with each other?

Edge computing needs to provide interoperability and interactivity among multiple heterogeneous devices and service architectures. Heterogeneity is a major characteristic of edge computing, which arises from the difference of device architectures, communication protocols, and network configurations. However, the edge should support heterogeneity. In the long term, edge computing has to define interoperability standards for applications and data exchange between its players. Moreover, machine-to-machine communication is an integral part of edge computing, and standard protocols that apply to all participating devices need to be devised. Interoperability is also essential in attracting sensors and IoT devices towards edge computing paradigm.

4.6 Business model

How will fair share be decided among edge players in the act of distributed resource utilization?

Multiple players are involved in the business model of edge computing, i.e., ISPs who own edge servers and network devices, cloud service providers that may or may not interplay, and end users/devices who may act as both client and server for edge services in ad-hoc environment. Edge computing needs a business model similar to that of pay-as-you-go model of cloud computing. However, unlike cloud computing, the decentralized location and utilization of resources make the task difficult. To create a complete business model, how resources will be accounted and monitored need to be determined. Moreover, how edge players will divide the payment among themselves and how user devices who participate in the edge by advertising their spare resources will gain incentives will also have to be determined.

5 Open Research Challenges

In this section, we highlight some of the most important challenges that impede the success of edge computing paradigms. The discussion on these challenges provides research directions for further investigations in the edge computing paradigms.

5.1 Seamless Edge Execution Handover

Seamless edge execution handover is vital in enabling the uninterrupted migration of execution between different edge servers when the mobile user is on the move. The mobility between different edge networks disrupts the execution when a mobile device moves across two different network coverage areas. In edge computing, if a mobile user moves away from the connected edge computational platform, then the performance of the application degrades because of the increased communication latency. The local service providers also do not allow the mobile user access to and the use of its resources from outside of the network because of security reasons. This situation leads to a number of edge execution handovers. Seamless edge execution handover becomes a challenging task because of the nondeterministic mobility behavior of the user and intrinsic limitations of the wireless medium. The seamless edge execution handover is a vital research problem that needs to be addressed for the success of the edge computing paradigm. The seamless handover solutions from wireless networks, such as one reported in [14] can be applied in designing mechanisms for seamless edge execution migration.

5.2 Monitoring, Accounting, and Billing

Monitoring of edge computing resources, accounting, and billing are necessary to ensure satisfactory quality of service and to properly charge the user for the offered services by the edge computing service provider. However, monitoring, accounting, and billing require a sustainable business model for the edge computing service providers. Designing such business model is a challenging research perspective because of the mobile nature of the user and limited scope of the service. Usually, a user may access the edge services for a limited time, such as during lunch hours on the university cafe, this short-time access makes the business model relatively complex. Further, when the user moves and execution migrates from one edge platform to another, the division of the charges among the involved edge computing service providers raise new challenges for monitoring, accounting, and billing. These challenges require a business model that incorporates the various levels of charging granularity.

5.3 Lightweight Security and Privacy

Security and privacy are paramount issues that impede the success of edge computing [15]. The offloading of application from the end user device to the edge server requires the transfer of data from the mobile device to the edge server. The data are exposed to intruders through security breaches. Lightweight security and privacy mechanisms are necessary because of the battery-powered nature of the end users' devices. Furthermore, the security and privacy solutions should also be agile to support the aim of edge computing by reducing over execution time. The diversity of the environment and complexity of the problem make designing the reliable lightweight security solution a challenging research task.

5.4 Real-Time Data Processing

Providing best services at the edge of user network in a certain environment where load and data are growing at tremendous rates has become a real challenge. Recent statistics revealed that 20.8 billion devices would be connected to each other by the end of 2020[1]. This fast rate of connectivity among devices can cause different scalability issues in terms of functionality, administration, and load. The connectivity among a large number of devices results in a flood of data production that can make it difficult for the edge node to perform real-time processing. Moreover, the large amount of data can increase the load at the network edge, which can degrade the performance of applications that require low latency in terms of execution. To enable the scalable platform, research efforts have been carried out; however, these efforts are in their infancy, and considerable attention should be given to solving the scalability issues in the near future.

1. http://www.gartner.com/newsroom/id/3165317

5.5 Social Collaboration

To enable social collaboration among different edge nodes belonging to different vendors to achieve a common goal regarding efficient analytics has become difficult. The main hurdles in social collaboration are standardization and competition. In a market where different companies offer different edge devices as a service provider, social collaboration could not be made possible because of heterogeneous device architecture and contentious issues. The enabling of social collaboration among edge nodes is a tough task that must be solved to enable efficient data analytics.

6 Conclusion

Edge computing extends cloud computing by bringing the services closer to the end user at the network edge. Although edge and cloud computing paradigms use similar attributes (multi-tenancy and virtualization) and mechanisms, the extension is a challenging task that brings several new challenges. This study is conducted with the aim of exploring edge computing paradigm.

In this article, we highlighted the significance of edge computing by providing real-life scenarios requiring strict constraints on application response time. Moreover, we categorized and classified edge computing literature and devised a taxonomy based on relevant parameters. We identified and outlined the requirements that need to be met to enable edge computing. Furthermore, several challenges that remain to be addressed are discussed as future research directions. Finally, we concluded that although the edge computing paradigm has some distinguishing characteristics, such as dense geographical distribution, mobility support, location awareness, proximity, low latency, and context awareness, the paradigm is in its early stages of development. Hence, considerable attention should be given to addressing the challenges to facilitate the adoption of edge computing, which will be a core component of the future computing landscape.

ACKNOWLEDGEMENTS

This work is funded by the High Impact Research Grant from the University of Malaya under references UM.C/625/1/HIR/MOE/FCSIT/03 and RP012C-13AFR. This work is also supported by the Deanship of Scientific Research at King Saud University through Research group No. (RG 1435-051).

REFERENCES

[1] E. Ahmed and M. H. Rehmani, "Mobile edge computing: opportunities, solutions, and challenges," p. 59–63, 2017.
[2] M. Satyanarayanan, "The emergence of edge computing," *Computer*, vol. 50, no. 1, pp. 30–39, Jan 2017.
[3] N. Kamiyama, Y. Nakano, K. Shiomoto, G. Hasegawa, M. Murata, and H. Miyahara, "Analyzing effect of edge computing on reduction of web response time," in *2016 IEEE Global Communications Conference (GLOBECOM)*, Dec 2016, pp. 1–6.
[4] X. Sun and N. Ansari, "Edgeiot: Mobile edge computing for the internet of things," *IEEE Communications Magazine*, vol. 54, no. 12, pp. 22–29, December 2016.
[5] M. Satyanarayanan, P. Simoens, Y. Xiao, P. Pillai, Z. Chen, K. Ha, W. Hu, and B. Amos, "Edge analytics in the internet of things," *IEEE Pervasive Computing*, vol. 14, no. 2, pp. 24–31, Apr 2015.
[6] B. P. Rimal, D. P. Van, and M. Maier, "Mobile edge computing empowered fiber-wireless access networks in the 5g era," *IEEE Communications Magazine*, vol. 55, no. 2, pp. 192–200, February 2017.
[7] F. Bonomi, R. Milito, J. Zhu, and S. Addepalli, "Fog computing and its role in the internet of things," in *Proc. of the first edition of the MCC workshop on Mobile cloud computing*. ACM, 2012, pp. 13–16.
[8] U. Shaukat, E. Ahmed, Z. Anwar, and F. Xia, "Cloudlet deployment in local wireless networks: Motivation, architectures, applications, and open challenges," *Journal of Network and Computer Applications*, vol. 62, pp. 18–40, 2016.
[9] A. Ahmed and E. Ahmed, "A survey on mobile edge computing," in *Intelligent Systems and Control (ISCO), 2016 10th International Conference on*. IEEE, 2016, pp. 1–8.
[10] N. Kumar, S. Zeadally, and J. J. Rodrigues, "Vehicular delay-tolerant networks for smart grid data management using mobile edge computing," *IEEE Communications Magazine*, vol. 54, no. 10, pp. 60–66, 2016.
[11] Y. Mao, J. Zhang, S. H. Song, and K. B. Letaief, "Power-delay tradeoff in multi-user mobile-edge computing systems," in *2016 IEEE Global Communications Conference (GLOBECOM)*, Dec 2016, pp. 1–6.
[12] W. Hu, Y. Gao, K. Ha, J. Wang, B. Amos, Z. Chen, P. Pillai, and M. Satyanarayanan, "Quantifying the impact of edge computing on mobile applications," in *Proceedings of the 7th ACM SIGOPS Asia-Pacific Workshop on Systems*. ACM, 2016, p. 5.
[13] C. Liu, Y. Cao, Y. Luo, G. Chen, V. Vokkarane, Y. Ma, S. Chen, and P. Hou, "A new deep learning-based food recognition system for dietary assessment on an edge computing service infrastructure," *IEEE Transactions on Services Computing*, vol. PP, no. 99, pp. 1–1, 2017.
[14] S. Fu, J. Li, R. Li, and Y. Ji, "A game theory based vertical handoff scheme for wireless heterogeneous networks," in *Proceedings of 10th International Conference on Mobile Ad-hoc and Sensor Networks (MSN), 2014*. IEEE, 2014, pp. 220–227.
[15] I. Stojmenovic and S. Wen, "The fog computing paradigm: Scenarios and security issues," in *2014 Federated Conference on Computer Science and Information Systems*, Sept 2014, pp. 1–8.

Ejaz Ahmed (S'13, M'16) received his Ph.D. (Computer Science) from University of Malaya, Kuala Lumpur, Malaysia in 2016. He is an associate editor in *IEEE Communication Magazine*, *IEEE Access*, *KSII TIIS*, and *Elsevier JNCA*. His areas of interest include Mobile Cloud Computing, Mobile Edge Computing, Internet of Things, and Cognitive Radio Networks. He has successfully published his research works in more than forty international journals and conferences.

Arif Ahmed received M.Tech degree in Computer Science Engineering (CSE) from the National Institute of Technology Silchar, India in 2014, and B. Tech in Information Technology (2nd Rank in the class) from Assam University, Silchar, India in 2012. He worked as a visiting scientist at the Centre for Development of Advanced Computing(C-DAC), Mumbai, India from August 2014 to February 2015. After returning from Mumbai, he taught as an assistant professor(Contractual) in CSE department of National Institute of Technology Silchar, India(from July 2015 to May 2016).

Ibrar Yaqoob received his Ph.D. degree in Computer Science from the University of Malaya, Malaysia, in 2017. He worked as a researcher at Centre for Mobile Cloud Computing Research (C4MCCR), University of Malaya. His research experience spans over more than three and half years. He has published a number of research articles in refereed international journals and magazines. His numerous research articles are very famous and among the most downloaded in top journals. His research interests include big data, mobile cloud, the Internet of Things, cloud computing, and wireless networks.

Junaid Shuja works as an Assistant Professor at Comsats Institute of Information Technology (CIIT), Abbottabad. He completed his Ph.D. from University of Malaya in 2017 and MS from CIIT Abbottabad in 2012. His primary research interest is code offload in heterogeneous execution architectures. Other research interests encompass topics like energy efficient data center, sustainable cloud computing, and emerging edge computing paradigm.

Abdullah Gani (M'01, SM'12) (M'01, SM'12) is a full professor in the Department of Computer System and Technology, University of Malaya. He received his Bachelor's and Master's degrees from the University of Hull, United Kingdom, and his Ph.D. from the University of Sheffield, United Kingdom. He has vast teaching experience due to having worked in various educational institutions locally and abroad: schools, teaching college, the Ministry of Education, and universities. His interest in research started in 1983, when he was chosen to attend a Scientific Research course in RECSAM by the Ministry of Education, Malaysia. More than 150 academic papers have been published in conferences and respectable journals. He actively supervises many students at all levels of study — Bachelor, Master, and PhD. His research interests include self-organized systems, reinforcement learning, and wireless-related networks. He worked on mobile cloud computing with a High Impact Research Grant for the period of 2011–2016.

Muhammad Imran is an assistant professor in the College of Computer and Information Science, King Saud University. His research interests include mobile ad hoc and sensor networks, WBANs, IoT, M2M, multihop wireless networks, and fault-tolerant computing. He has published a number of research papers in peer reviewed international journals and conferences. His research is financially supported by several grants. He is serving as a Co-Editor-in-Chief for EAI Transactions on Pervasive Health and Technology. He also serves as an Associate Editor for the Wireless Communication and Mobile Computing Journal (Wiley), the Inderscience International Journal of Autonomous and Adaptive Communications Systems, Wireless Sensor Systems (IET), and the International Journal of Information Technology and Electrical Engineering. He has served/serves as a Guest Editor for IEEE Communications Magazine, IJAACS, and the International Journal of Distributed Sensor Networks. He has been involved in a number of conferences and workshops in various capacities, such as a Program Co-Chair, Track Chair/Co-Chair, and Technical Program Committee member. These include IEEE GLOBECOM, ICC, AINA, LCN, IWCMC, IFIP WWIC, and BWCCA. He has received a number of awards such as an Asia Pacific Advanced Network fellowship.

Muhammad Shoaib received his Ph.D. degree in Communication and Information System from Beijing University of Posts and Telecommunications, China (2010). His areas of research include video compression techniques, multilayer video coding, commercial Data Center facilities and IP packet based network, infrastructure and security. He is currently working as an Assistant Professor in the College of Computer and Information Sciences (Information Systems Department) in King Saud University.